CHAPTER ONE

Research and Analysis for Real-World Applications

CATHERINE M. BANKS

1.1 Introduction and Learning Objectives

Modeling and simulation (M&S) has made a name for itself as a discipline with its own body of knowledge, theory, and research methodology and as a tool for analysis and assessment. Significantly, M&S has attained this broad and meaningful position in a few short decades paralleling the technological advances of mainframe and desktop computers, the ever-expanding internet, and the omnipresent digital communications infrastructure. In 1999, the National Science Foundation (NSF) declared simulation the *third branch of science* (1). In a 2006 NSF report entitled, *Simulation-Based Engineering Science: Revolutionizing Engineering Science through Simulation*, a focused discussion ensued on the challenges facing the United States as a technological world leader. The report proffered four recommendations to ensure U.S. maintenance of a leadership role in M&S as a strategically critical technology. Foremost was the call for the NSF to "underwrite an effort to explore the possibility of initiating a sweeping

Handbook of Real-World Applications in Modeling and Simulation, First Edition.
Edited by John A. Sokolowski and Catherine M. Banks.

overhaul of our engineering educational system to reflect the multidisciplinary nature of modern engineering and to help students acquire the necessary M&S skills" (2). As noted in the *Introduction* of this text, a national movement is underway to ensure the role of M&S as a future technology. M&S education is a must for anyone who desires to be a part of that future technology. And it begins with acquiring an understanding of the four precepts on which M&S is premised: modeling, simulation, analysis, and visualization:

Modeling or creating an approximation of an event or a system.

Simulation or the modification of the model in which the simulation allows for repeated observation of the model as well as the methodology, development, verification and validation, and design of experiments.*

Visualization or the representation of data and the interface for the model as appropriate for conducting digital computer simulations providing an overview of interactive, real-time 3D computer graphics, and visual simulations using high level development tools.

Analysis of the findings or simulation output to draw conclusions, verify, and validate the research to make recommendations based on various simulations of the model as well as the inclusion of the constraints and requirements of engaging M&S as a way of declaring the limitations of the research.

Technological advancements have paved the way for new approaches to modeling, simulation, and visualization. Modeling now encompasses high degrees of complexity and holistic methods of data representation. Various levels of simulation capability allow for improved outputs and analysis of discrete and continuous events. State-of-the-art visualization allows for graphics that can represent details so intricate as to be found within a single shaft of hair (3). Once the domain of the engineering and computer science disciplines, M&S is now accepted as a multidisciplinary field of study capable of an expanding body of knowledge and user-friendly applications to address any research that calls for integrating quantitative and qualitative research methods and diverse modeling paradigms. M&S has moved far from static modeling; it is capable of representing the animate and the inanimate, and intangibles such as aspects of life, as well as life (human modeling) itself. Thus, M&S serves as a means of analyzing, assessing data to provide information for decision making, and/or teaching and training.

*Definitions of simulation vary and include a method for implementing a model over time, technique for testing, analysis, or training in which real-world systems are used, or where real-world and conceptual systems are reproduced by a model, an unobtrusive scientific method of inquiry involving experiments with a model rather than with the portion of reality that the model represents, a methodology for extracting information from a model by observing the behavior of the model as it is executed.

1.1.1 LEARNING OBJECTIVES

This chapter presents a broad look at M&S for research and analysis beginning with

- a contemporary look at M&S and its applied domains—background;
- a discussion of the theoretical foundation—M&S theory and toolbox;
- overview of research methods—research and analysis methodologies;
- case study—engaging a research methodology to analyze a problem;
- opportunities to test your understanding—exercises.

The primary learning objective of this chapter is for the reader to appreciate the breadth of opportunity M&S presents as a research and analysis tool. Inherent in the process of modeling is the required in-depth research of the event or system being modeled. This is because *models are driven by data* and so the data collection must be done with great accuracy. It can be said that a model is only as good as the data used to develop it. Specific to analysis is simulation development and the outputs that facilitate a variety of opportunities to review—also known as analyze—the intent of the modeling effort such as the analysis of a research question. And this is the case because M&S allows for a retesting of the hypothesis by allowing for iterations of the model's inputs. Thus, analysis can include determining attributes or time-sensitive changes to answer questions of a more predictive nature. For example, a model can replicate a protest scene with data representing protester attrition due to fear of arrest or fear of being accosted by counter-protest law enforcement (police). The common sense conclusion is that the police will eventually bring an end to the incident. But when? And how many policemen are needed to do this? What is the ratio of protestor to police needed to quash the protest? There are other factors that the model must represent: attributes of the environment, intent of the protest, the nature of the leadership, and the overall attitude of the protesters (pacifist or violent). These data inputs, and various iterations of the inputs via simulations, allow for potential outcomes or predictive assessments of the situation. Only M&S has the capability to redraw and retest the model and research question to provide specifics as to ratios of protesters to police or tipping points for change.

The secondary learning objective is to have a comprehensive grasp of the background, theory, paradigms, and domains (applications) of M&S. Putting these pieces together affords the M&S professional a holistic approach to the *developer—user* aspects of M&S. And importantly, it ensures the fundamental M&S protocols of verification and validation: Did we build the right thing (as to function and purpose)? Did we build it right (as to degree of correctness)?

1.2 Background

When did M&S make its first appearance? Is it a new field of study coupling engineering and computer science knowledge and skills, thus making it a cog in the wheel of technology evolution? Indeed, as a stand-alone discipline it is relatively young, but as a tool to examine, explore, and train it has existed for centuries offering much more than engineering and predating computer science. In fact, one can reasonably argue that the origins of modeling began in the ancient world in the form of live training as conducted by the Roman armies from c.500 BCE–1500 CE.

This was followed by an age of sophisticated art and complex architecture, c.1200–1600. Artists of the Renaissance made use of modeling as a means of conceptualizing their designs before beginning a project. One of the most ardent users of modeling was Leonardo da Vinci. His collection of work includes paintings, sculptures, building designs, advanced weaponry, flying machines, and anatomical studies. As an engineer, he made repeated use of modeling to test the design of many of his inventions and projects. His understanding of the system of systems engineering was futuristic for his period in history. Still, he determined that by understanding how each separate machine part functioned, he could modify it and combine it with other parts in different ways to improve existing machines or create new machines. da Vinci provided one of the first systematic explanations of how machines work and how the elements of machines can be combined. Through the following centuries the military continued to use modeling as a means of training with live exercises and with games that would resemble table-top exercises.

The technical origins of M&S go back to 1929 with the *Link Flight Simulator*. As a training tool, this simulator proved to be greatly cost cutting and it was eventually adopted by all branches of the military. Throughout the twentieth century, the Department of Defense laid claim to M&S by engaging simulation training in large-scale exercises. By 1983, the **Defense Advanced Research Projects Agency** (**DARPA**) had initiated simulator networking (**SIMNET)** with an emphasis on tactical team performance on the battlefield. Advancements in computer software and hardware as well as artificial intelligence and software agents hastened the pace of the maturation of M&S as a discipline and opened the way for M&S as a multidisciplinary application or tool for research and analysis. By the turn of the twentieth century, advanced academic programs enabled engineering students to graduate with a Doctor of Philosophy (PhD) in M&S.[†]

The advancement in technical capacity as well as research and development (R&D) allows M&S to have at its disposal enhanced capabilities for modeling, simulating, and analyzing complex phenomena. The technical features, coupled with a clearer understanding and application of the numerous modeling paradigms, allow the modeler (developer) to represent both **complicated systems** and **complex systems** and this is important because there are significant differences in these systems. To understand how they differ, a review of what comprises a system is needed.

[†] For a detailed discussion of the history of M&S, see Banks CM, What is modeling and simulation, in *Modeling and Simulation: A Multidisciplinary Approach* by Sokolowski JA and Banks CM, editors. New York: John Wiley and Sons, Inc., 2010.

A **system** is a construct or collection of different elements that together produce results not obtainable by the elements alone. The elements to a system vary ranging from people to hardware to facilities to political structures to documents—any and all of the things required to produce system-level qualities, properties, characteristics, functions, behavior, and performance. Recall it was da Vinci who recognized that the value of the system as a whole is the relationship among the parts. A modeler must understand both the *parts* and the *whole* of a system.

There are two types of systems: *discrete* in which the variables change instantaneously at separate points in time and *continuous* where the state variables change continuously with respect to time (these systems and the simulations used to represent them will be discussed in greater detail under Section 1.3.1). So how do complicated and complex systems differ? They diverge on the basis of the level of understanding of the system; for example, a human system may have few parts but it is complex because it is difficult to ascertain absolutes in the data as human systems data is organic and dynamic. Thus, one cannot predict the behavior of a human system with any certainty. On the other hand, a finite element model or physics-based model may be complicated due to its numerous parts, but it is not complex in that it is predictable and the data to model such a system is not soft or fuzzy (unpredictable).

For the purposes of analysis there are three principle approaches to the study of a system: (i) the actual system versus a model of the system, (ii) a physical versus mathematical representation of the system, and (iii) analytical solution versus simulation solution (which exercises the simulation for inputs in question to see how they affect the output measures of performance) (4). Because M&S provides various means to analyze a system and it has advanced to the level of representing both complicated and complex systems, M&S applications—the *user* side—have increased.

It is on the user side that M&S is growing as a means for **analysis** or the investigation of the model's behavior, **experimentation** that occurs when the behavior of the model changes under conditions that exceed the design boundaries of the model, and **training** for the development of knowledge, skills, and abilities that are gained through the operation of the system represented by the model. These three user goals are achieved via **stand-alone simulation**, which comports to the notion of experiential learning or training as one proceeds, or **integrated simulation** such as SIMNET, which can be used to enrich and support real systems. Note that in both stand-alone and integrated simulations the real system cannot be engaged. The reasons for this vary: the system may not be accessible, it may be dangerous to engage, it may be unacceptable to engage, or, simply, the system does not exist. Thus, the simulation imitates operations of these various systems and facilitates analysis, experimentation, and training, which would otherwise be unattainable.

Making use of both stand-alone and integrated simulations are numerous M&S subfields or domains. The **domain** of an M&S process refers to the subject area of the process. There are numerous domains and as M&S becomes fluent in user community content and functionality more domains will be engaged. This is evident in the fact that in a few short years the M&S community of domains has extended far beyond its familiar origins of military applications. As a result, M&S has extended deep roots into transportation, decision support, training and education (also known as game-based learning), medical simulation, homeland security simulation, M&S for the social sciences, and virtual environments as equally significant domains. Chapters 3–8 of this book provide detailed discussions of these prominent domains: *transportation, homeland security risk modeling, operational research, business process modeling, medical*, and *military*. Before delving into these chapters/individual domains, it would be beneficial to discuss the fundamentals of M&S: its theory and toolbox.

1.3 M&S Theory and Toolbox

Some in the engineering community consider M&S to be an infrastructure discipline necessary to support integration of the partial knowledge of other disciplines needed in applications. Its robust theory is based on dynamic systems, computer science, and ontology of the domain, that is, representing knowledge as a set of concepts within M&S and the relationships between those concepts.[‡] It can be used to reason about the entities within M&S and/or it may be used to describe M&S. It is the theory and ontology that characterize M&S as distinct in relation to other disciplines. Theoretically, the concept of the model allows for it to be the physical, mathematical, or logical representation of a system, entity, phenomenon, or process. It should be remembered that a system might be complicated or complex and unengaged for a variety of reasons. Models are stand-ins for those systems. And it is this substitute model that will serve for the purposes of research and analysis of the system. So how does one go about developing a model so as to abstract from reality a description of the system? Begin by acknowledging the can do's and limitations of modeling.

It would be foolhardy to think that one could develop a model representing all aspects of the system being studied as it would be timely, expensive, and complex—perhaps impossible. Rather, the model should be developed as simply as possible representing only the system aspects that affect the system performance being investigated in the model, perhaps depicting the system at some point of abstraction. The intent of the model is to represent the system as reliably as possible.

There are two approaches to model development: **physical**, such as a scale model of a car to study the effects of weight on velocity, and **notional**, which is basically a set of mathematical equations or logic statements that describes the behavior of the system.

It is the **simulation** that describes the behavior of the system using either a mathematical model or a symbolic model (5). Simulation can imitate the operation of a real-world process or system over a period of time (6).

[‡]Portions of this chapter are based on Banks CM, Introduction to modeling and simulation, in *Modeling and Simulation Fundamentals: Theoretical Underpinnings and Practical Domains* by Sokolowski JA and Banks CM, editors. New York: John Wiley and Sons, Inc., 2010.

With the execution of a simulation comes the **run** and/or **trial**, which is a single execution; a series of executions is called the **exercise**. Thus, run/trial and exercise are similar in meaning, but they imply different uses of the simulation runs. There are four phases to M&S development and each phase has a technology used to support its development.

Model phase makes use of **modeling technologies** in developing a model of a system. The model will include data, knowledge, and insight about the system. The type of system being modeled will determine the model construct and modeling paradigm.

Code phase engages **development technologies**; this is because simulation is a software design project. Computer code must be written that represents algorithmically the mathematical statements and logical constructs of the model. Obviously, code phase is heavily drawn from the modelers software engineering expertise.

Execute phase proceeds via **computational technologies**, which, for simple simulations, means implementing the simulation code on a personal computer and for complex simulations implementing the simulation code in a distributed, multiprocessor, or multicomputer environment where the different processing units are interconnected over a high speed computer network. Modelers need to understand these underpinnings of computer architectures, networks, and distributed computing methodologies.

Analyze phase is conducted with the use of **data/informational technologies** to produce the desired performance information that was the original focus of the research. Models premised on variability and uncertainty are likely to apply probability and statistics in the analysis.

With model development underway, the modeler must also be concerned with the simulation's attributes, of which there are three: fidelity, validity, and resolution.

Fidelity conveys how closely the model or the simulation matches reality. High fidelity signals the model or simulation that closely matches or behaves like the real system. This is difficult to achieve because models can not capture every aspect of a system. It should be remembered that models development should center around representing only the aspects of a system that are to be investigated. Low fidelity is tolerated with the less significant aspects of the system.

Validity conveys three constructs of accuracy of the model: (i) reality—how closely the model matches reality; (ii) representation—some aspects are represented and some are not; (iii) requirements—different levels of fidelity required for different applications.

Resolution (or granularity) is the degree of detail with which the real world is simulated. Obviously, more detail yields higher resolution. Simulations can go from low to high resolution.

Scale (or level) is the size of the overall scenario or event the simulation represents. Thus, the larger the system, the larger the scale of the simulation.

1.3.1 SIMULATION PARADIGMS

M&S has three primary simulation paradigms that the modeler can chose to best represent his given system.[§]

Monte Carlo simulation randomly samples values from each input variable distribution and uses that sample to calculate the model's output. This process of random sampling is repeated until there is a sense of how the output varies given the random input values. Monte Carlo simulation models system behavior using probabilities.

Continuous simulation allows for system variables that are *continuous functions of time*. With continuous simulation *time* is the independent variable, and the system variables evolve as time progresses. Thus, this type of simulation would need to make use of differential equations in developing the model.

Discrete event simulation allows for system variables that are *discrete functions in time*. The discrete functions result in system variables that change only at *distinct instants of time*. The changes are associated with an occurence of a system event. As a result, discrete event simulations advance in time from one event to the next event. This simulation paradigm adheres to queuing theory models.

Another aspect of simulation design is the simulation mode or type. These modes vary in operator and environment. For example, the model and simulation can include real people doing real things, or real people operating in unreal or simulated environments, or real people making inputs into simulations that execute those inputs by simulated people. There are three modes: live, virtual, constructive, or a combination thereof.

Live simulation involves real people operating real systems. This simulation strives to be as close as possible to real use and it often involves real equipment or systems. The purpose of live simulation training is to provide a meaningful and useful experience for the trainee.

Virtual simulation involves real people operating in simulated systems. These systems are recreated with simulators and they are designed to immerse the user in a realistic environment. This type of training is designed to provide experiential learning.

Constructive simulation involves real people making inputs into a simulation that carries out those inputs with simulated people operating in

[§]For a detailed discussion on simulation, see Sokolowski JA, The practice of modeling and simulation: Tools of the trade, in *Modeling and Simulation for Medical and Health Sciences* by Sokolowski JA and Banks CM, editors. New York: John Wiley and Sons, Inc., 2011.

simulated systems. As real people provide directives or inputs, activity begins within the simulation. There are no virtual environments or simulators and the systems are operated by nonparticipants.

1.3.2 TYPES OF MODELING

There are numerous types of modeling within the M&S toolbox that range from the mathematical to the hybrid:

Physics-based modeling is solidly grounded in mathematics. A physics-based model is a mathematical model where the model equations are derived from basic physical principles. Unique to physics-based models is the fact that the physics equations are models themselves in that many physics-based models are not truly things, but intangibles; hence, they are representations of phenomena.

Finite element modeling (FEM) is the method for modeling large or complicated objects by decomposing these elements into a set of small elements and then modeling the small elements. This type of modeling is widely used for engineering simulation, particularly, mechanical and aerospace engineering. These subdisciplines conduct research that requires structural analysis or fluid dynamics problems. FEM facilitates the decomposition of a large object into a set of smaller objects labeled *elements*.

Data-based modeling results from models based on data describing represented aspects of the subject of the model. Model development begins with advanced research or data collection, which is used in simulations. Data sources for this type of modeling can include actual field experience via the real-world or real system, operational testing and evaluation of a real system, other simulations of the system, qualitative and quantitative research, as well as best guesses from subject matter experts (SMEs). The model is developed with the view that the system is exercised under varying conditions with varying inputs. As the outputs unfold, their results are recorded and tabulated so as to review appropriate responses whenever similar conditions and inputs are present in the model.

Aggregate modeling facilitates a number of smaller objects and actions represented in a combined, or aggregated, manner. Aggregate models are used most commonly when the focus of the M&S study is on aggregate performance. The model can also scale and number represented entities that are large and can compromise the time required to conduct a simulation. These models are most often used in constructive models; they are not physics-based models.

Hybrid modeling entails combining more than one modeling paradigm. This type of modeling is becoming common practice among model developers. Hybrid modeling makes use of several modeling methods; however, they are disadvantaged in that composing several different types of models correctly is a difficult process.

1.3.3 MODELING APPLICATIONS

Just as there are differing simulation paradigms, so too there are a variety of modeling applications. Below are some of the most widely used applications.**

Agent-based modeling focuses on these analysis, or agents, and the sequence of actions and interactions of the agents over a period of time. Agents may represent people, organizations, countries—any type of social actor. These actors may act in parallel, may be heterogeneous, and may learn from their actions. Each agent responds to the prior action of one or more of the other agents or the environment in the model (or system). This, in turn, produces an extended and often emergent sequence of behaviors, which can be analyzed for different things. The action and inaction agents are regarded as variables. This type of modeling is intrinsically social in that the actions and characteristics of the agents are influenced by the actions and characteristics of the other agents in the social system. To develop an agent-based model, the modeler must first define the basic behavior of an agent. Typically this is done using a series of simple rules that the agent must follow. These rules help describe the fundamental goals that the agent is trying to achieve. This methodology is an effective way to simulate complex social behaviors through the application of relatively simple rules that each agent follows.

Game theory modeling is associated with rational decision making among players, be it on a (sports) field of play or political competitions. This type of modeling serves as a tool to study the interactions of individuals (players) in various contexts. The model allows the analyst an opportunity to observe interactions between and among the players so that each, as an individual decision maker, can determine what he deems to be the best course of action. The model also facilitates the ability to analyze strategic behavior where there are conflicts of interest. This type of model output allows the analyst to categorize interpersonal behavior within a spectrum of cooperative or competitive. There are two types of game theory: *cooperative game theory*, whereby the players can communicate to form winning coalitions, and *noncooperative game theory*, which focuses more on the individual and his handicap of not knowing what the other players will do.

System dynamics deals with the simulation of interactions between objects in dynamic (active) systems combining theory, methods, and philosophy to analyze the behavior of systems regardless of the nature of the system. This is because system dynamics provides a common foundation that can be applied wherever there is a need to understand and influence how things change through time. Thus, system dynamics modeling lends itself to macro level representations of a system that can address the interdependence of the actors, events, or variables within the system. The modeling consists of two components: (i) the causal loop diagram describing how various system

**For a detailed discussion on modeling applications, see Sokolowski JA and Banks CM, *Modeling and Simulation for Analyzing Global Events*. New York: John Wiley and Sons, Inc., 2009.

variables relate to one another from a cause and effect standpoint (a drawback to this type of diagramming is that it does not allow for the accumulation of variable totals) and (ii) stocks and flows that are used to overcome the causal loop drawback. (*Stocks* are accumulation points within the system that allow one to measure the amount of a variable at any given time; *flows* are inputs and outputs of stocks that represent the rate of change of the stock).

Behavioral modeling captures human activity in which individual or group behaviors are derived from the psychological or social aspects of humans. This modeling hosts a variety of approaches; most prevalent are the computational approaches found in social network models and multiagent systems. This modeling facilitates the incorporation of socially dependent aspects of behavior that occur when multiple individuals are together. Chapter 2 of this book provides a detailed discussion and case study on behavioral modeling.

Social network modeling supports understanding the connections between and among people. It also allows for an explanation of a flow of information, or the spread of contagion, or the identification of individuals who are isolated from the group (also known as outliers). Grouping patterns (algorithms) allow for the separation of large networks into smaller subsets. This draws closer the members who share identifying marks or attributes. Integral to social network modeling is the analysis of patterns of relationships among the members in the social system to include varying levels of the analysis such as person to person or groups to groups. This type of modeling relies on actors that are concrete and observable; thus, the relationships within the social network are usually social or cultural. These types of relationships bind together the actors or entities making them interdependent entities.

1.4 Research and Analysis Methodologies

Since a model is a representation or characterization of a system, the data used to develop the model and the inputs to the simulation are in effect what validate and verify the construct of the model and the content of the simulation outputs. As such, the underpinnings of the verification and validation of the model are tied directly to the research and the comprehension of (i) what is being modeled—the system and (ii) the hypothesis that is to be tested—the research question. Thus, conceptual model development is a good place to start as it initiates model development with ideas and suggestions as to what needs to be modeled. As research ensues model development will refine. If a research question is already in place, the process is expedited to some extent with a more focused approach to the research. How does one approach conducting research for model development? By first recognizing the two primary forms of research and how the data or information they yield is applied to the model.[††]

There are basically two types of data and each hosts different methods of data mining (research): qualitative and quantitative. The debate over which is more

[††]For additional case study development examples, see Sokolowski JA and Banks CM, *Modeling and Simulation for Analyzing Global Events*. New York: John Wiley and Sons, Inc., 2009.

sound is premised on two opposing notions: *all research ultimately has a qualitative grounding* versus *there is no such thing as qualitative data because everything is either 1 or 0*. How do they differ? The data reviewed in **qualitative research** is subjective and open for interpretation by the analyst; thus, some qualitative data such as the words, pictures, and behaviors can be assessed differently by two analysts. It is clear that the training and thinking of the analyst affects his interpretation of the data. On the other hand, **quantitative research** involves the analysis of numerical data, and therefore it is objective. The value of a whole number is the same no matter who assesses its value. It is important to note these differences as they have a clear effect on model development.

Qualitative research begins with gathering information by direct observation, analysis of documents and sources, and interviews. This lends itself to smaller, focused samples and it makes understanding or interpretation of human behavior possible. Qualitative research is primarily exploratory and often remains open-ended, not conclusive. It can yield information that is very detailed and often it is data that is difficult to categorize because it is individualized. Therefore, the modeler must recognize that qualitative research is going to yield large volumes of data that is difficult to generalize, and it is data that is subject to interpretation. Modelers are challenged with incorporating qualitative data into the model, but it is essential that they acquire this skill. Quantitative data, on the other hand, is much friendlier to model development because this data expresses quantity, number, and measurement. The research is a systematic investigation. Because this data engages mathematical models and theories, quantitative research is prominent in the sciences (biology, chemistry, physics, mathematics, psychology, and engineering). Thus, quantitative data examines events through the numerical representation, and/or statistical analysis. The goal of this approach is to quantify behavior by measuring variables on which they hinge and intersect, comparing the variables, and pointing out correlations. That brings us back to the debate and the differences between these two research methods and data yields.

Recall the two opposing notions: the notion *all research ultimately has a qualitative grounding* contends that quantitative research obscures reality by omitting the *human-ness* of an event as reflected in nonnumeric, nonmeasurable factors. Conversely, the notion *there is no such thing as qualitative data because everything is either 1 or 0* is favored by those who engage quantitative data to *legitimize* research. There is, however, a third path to take in the form of a combination of quantitative and qualitative data gathering referred to as **mixed-methods research**. This combined mix of research and data facilitate a summarization of large bodies of subjective or qualitative data and a generalization based on objective or quantitative projections. And this meshes perfectly with M&S because M&S can accommodate this hybrid approach to investigation.

With model in hand, the simulations will no doubt result in the generation of large quantities of **system performance data**. The data is stored in a computer system as large arrays of numbers. Converting those numbers into meaningful information that describes the behavior of the system is the first step in **analysis**. The analysis is in a sense an interpretation of the simulation's output of the modeled system over a specified period of time with a given set of conditions and inputs. Simply, analysis takes place to draw conclusions, verify, and validate the research, and make recommendations based on various simulations of the model. In sum, research is needed to form the **inputs** to a model; they can be qualitative in nature as attributes needed to make the model match a social environment or quantitative in nature lending themselves to mathematical representations. The **outputs** to the simulation are the behaviors of the model. It is the behaviors of the model that are analyzed as to better understand and/or explain a system. The outputs also serve as a means to test the viability of a hypothesis and select the best representation of the system.

Qualitative data adds significantly to models in which human behavior is being assessed, making it integral to model development in many M&S domains. For example, military M&S is now expanding upon its traditional usage of M&S by incorporating new approaches to research and analysis toward simulations heavily infused with aspects of human behavior modeling. By incorporating fuzzy or squishy data yielded from qualitative research, these models now represent human behavior in a realistic, holistic, and relevant manner. The case study is another example of integrating both forms of data. It also outlines a methodical approach to M&S progression beginning with formulating the research question to concluding the analysis.

Case Study: A Methodology for M&S Project Progression

The following is a suggested methodology for an M&S project progression to be used in the analysis of a system to observe (explain) its behavior, answer a research question, or serve as proof of a hypothesis. There are six steps to completing the project.

Step One: Developing the Research Question and Methodology

Models allow for observation of systems—information gathering. They can replicate a system to ascertain the cause of a dilemma—problem assessment. And they can characterize variables affecting a system in an effort to answer *What-if's* about the system's behavior—question resolution. Thus, it is important to understand just what information, problem, or question is in need of investigation. This will provide focus for the research and determine what modeling paradigm would best serve the investigation. For example, a public medical administrator wants to predict the health effects of obesity-related illnesses in his community. The investigation is certainly worth exploring as obesity is fast becoming a national cause for concern relative to the overall health of the country; it will become a major expense to public health. No doubt this issue will have an impact on the medical community as the percentage of patients with obesity-related problems will exponentially outnumber the increase in trained medical professionals; medical subfields will be directly and indirectly affected, healthcare centers will be affected, and these centers may or may not be prepared to manage the patient load. These factors represent only the administration of medical care aspects of the issue. What about the patient side? Is there a demographic profile to construct such as who are the emergent obese patients, who are the urgent obese patients, and who are the at-risk obese patients? In determining the profile is there a propensity for any given health issues for this population, that is, genetic or cultural? If so, will those needs be cared for by the present medical community? If not, then how will the medical community address this deficiency?

How will the state and local governments address these needs? From this example of brainstorming the topic, the modeler develops a list of questions that will serve to define the research agenda. To narrow it even further, a research question(s) is needed to focus the study.

For this topic, the public medical administrator wants to ascertain the effects of obesity-related illnesses on the hospitals in his community. He is interested in how that demographic profile and its propensity for specific obesity-related health issues will impact the current capacity of the public health facilities.[‡‡] He wants to make the case for better public health awareness (education) and execute deliberate adjustments to the public health system as a means to serving the emergent patient case-load and providing proactive care for urgent and at-risk patients. From this preliminary brainstorming activity, conceptual model and project can be developed.

For example, this case study is tasked with characterizing the demographic profile of the community in its current stage, forecasting obesity-related illnesses vis-à-vis that specific population, representing the current structure of the medical community, and proffering solutions for adjustments to that community as a means of providing medical care and education for this population. The simulations developed for this model can explore different outcomes on the basis of changes to the health conditions and behaviors of the population. The research will center on the effects of obesity on current public health capacity, future public health capacity given no positive change on the population, and future public health capacity with positive change due to personal care, education, and modifications in governmental policy for public health administration. Therefore, a viable research question could be this: *Does current state for healthcare education and public health for obese populations parallel the current trends in rising obesity-related health concerns vis-à-vis present public health capacity?* Answering this question provides a means to proffer *What-ifs: What is the tipping point in which present pubic healthcare capacity can no longer provide adequate patient care for obesity-related illnesses? What changes in policy need to be implemented*? *Given the current trends in healthcare education and training where does the medical community stand?* This single research question facilitates the development of numerous simulations on a policy by policy basis or an overall comprehensive look. These simulations can also provide specific outputs to measure needed changes in education for various populations: the general public, medical professionals (reeducation to meet current demands), rising medical professionals (new medical curriculums). It can also evaluate modifications to public healthcare administration and capacity. With the research question in hand, Step Two can begin.

Step Two: Research

Integral to this case study is a review of many aspects of society (demographics, population healthcare risk factors and problem-propensity, population socioeconomics, government capacity (its ability to provide services), public healthcare capacity) all of which affect the future of public health care from the patient/medical/governmental perspectives. The research will yield a baseline model of a complex system of that society. From that baseline model simulations can be developed to answer the *What-ifs* listed above. The modeler's goal is to develop a mathematical formula derived from the research for a predictive model to assess needed changes in governmental policy to address future healthcare challenges of obesity-related illnesses.

[‡‡]Capacity is defined as the ability of public healthcare to serve/treat its community; capacity includes facilities (hospitals, centers, and offices), adequate equipment at the facilities, personnel (administrative and medical professionals), public education (workshops, seminars), professional development for medical personnel, and IP infrastructure (for internal use as in patient databases and external use as in communications).

Step Three: Mapping Data

Mapping data is essentially the translation of a qualitative description of the factors capturing the context of the event into a numerical representation. Quantitative methods that make use of numbers (such as the number of deaths from specific illnesses or the number of trained medical professionals in a specific subfield) do not provide the contextual significance of these data, which could lead to erroneous causation representations. This method provides for a qualitative to quantitative mapping of factors that maintains their context in relation to the studied event. Each qualitative factor is independently evaluated and scored by one or more SMEs for its relevance and significance in describing the behavior of the system. The scoring may be done using a Likert scale rating that represents the value of each factor's contribution to the model. As an alternative, a Bayesian assessment may be used to capture this value. Either method maintains the context of how the individual factors influence the modeled event.

In describing these complex events there may be scores of factors that must be considered. Possessing a large number of them may make it difficult for a modeler to map the multiple relationships in a concise manner. Grouping them into categories will facilitate this mapping. Index values representing each grouping can then be developed and used in the modeling process without the loss of contextual relevance. For example, research for the case study outlined above will yield factors that can be grouped or binned into indices such as

1. *Polity index*: A measure of the government's ability to provide fiscal support for public health ranging from patient care to public education to professional development of medical professionals.
2. *Population demographics societal index*: A measure of the population from the perspective *obtaining* healthcare to include age, gender, ethnicity, and socioeconomics.
3. *Population demographics medical index*: A measure of the population from the perspective *providing* healthcare to include age, gender, ethnicity, personal health history, family health history, propensity for medical problems, poor health habits, height, weight, family dynamics.
4. *Public healthcare capacity index*: A measure of the medical community's ability to provide health care to include facilities (hospitals, centers, and offices), adequate equipment at the facilities, personnel (administrative and medical professionals), public education (workshops, seminars), professional development for medical personnel, and IP infrastructure (for internal use as in patient databases and external use as in communications).
5. *Obesity-related illnesses trends index*: The present prominent and predicted prominent illnesses.

Values for the indices are the prerogative of the modeler and the researcher (SME); however, using a simple scale to rate each factor may be the most effective way of representing the data. For example, rating each factor on a Likert scale of -5 to 5 with $+5$ representing a high positive influence and -5 representing a high negative influence will serve to characterize the variables that comprise the indices. Those ratings can then be average under the appropriate index to produce a final index value which can be normalized between 0 and 1. These indices will be used to seed key parameters in the model. This mapping of qualitative data preserves the context behind what is actually going on in various indices and it provides a means to quantitatively represent this context.

Step Four: Selecting a Modeling Paradigm and Executing the Model

Choosing the appropriate modeling paradigm will depend on the specific purpose of the model. Systems engineering view of complex systems provides a means for depicting complex systems and displaying the relationships among variables through the use of system dynamics modeling. A system dynamics approach facilitates exploring a macro level view of system behavior where sufficient data is available to calibrate the model to existing conditions to serve answering additional research questions (*What-ifs*). System dynamics modeling provides a graphical means to represent the different indices and their respective variables in causal and correlative relationships within the system. Developing an actual system dynamics model is beyond the scope of this chapter; however, the modeler should recognize that the exercise of creating a system dynamics view of an event or system necessitates an in-depth translation of a conceptual representation of the event into a more detailed mathematical relationship among the factors. With these four steps completed, analysis of the model's outputs takes place.

Step Five: Responding to the Research Question

This step begins with an analysis of the model's outputs, the various simulation findings, and comparisons of the findings, all in an effort to respond to the research question. Because this research was undertaken using M&S, numerous answers to the research question can be achieved. This is the beauty of simulation; it can accommodate changes to what are the direct and indirect variables of the system. As such, the answer to the research question can include proffering suggestions for future policy or actions and these predictions have a high degree of reliability because they are mathematically sound solutions.

Step Six: Model Validation

Step six is the process of comparing simulation results derived from a model against the real-world system that the model is meant to represent. It falls to the judgment of the developer and/or user to determine if the simulation results are close enough to the real system. If the answer is yes, then the model is considered a valid representation of the real system or process. Determining it to be *close enough* is obviously a subjective term that must be interpreted by the person employing the model.

Validating models of physical phenomena is generally straight forward since the laws that govern those systems are usually well known and mathematically precise and often it is just a matter of matching it to a 100% predictable outcome. On the other hand, comparing the simulation results against a real-world system such as public healthcare is more problematic. When modeling real-world systems are in existence, the modeler can compare the results of the simulation with the current information to judge validity. The modeler can then attempt to extrapolate that model to investigate not only *What-ifs*, but also what could happen in the future. This process is known as **predictive modeling** and it is a means to answer a specific question or set of questions.[§§]

SUMMARY

In 1998, the Institute of Industrial Engineers (IIE) listed the advantages and disadvantages of using M&S (7). The list includes in part *the ability to choose correctly, the ability to explore possibilities, and the ability to diagnose a problem.*

[§§] For more information on predictive modeling and validation, see Balci O, Verification, validation, and testing, in *Handbook of Simulation: Principles, Advances, Applications, and Practice*, by Banks J, editor. New York: John Wiley & Sons, 1998, pp. 335–393.

From these advantages it is easy to see *why* M&S is becoming a modality for research and analysis of real-world applications. This chapter is aimed at providing the *how* to M&S.

The overall theme of the chapter emphasizes the breadth of opportunity M&S presents as a research and analysis tool. For modelers, this is accomplished when the M&S underpinnings such as theory, simulation paradigms, types of models, and modeling applications are implemented. The chapter case study included a six-step methodology to crafting a research question, developing the model, and validating the output. Modelers who appreciate the *why* and the *how* of M&S often achieve a holistic approach to research and analysis of real-world events.

KEY TERMS

Analysis: findings or simulation output to draw conclusions, verify, and validate the research
Constructive Simulation: involves real people making inputs into a simulation that carries out those inputs by simulated people operating in simulated systems
Continuous Simulation: allows for system variables that are *continuous functions of time* making *time* the independent variable and the system variables evolve as time progresses
Discrete Event Simulation: allows for system variables that are *discrete functions in time*
Domain: in an M&S process refers to the subject area of the process
Fidelity: speaks to how closely the model or the simulation matches reality
Integrated Simulation: experiential learning used to enrich and support real systems
Live Simulation: involves real people operating real systems
Modeling: creating an approximation of an event or a system
Monte Carlo Simulation: randomly samples values from each input variable distribution and uses that sample to calculate the model's output to provide a sense of how the output varies given the random input values using probabilities
Qualitative Research: subjective and open for interpretation by the analyst includes words, pictures, and behaviors that can be assessed differently by analysts
Quantitative Research: involves the analysis of numerical data and it is objective in nature
Simulation: modification of the model in which simulation allows for the repeated observation of the model
Stand-Alone Simulation: comports to the notion of experiential learning or train as one proceeds
System: construct or collection of different elements that together produce results not obtainable by the elements alone
Validity: conveys three constructs of accuracy of the model: reality, representation, and requirements

Virtual Simulation: involves real people operating in simulated systems
Visualization: representation of data and the interface for the model

EXERCISES

1.1 Define modeling, simulation, and visualization.
1.2 Define analysis and explain its importance.
1.3 Explain the differences between complicated and complex systems.
1.4 Explain the differences between stand-alone simulation and integrated simulation.
1.5 Explain the two approaches to model development.
1.6 Explain the four phases of M&S development.
1.7 How do fidelity and validity differ?
1.8 Explain the simulation paradigms.
1.9 Explain the types of modeling.
1.10 Explain the types of modeling applications.
1.11 How do qualitative and quantitative data and research approaches differ?
1.12 Discuss the six-step methodology to M&S project progression.

REFERENCES

1. 2006. Simulation-based engineering science: Report of the National Science Foundation Blue Ribbon Panel. Available at http://www.nsf.gov/pubs/reports/sbes_final_report.pdf. Accessed 2011 June.
2. 2006. Simulation-based engineering science: Report of the National Science Foundation on Revolutionizing Engineering Science through Simulation. Available at http://www.nsf.gov/pubs/reports/sbes_final_report.pdf. Panel 2006. Accessed 2011 June.
3. Sokolowski JA, Banks CM. A proposed approach to modeling and simulation education for the medical and health sciences. In: Proceedings of the 2010 summer simulation conference; 2010 July 11–15; Ottawa, Canada: 2010.
4. Law AM, Kelton WD. *Simulation, Modeling, and Analysis*. 4th ed. New York: McGraw-Hill, Inc; 2006.
5. Fishwick PA. *Simulation Model Design and Execution: Building Digital Worlds*. New Jersey: Prentice Hall; 1995.
6. Banks J, editor. *Handbook of Simulation: Principles, Methodology, Advances, Applications, and Practice*. New York: John Wiley & Sons, Inc; 1998.
7. Colwell RR. Complexity and connectivity: a new cartography for science and engineering. Remarks from the American Geophysical Union's Fall Meeting; San Francisco: 1999.