SAND97-2607C
SAND--97-2607C
CONF-980610--

Draft of a Paper for a Special Session on
Uncertainty Analysis in Computational Heat Transfer
(K-12 and K-20 Committees)
1998 ASME/AIAA Joint Thermophysics and Heat Transfer Conference
Albuquerque, New Mexico
June 15-18, 1998

DRAFT

Uncertainty and Error in Computational Simulations

William L. Oberkampf† Kathleen V. Diegert§
Kenneth F. Alvin†† and Brian M. Rutherford§
Sandia National Laboratories
Albuquerque, NM 87185-0825

Abstract

The present paper addresses the question: "What are the general classes of uncertainty and error sources in complex, computational simulations?" This is the first step of a two step process to develop a general methodology for quantitatively estimating the global modeling and simulation uncertainty in computational modeling and simulation. The second step is to develop a general mathematical procedure for representing, combining and propagating all of the individual sources through the simulation. We develop a comprehensive view of the general phases of modeling and simulation. The phases proposed are: conceptual modeling of the physical system, mathematical modeling of the system, discretization of the mathematical model, computer programming of the discrete model, numerical solution of the model, and interpretation of the results. This new view is built upon combining phases recognized in the disciplines of operations research and numerical solution methods for partial differential equations. The characteristics and activities of each of these phases is discussed in general, but examples are given for the fields of computational fluid dynamics and heat transfer. We argue that a clear distinction should be made between uncertainty and error that can arise in each of these phases. We believe that the present definitions for uncertainty and error are inadequate and, therefore, we propose comprehensive definitions for these terms. Specific classes of uncertainty and error sources are then defined that can occur in each phase of modeling and simulation. The numerical sources of error considered apply regardless of whether the discretization procedure is based on finite elements, finite volumes, or finite differences. To better explain the broad types of sources of uncertainty and error, and the utility of our categorization, we discuss a coupled-physics example simulation. We then discuss how the methodological ideas developed can be applied in the modeling and simulation of a weapon in an abnormal environment. Specifically, we consider the conceptual problem of a damaged weapon in an aircraft crash and fuel fire environment.

†Aerosciences Dept., Org. 9115, Associate Fellow, AIAA.
§Statistics and Human Factors Dept., Org. 12323
††Structural Dynamics Dept., Org. 9234.

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the U. S. Department of Energy under contract No. DE-AC04-94AL85000.

DISCLAIMER

I. Introduction

During the last several years there has been an increasing level of attention to critically assessing the accuracy and credibility of computational simulations. Although the fields of fluid dynamics, heat transfer, and structural mechanics do have same level of interest in this topic, the general trend is a welcomed one. It shows that computational simulation is maturing from a research activity to a useful tool that impacts the design of engineered systems of all types. The primary method of assessing the accuracy of simulations has been to compare computational predictions with experimental data. This is known as the process of validation of computational simulations. One of the difficulties of experimental validations, however, is the continually increasing cost and time required to conduct these experiments. With the rapidly decreasing cost of computer power, there is great economic and competitive pressure to conduct simulations with fewer comparisons with experiments. Terminology such as "virtual prototyping" and "virtual testing" are now being used to describe the numerical simulation of "testing" of new hardware and even entire systems. An additional difficulty with experimental validation is that certain types of validation experiments can not be physically or safely conducted. Various examples are: matching the aerothermodynamic environment of an atmospheric reentry vehicle in a wind tunnel; structural failure of the containment vessel of a nuclear power plant; failure of a bridge or dam during an earthquake; and exposure of a nuclear weapon to an accident environment such as an aircraft fuel fire. Physical events such as these cannot be conducted with full fidelity for the purpose of validation, thus assessing the accuracy of simulations for these events is crucial.

The issues, methodologies and terminology for assessing the accuracy of complex simulations are being discussed and debated in the literature in a very wide variety of engineering disciplines [1-9]. These topics are closely related to code verification and validation (V&V). It is our view that all of these topics, at a high level, can be considered as part of the field of uncertainty estimation. Uncertainty estimation has its roots in probability theory and statistics and has primary application in areas such as quality control in manufacturing process, estimation of experimental uncertainty, and probabilistic risk assessment of large systems. Many of the same mathematical tools of uncertainty estimation can be used to represent different types of uncertainty in modeling and computational simulation.

Uncertainty due to different sources in computational simulations are also now being addressed by a number of researchers [10-15]. Examples include; numerical solution error estimation in finite element simulations, use of Richardson's extrapolation method for estimating grid convergence errors, uncertainty due to different turbulence models in computational fluid dynamics, and Monte Carlo estimation of structural response due to stochastic uncertainty in material properties. All of these are examples of different types of contributors to uncertainty in various computational simulations. Some of these can be categorized as modeling uncertainties and others as numerical solution errors. The computational simulation literature has done little to categorize the different types of sources of uncertainty and error. Indeed, there is little discussion, much less agreement, as to what is included in uncertainty, and how that is related to error.

We believe the estimation of total modeling and simulation uncertainty can be divided into two parts. First is the identification of all the possible types or classes of sources of uncertainty and error. Second is a general mathematical procedure for combining, integrating and propagating individual sources of uncertainty and error through the entire modeling and simulation process. At present, neither of these steps has been developed in any general sense. The field that has probably made the most progress on both of these steps deals with the thermal-hydraulic analyses for safety of nuclear power plants. For example, many failure scenarios and event tree analyses for these type systems have been constructed, along with probabilistic assessment of events and their consequences. We believe, however, there are three short comings to this previous work. First, there has not been clear delineation of the classes of sources of uncertainty, or a distinction between uncertainty and error in the modeling and simulation. Second, very little effort has been devoted to the impact on total system uncertainty due to uncertainty in the model itself, also referred to as model form uncertainty. However, recent work by Draper [16] and Lasky [17] suggests an increasing focus on this issue. Third, the vast majority of this work is based on statistical, or probabilistic, mathematical representations of uncertainty. The primary emphasis has been on uncertainty distributions in input parameters, initial conditions, and boundary conditions. We believe, however, that non-probabilistic mathematical representations may be more appropriate when the uncertainty derives from lack of knowledge or errors, such as numerical solution errors.

The present paper deals with the first part of the estimation of total modeling and simulation uncertainty; identification of all possible classes of sources of uncertainty and error. We begin by developing a new structure of the general phases of modeling and simulation. The phases proposed are: conceptual modeling of the physical system, mathematical modeling of the system, discretization of the mathematical model, computer programming of the discrete model, numerical solution of the model, and interpretation of the results. This new view is built upon combining phases recognized in the disciplines of operations research and the numerical solution of partial differential equations. Characteristics and activities of each of these phases is discussed with regard to a variety of disciplines in computational mechanics and thermal sciences.

Building on this structure, we argue that a clear distinction should be made between uncertainty and error that can arise in each of these phases. We believe that the existing distinctions between uncertainty and error are inadequate and, as a result, we propose comprehensive definitions for these terms. Specific classes of uncertainty and error sources are then defined that can occur in each phase of modeling and simulation. The present discussion generally shows how uncertainties and errors in one phase might propagate to subsequent phases, but this paper does not address the technical issues involving propagation of uncertainty and error. The numerical sources of error considered apply regardless of whether the discretization procedure is based on finite elements, finite volumes, or finite differences. We also propose a term, modeling and simulation dubiety, to represent the level of doubt in the total simulation caused by both uncertainty and error. To better explain the broad types of sources of uncertainty and error, and the utility of our categorization, we discuss a coupled-physics example simulation. We discuss how the methodological ideas developed can be applied in the modeling and simulation of a weapon in an abnormal environment. We consider the conceptual problem of a damaged weapon in an aircraft crash and fuel fire environment. This example considers the widest possible range of a fully coupled thermal-material response simulation with regard to detonation safety of the weapon.

2. Modeling and Simulation

Before we review the literature, we will define what we mean by modeling and simulation. We will use broad definitions for these terms because the issues we address in this paper will cover a wide range of computational mechanics, thermal sciences, and physics. We use the definition of *model* given by Neelamkavil [18]: "A model is a simplified representation of a system (or process or theory) intended to enhance our ability to understand, predict, and possibly control the behavior of the system." By *modeling* we mean the construction or improvement of a model. Different types of models will be defined in later sections. We also use Neelamkavil's definition of *simulation* [18]: "A simulation is the process of imitating (appearance, effect) important aspects of the behavior of the system." In other words, simulation is the exercise of a model. Here we are specifically interested in the exercise of computer models, i. e., computer codes built from mathematical models.

2.1 Review of the Literature

In attempting to identify all possible classes of sources of uncertainty and error we spent significant time reviewing a broad range of literature in modeling and simulation. Modeling and simulation is conducted in essentially every technical discipline. The operations research community, because it deals with the widest range of system types and processes, has developed many of the general principles and procedures for modeling and simulation. Researchers in these fields have made significant progress defining and categorizing the various phases of modeling and simulation [19-22]. The areas of emphasis in OR include problem entity definition, conceptual model definition, data and information quality, and how simulation results are intended to aid in decision making. From an engineering perspective, however, many feel this work is extraneous because it does not deal with solving partial differential equations. We have found the OR research very helpful in providing a constructive philosophical approach for identifying uncertainty sources and errors and some of the basic terminology. In addition, the OR literature has developed the fundamental principles for verification, validation, and accreditation for modeling and simulation, although we do not discuss this topic in depth.

In 1979, the Society for Computer Simulation Technical Committee on Model Credibility developed a diagram identifying the primary phases and activities of modeling and simulation [23] (Fig. 1). The diagram shows that analysis is used to construct a conceptual model of reality. Programming converts the conceptual model into a computerized model. Then computer simulation is used to predict reality. Also shown in the diagram are the activities of model qualification, model verification, and model validation. Although simple and direct, the diagram clearly shows the relationship of two key phases of modeling and simulation to each other, and to reality, i. e., the system or process being considered. When complex engineering systems or physical processes are considered, the diagram ignores some major activities, specifically the solution of the partial differential equations describing the system.

Through the 1980's Sargent [24, 25] made some improvements to this view of modeling and simulation, but no fundamental changes to the ideas in Fig. 1 were made. His major contributions were primarily in the areas of procedures for verification and validation of models and simulations. Nance [26] and Balci [27] also made a significant expansion to the phases of modeling and simulation. Fig. 2 shows their concept of the life cycle of a simulation study. Major phases added to the earlier description were System and Objectives Definition, Communicative Models, and Simulation Results. Communicative Models were described as "a model representation which can be communicated to other humans, can be judged or compared against the system and the study objective by more than one human" [26]. These three additions helped clarify important phases in modeling and simulation.

As can be seen from Figs. 1 and 2, the emphasis in these descriptions is in assessing the credibility of the model and simulation, and in its improvement. The primary methods to accomplish this are verification, validation, qualification and testing of nearly all phases of the process. The emphasis in the present work is on *estimation* of total modeling and simulation uncertainty with limited experimental data. With this emphasis, Figs. 1 and 2 lack two key features. First, the inability to clearly identify where sources of uncertainty and error might *originate*. Second, the notion of *propagation* of different types of uncertainties and errors through the modeling and simulation process.

In the computational fluid dynamics literature Mehta [7] appears to be the only one to have incorporated some of the OR concepts of modeling and simulation. Figure 3 shows Mehta's diagram for the phases of modeling and simulation. As can be seen, it has features similar to Figs. 1 and 2, but now the emphasis is on sources of uncertainty and total modeling and simulation uncertainty. Although specific types of sources of uncertainty are not identified in Fig. 3, Mehta does describe several sources. He lists general uncertainties in the fluid dynamics models and in the computational analysis.

A large number of other investigators have investigated sources of uncertainty and error, but they have not made direct associations to the various phases of the modeling and simulation process. Several of these investigators will be referred to in Section 3: Sources of Uncertainty and Error.

2.2 Phases of Modeling and Simulation

Figure 4 shows our representation of the phases of modeling and simulation appropriate to systems or processes analyzed by the numerical solution of PDE's. The initial phase is called the definition of the physical system. This phase also includes the specification of the requirements or objectives of the modeling and simulation, as described by Nance [26] and Balci [27]. The physical system can be either an existing system or process, or a proposed system or process, for example a proposed design. The physical system could be as simple as laminar flow through a pipe, or it could be as complex as fire spread through an aircraft cabin.

<u>Conceptual Modeling</u> The conceptual model phase determines what scenarios of physical events, or sequence of events, will be considered and what types of coupling of different physical processes will be considered. These determinations are based on the requirements determined in the first phase. During this phase, no mathematical equations are written, but the fundamental assumptions of the events and physics are made. Only conceptual issues are considered, with heavy emphasis on determining all possible factors that could possibly affect the requirements set for the modeling and simulation. An additional important feature of this phase is that all possible physics-couplings are listed that may influence the results, even if it is considered unlikely that they will be considered later on in the analysis. This is critical because if a possible physics coupling is not considered in this phase, it can not be resurrected later in the process. This feature is similar to the fault-tree structure in probabilistic risk assessment of high consequence systems, such as in nuclear reactor safety analyses. Even if a certain sequence of events is considered extremely remote, it should still be included as a possible event sequence in the fault-tree. Whether or not the event sequence will eventually be analyzed is not a factor in including it in the conceptual modeling phase.

<u>Mathematical Modeling</u> The next phase is the mathematical modeling phase. During this phase the precise mathematical, i. e., analytical, statement of the problem, or series of event-tree-driven problems, to be solved is developed. Any complex mathematical model of a problem, or physical system, is actually composed of many mathematical submodels. The complexity of the models depends on the physical complexity of each phenomena being considered, the number of physical

phenomena considered, and the level of coupling of different types of physics. The mathematical model formulated in this phase is considered to be the complete specification of all of the partial differential equations for all elements of the system. For example, if the problem being addressed is a fluids-structure interaction, then all of the coupled fluid-structures PDE's must be specified, along with any material property changes either the fluid or structure might undergo because of their interaction. The integral form of the equations could also be considered, but this type formulation will not be addressed in the present discussion. Along with the PDE statement of the mathematical model, all of the appropriate initial and boundary values, and the required auxiliary, or closure, models must be specified for the physical events considered.

Our emphasis on comprehensiveness in the mathematical model should not be confused with a model's attempt to represent physical complexity. The predictive power of a model depends on its correct identification of the dominant controlling factors and their influence, *not* upon its completeness. A model of limited, but known, applicability is generally more useful than a more complete model. This dictum of engineering seems to be forgotten in modern times because of rapidly increasing computing power. The clear tendency, seen in all fields of engineering, is to use more complex models and then "beat it to death" with the computer. Examples are abundant, but just to mention a few: use of Navier-Stokes equations to compute the lift on a streamlined body at low angle of attack; use of time iterative Navier-Stokes equations to compute attached supersonic flow over a vehicle; and use of finite elements to compute stresses in thin shells.

An additional point concerning incompleteness of models should be made. Any mathematical model, regardless of its physical level of detail, is *by definition* a simplification of reality. Any real-world system, or even individual physical processes, contain phenomena that are not represented in the model. Statements such as "full physics simulations" can only be considered as marketing jargon. Our point was succinctly stated by George Box [28]: "All models are wrong, some are useful."

Discretization of the Model The next phase is the conversion of the PDE form of the mathematical model into a discrete, or numerical, model. This phase takes into account the conversion of the mathematics from a calculus problem to an arithmetic problem. In the discretization phase, all of the spatial and temporal differencing methods, discretization of the boundary conditions, discretization of the geometric boundaries, and grid generation methods are specified in analytical form. In other words, algorithms and methods are prescribed in mathematically discrete form, but the spatial and temporal step sizes are not specified. This step focuses on the conversion from continuum mathematics to discrete mathematics, not on numerical solution issues. We strongly believe that the continuum mathematical model and the discrete model should be separately represented in the phases of modeling and simulation [29]. This phase deals with questions such as consistency of the discrete equations with the PDE's, mathematical singularities, and differences in zones of influence between the continuum and discrete systems.

Programming of the Discrete Model The next phase, which is common to all computer modeling, is the computer programming phase. This phase converts the algorithms and solution procedures defined in the previous phase into a computer program. This phase has probably achieved the highest level of maturity because of many years of programming development and software quality assurance efforts [30, 31]. These efforts have made a significant impact in areas such as commercial graphics, mathematics, and accounting software, telephone circuit switching software, and flight control systems. Little impact, however, has been made in corporate and university developed software developed for computational fluid dynamics, solid dynamics, and heat transfer simulations.

Numerical Solution of the Discrete Model The next phase is the numerical solution of the programmed computer model. In this phase, individual numerical solutions are obtained. This phase should be thought of as the most specific of all phases of modeling and simulation. In this phase there are *no quantities* left arithmetically undefined or continuous. For example, grid spacing is uniquely defined, all parameters such as Reynolds number and chemical reaction rate constants are specified, and time and space exist only at points. If multiple computational solutions are required for the analysis, as is commonly the case, then the numerical solution would not be unique, but would mean many solutions. Consider, for example, a conduction heat transfer analysis where the thermal conductivity is specified by some probability distribution. Then hundreds or thousands of Monte Carlo solutions may be required to address the question posed in the definition of the problem.

Interpretation of Results The final phase concerns the interpretation of computational results. This phase involves determining the methods for presentation of computed results into a form that is usable by a human. This phase can also be described as the construction of continuous functions based on the discrete solutions obtained in the previous phase. Here the continuum mathematics formulated in the mathematical modeling phase is approximately reconstructed. This phase is specifically called out because of the sophistication of the software that is being developed to comprehend modern complex simulations. This area includes graphical visualization of results, animation of results, use of sound for improved interpretation, and the analysts "going into the solution space" using virtual reality. Some may argue that this phase is simply "post-processing" of the computational data. We believe, however, this description does not do justice to the rapidly growing importance of this area. In addition, by referring to this phase as interpretation of results, we are able to include types of errors that are not simply due to the modeling and simulation of the system, but to the conclusions drawn from the simulation results. These topics will be discussed in the next section.

3. Sources of Uncertainty and Error

We will now discuss the types of uncertainties and errors that are associated with each phase of modeling and simulation. As one might suspect, developing the definition for each of the phases was not done independently from developing the ideas for types of uncertainties and errors. Essentially all of the individual sources of uncertainty and error described below have been pointed out by researchers in the past. Some, like computer round-off error, are very well understood, even to the point that most computational analysts do not make note of it. Others are poorly understood or characterized; for example, should a deficiency be treated as an uncertainty or an error. In the sections that follow we first develop comprehensive definitions for uncertainty and error that are appropriate for modeling and simulation. Second, we describe a hypothetical modeling and simulation sample problem which will be used as an example during the description of uncertainties and errors. Third, we describe a general framework for classes, or types, of uncertainties and errors for each phase of modeling and simulation. In addition, we use the example problem to give specific examples for each class of uncertainties and errors. We strongly believe that a more comprehensive taxonomy for uncertainties and errors must be developed in order to mathematically estimate total simulation uncertainty for complex systems.

3.1 Definitions of Uncertainty and Error

As we attempted to identify general types of uncertainty and error, we found ourselves asking more and more fundamental questions as to what is the distinction between uncertainty and error. Although the meaning of these terms seems to be intuitive, upon careful thought it is found their meaning is not precise, or very context specific. We observed that the majority of text books and research papers do not define what they mean by uncertainty and error. Only a few authors carefully define uncertainty and error, but their definitions are in the restricted context of their subject. The most developed definition or understanding of uncertainty is in regard to experimental measurements. Although this is helpful, we require definitions that apply to the much broader topic of modeling and simulation.

We define uncertainty as a *potential* deficiency in any phase or activity of the modeling process that is due to *lack of knowledge*. The first feature which our definition stresses is "potential", meaning that the deficiency may or may not occur. In other words, there may be no deficiency, say in the prediction of some event, even though there is a lack of knowledge. Whether the deficiency occurs or not is most commonly represented by some type probability distribution of occurrence. The second key feature of uncertainty is that its fundamental cause is incomplete information. Following Klir [32], incomplete information can be caused by vagueness, nonspecificity, or dissonance. By vagueness we mean lack of precise definition, unclearness and indistinctness. Nonspecificity refers to the variety of alternatives in a given situation that are left unspecified. Dissonance refers to the disagreement resulting from the attempt to classify an element of a given set into two or more disjoint subsets of interest. Since the cause of uncertainty is partial knowledge, increasing the knowledge base can reduce the uncertainty. When uncertainty is reduced by an action, such as observing, performing an experiment, or receiving a message, that action is a source of information. The amount of information obtained by the action is measured by the resulting reduction in uncertainty. This concept of information is called "uncertainty-based information." Examples of this are: improving the accuracy of prediction of heat flux in a steel bar by improving the knowledge of the thermal conductivity of the bar in the predictive model; improving the prediction of the convective heat transfer rate in turbulent flow by improving the turbulence model; and improving the prediction accuracy for melting of structure in a open-pool fuel fire by improved knowledge of the atmospheric winds.

We define error as a *recognizable* deficiency in any phase or activity of modeling and simulation that is *not* due to lack of knowledge. Our definition stresses the feature that the deficiency is identifiable or knowable upon examination, that is, the deficiency is not determined by lack of knowledge. By this we mean that there is an agreed-upon approach which is considered to be more accurate. If divergence from the correct or more accurate approach is pointed out, the divergence is either corrected or allowed to remain. This implies a segregation of error types; error can be either *intentional* or *unintentional*. Examples of intentional errors are: finite precision arithmetic in a computer; physical approximations made to simplify the modeling of a physical process; a specified level of iterative convergence of a numerical scheme; conversion of the governing PDE's into discrete equations. When the analyst introduces these intentional errors into the modeling or simulation process, there is typically some idea of the magnitude of the error introduced. Unintentional errors are blunders, or mistakes. That is, the analyst intended to do one thing in the modeling and simulation. but, for example, due to human error, did another. There are no straightforward means to estimate or bound the contribution of unintentional errors. Sometimes the unintentional error is capable of being discovered by the person who committed it; e.g., a double check of coding reveals that two digits have been reversed. Sometimes blunders are due to inadequate human interactions, and can only be resolved by communication. For example, one person misunderstood the required input format for a code written by another person. In this case, a rigorous review process by both individuals should uncover the error.

3.2 Description of the Example Problem

Consider the coupled thermal-material analysis of a weapon in an open-pool fuel fire environment. Assume that the weapon may be damaged, but the level of damage is unknown. This example would be characteristic of a weapon carried by an aircraft, that crashed during take-off or landing. Assume that the type of weapon is known, but no other information about the weapon before the accident is known. The weapon contains high explosive that is normally a solid and it has an integrated electrical-mechanical arming, fusing, and firing system. And finally, assume that the purpose of this analysis is to compute a probabilistic estimate of whether the high explosive will detonate. Stated somewhat differently, compute the probabilistic risk assessment of the detonation safety of the weapon in this crash and burn scenario.

The purpose of our example is to point out the myriad of factors and possibilities that enter into a complex, real world, engineering simulation. We will only list aspects of this example in our discussion of uncertainties and errors in the conceptual modeling and mathematical modeling phases. These phases require the resolution of many specific probabilistic issues. Although we make no computations here, the magnitude of the computing effort required should become clear.

3.3 Conceptual Modeling Uncertainties

From the description of the conceptual modeling phase given in Section 2, we believe that the dominant "deficiency" is uncertainty, as opposed to "error". Deficiencies can occur in any of the phases of modeling and simulation, but the credibility of each phase is primarily limited either by uncertainties or errors. Conceptual modeling uncertainties arise in the formulation of the analysis of the event, or process, and in the lack of knowledge of the event. Figure 5 shows the two types of uncertainties associated with conceptual modeling; scenario abstraction and lack of system knowledge. By scenario abstraction we mean the determination of all possible physical events, or sequence of events, that may affect the goals of the analysis. For relatively simple systems, such as fluid flow not interacting with any structures or materials, scenario abstraction can be straight forward. For complex engineered systems exposed to a variety of interacting factors, scenario abstraction is a mammoth undertaking. The best example we can give for how this should be done for complex systems is the probabilistic safety assessment of nuclear power plants. As the many-branched event tree is constructed for complex scenarios, the probability of occurrence of certain events becomes extremely low. Typically little analysis effort is expended on these extraordinarily rare possibilities. If one is dealing with very high consequence systems, however, these extremely improbable scenarios must be examined. Not including or recognizing these branches of the event tree can cause substantial loss in the credibility of the modeling and simulation.

The second class of uncertainty listed, lack of system knowledge, refers to uncertainties that are primarily due to limited information about the system. This class clearly affects and interacts with the scenario abstraction effort, but here we stress lack of information for a branch of the tree rather than the possible existence of the branch. Two important examples for this class of uncertainty should be mentioned. First is the lack of knowledge of the initial state of key elements of the system. If it is a complex engineered system then knowledge of the factors, such as the following, becomes important: was the system incorrectly manufactured or assembled, how well was the system maintained, and was the system damaged in the past and not recorded. Second, is lack of knowledge of future conditions impacting the system. Examples of these are atmospheric environmental conditions and human interaction with the system during the event. These are examples where it is not possible to reduce lack of knowledge, and reduce the uncertainty, by improved sampling of past events. However, the uncertainty can sometimes be reduced by certain action taken with respect to the system that limits or further defines the state of key elements of the system. Often these are policy or procedural decisions.

For the example problem of a weapon in a fire we list a number of scenario abstraction and lack of system knowledge sources of uncertainty. Rather than attempt to list all of the possibilities, we just give an indication of the types of uncertainties that should be be considered in this phase.

- Lack of information concerning the weapon before the accident
- Manufacturing variability of components and the complete system
- Manufacturing, assembly, and handling errors affecting the system
- Maintenance of the weapon and components that effect the state of the weapon before the crash
- Structural damage due to the crash before the start of the fire
- Structural and electrical damage to the arming, fusing and firing system before the start of the fire
- Detonation sensitivity of the explosive affected by age
- Number of weapons carried on-board the aircraft that affect individual weapons
- Adjacent weapon detonating
- Uncertainty in damaged geometry
- Uncertainty in impact area characteristics (e. g., water, trees, city)
- Uncertainty in material properties of components and subsystems
- Fuel source and quantity
- Wind speed, temperature and other environmental conditions
- Fire intensity and duration
- Uncertainty in emissivity of surfaces before and during the fire
- Atmospheric electrical source of energy to the arming, fusing and firing system (e.g.,lightning)
- Undesirable effects of accident response teams to the accident (e.g., additional damage)
- Unintended effects of accident response teams to the accident (e.g., introduction of foams or electrical power to the crash site)
- Accident response hampered by unsafe state of the weapons
- Inadequate use of expert opinion in scenario abstraction (e.g., insufficient diversity)

3.4 Mathematical Modeling Uncertainties

Mathematical modeling contains both uncertainties and errors, but we believe that uncertainties are typically more important than errors in this phase. (Note that for the remainder of the paper when we refer to "errors" we will only be referring to *intentional* errors, unless otherwise stated.) Uncertainties and errors that occur in this phase arise from three mathematical sources (Fig. 5): the continuum equations for conservation equations of mass, momentum, and energy; all of the auxiliary equations which supplement the conservation equation; and all of the initial and boundary conditions required to solve the PDE's. The predominant uncertainties that occur in mathematical modeling are those due to limited knowledge of the actual physics involved, or inadequate knowledge to represent elements in known physics. The primary errors are due to mathematically representing the physics in more simplified form that is appropriate for the results required from the modeling and simulation. Both of these together are sometimes referred to as "model form errors" or "model structural errors".

Examples of uncertainties that occur in the conservation equations are: limited knowledge of the physics of multiphase flow, limited knowledge of turbulent reacting flow, and uncertainty if a boundary layer will be laminar or turbulent. Auxiliary physical equations in the mathematical model are equations, such as, expressions for thermal conductivity, turbulence models, and chemical reaction equations. Examples of uncertainties in these models are poorly known probability distributions of material properties due to manufacturing variability and unreliable turbulence models. It may be argued that accuracy turbulence models should be considered as errors instead of uncertainties. This is based on the argument that the accuracy of turbulence models could be ordered, e. g., algebraic models, two equation models, and Reynolds stress models. In a general sense, this ordering could be accepted, but for individual flow fields there is no guarantee that any one model will be better than any other model. Examples of uncertainties in initial and boundary conditions are: inaccurately known initial temperature distribution in a solid, imprecisely known geometry of materials because of manufacturing and assembly variances, and poorly known wind conditions in a pool fire.

Errors in mathematical modeling can also be identified. Some examples are; assumption that a flow field can be modeled as a two-dimensional flow when three-dimensional effects are important, assumption of a steady flow when the flow is actually unsteady, assumption of continuum fluid mechanics when non-continuum effects are important, and the assumption of a rigid boundary when the boundary in flexible. It is observed that all of these examples are of the character that physical modeling approximations were made to simplify the mathematical model and the subsequent solution.

For the example problem of a weapon in a fire we list a number of mathematical modeling uncertainties and errors:

- Use of 2D models for 3D problems
- Use of steady state models for non-steady state solutions
- Poorly known fluid dynamic turbulence models coupled with combusting flow
- Uncertainty in coupled mechanics - interaction of structural and thermal and possibly electrical
- Uncertain thermodynamic and transport properties of all materials
- Inaccurate probability distributions of the geometry of components because of small sample sizes
- Use of different submodels - crack propagation, joints, thermal conductivity
- Possibility of missed interactions at low levels of details, especially in damaged node
- Errors in coupled solution procedures; ex: structural, thermal, structural, thermal interact via forces
- Inadequate temporal coupling between thermal and structural mechanics coupling
- Use of transport, thermodynamic, and material properties outside the range of validity
- Ignoring electrical resistance heating in components due to unexpected activation of a power supply
- Insufficient level of spatial and temporal modeling for physics involved
- Inaccurately known thermal contact resistances due to both manufacturing, assembly, and crash damage
- Incomplete modeling of interaction of non-linearities (e. g., turbulence and combustion)
- Inaccurate interpolation of transport, thermodynamic, and material properties
- Inappropriate statistical models to represent non-deterministic phenomena

3.5 Discretization Errors

The discretization phase converts the continuum model of the physics into a discrete mathematics problem. Since this is fundamentally a mathematics approximations topic, errors and not uncertainties are the dominate issue in this phase. Some may question why this conversion process should be separated from the solution process. We argue that this conversion process is the root cause of more difficulties in the numerical solution of PDE's than is generally realized. Our view is based on the increasing difficulty of the nonlinear features of PDE's being numerically solved. Taking a historical perspective, early numerical methods and solutions were developed for linear PDE's, such as simple heat conduction, Stokes flow, and linear structural dynamics. Modern numerical solutions have attacked nonlinearities such as high Reynolds number laminar flow and shock waves and, in hindsight, these have proven more difficult than anticipated. Additional nonlinear physics such as turbulent flow, combustion, multiphase flow, phase changes of gases, liquids and solids, fracture dynamics, and chaotic phenomena are also being attacked, most with limited success. When strongly nonlinear features are coupled, the mathematical underpinnings become very thin and the successes become few. Recent investigators [33-35] have clearly shown that the numerical solution of nonlinear ordinary and partial differential equations can be quite different from exact analytical solutions even when using well established methods well within the numerical stability limits of the methods. Yee et al [36] have referred to this phenomena as the "dynamics of numerics" as opposed to the "numerics of dynamics." It is becoming increasingly clear that the mathematical features of strongly nonlinear and chaotic systems can be fundamentally different between the continuous and discrete form, regardless of the grid size [37, 38]. Oberkampf and Blottner [29] have pointed out that the zones of influence between the continuum and numerical counterparts are commonly different, even in the limit as the mesh size approaches zero.

As shown in Fig. 5, we identify three sources of discretization error; discretization of the conservation laws, the boundary conditions, and the initial conditions. The types of errors we are pointing out here are typically very difficult to isolate. One method of identifying these type errors is to analytically prove whether the finite difference method is consistent, that is, does the finite difference method approach the continuum equations as the step size approaches zero. For simple differencing methods, this is quite straightforward. For complex differencing methods such as essentially non-oscillatory (ENO) schemes and second order, multidimensional, upwind schemes, the determination of consistency of the algorithms for a wide range of flow conditions and geometries is difficult. Related issues dealt with in this phase: are the conservation laws satisfied for finite grid sizes, does the numerical damping approach zero as the mesh size approaches zero, and do aliasing errors exist for zero mesh size. Discretization of PDE's are also involved in the conversion of von Neumann and Robin's, i. e., derivative, boundary conditions to difference conditions. We include the conversion of continuum initial conditions to discrete initial conditions, not because there are derivatives involved, but because spatial singularities may be part of the initial conditions. An example of this is the decay of vortex whose initial condition is given as a singularity. Our point is also valid, indeed much more common, when singularities or discontinuities are specified as boundary conditions. Some may argue that these discontinuities and boundary singularities do not actually occur in nature, so the issue of accuracy of representation of these is superfluous. This misses the point completely. If these nonlinear features exist *in the mathematical model of the physics*, the issue is whether the discrete model represents them accurately; not whether they exist in nature. In other words, it is the difference between verification (solving the problem right) and validation (solving the right problem).

3.6 Programming Errors

The credibility of the programming phase is most influenced by unintentional errors, i.e., mistakes. In Fig. 5 we have categorized these mistakes into two types; input and programming errors. We will only briefly discuss these type errors because this topic is thoroughly covered in many software quality assurance texts [30, 31].

Computational researchers and analysts experienced only with model problems, even large scale model problems, typically do not appreciate the concern with input errors. They feel it is simply a matter of carelessness that can easily be remedied by quality assurance practices. This, however, is not the case with very large codes, particularly coupled multi-physics codes, that heavily rely on sophisticated computer aided design/solid modeling codes for input. The complexity of the input data and the resulting room for error with these type codes, is extraordinary. This has been recognized for some time in the nuclear reactor safety thermal-hydraulic analysis field. Formal, structured, and rigorous procedures have been developed to ensure the input data accurately reflects the intended input.

The capturing and elimination of programming errors, although not generating much excitement with computational researchers, remains a major cost factor in producing highly verified software. Even with the maturity of the software quality assurance methods. the difficulty of assessing software quality is becoming more difficult because of massively parallel computers. The complexity of optimizing compilers for these machines, the complexity of message passing, and memory sharing is, in our opinion, increasing faster than the capability of software quality assessment tools. As a case in point, debugging computer codes on massively parallel computers is moving toward becoming a non-deterministic process. That is, the code does not execute the same from one run to another because of other jobs executing on the MP machine. It is still a fundamental theorem of programming that the correctness of a computer code cannot be proven. except for trivial codes. Credibility can only be built by structured coding practices and continued testing, i. e., verification that the coding correctly represents the discrete model.

3.7 Numerical Solution Errors

Numerical solution errors have been investigated longer and in more depth, than any of the errors associated with the numerical solution of PDE's. Indeed, they have been investigated since the beginning of numerical solutions; Richardson in 1910 [39]. These deficiencies in the solution of the discrete equations are properly called errors because they are approximations to the solutions of the original PDE's. As shown in Fig. 5, we categorize these errors into four categories: spatial grid convergence, time step convergence, iterative convergence, and computer round-off. Of these, perhaps the only one that needs explanation is iterative convergence. By this we mean the finite accuracy to which nonlinear algebraic, or transcendental, discrete equations are solved. Iterative convergence error normally occurs in two different phases of the numerical solution. First is the iterative convergence that must be achieved within a time step. Examples are: intra-time step iteration to solve the unsteady heat conduction equation when the thermal conductivity depends on temperature; intra-time step iteration to determine the liquid-solid boundary in a melting or solidification problem; and the iterative solution for nonlinear analytic expressions for transport or thermodynamic properties. On finite volume schemes, for example. conservation of mass, momentum, and energy can be violated with inadequate iterative convergence at each time step. The second type iterative convergence addresses the accuracy of global iterative convergence of an elliptic PDE. Tolerance specifications must be given for the convergence accuracy of each iterative procedure used in a code.

Although we categorize four sources of solution error, it should be noted that they are of two types. The first type is due to the finite discretized solution of the PDE's; spatial grid convergence and time step size convergence are of this type. The second type is due to the approximate solution of the discrete equations, that is, what errors are made in the solution to the given discrete equations. Iterative convergence and round-off error are of this type and they account for the difference between the exact solution to the discrete equations and the computer solution obtained.

All texts dealing with the numerical solution of PDE's address the topic of estimating the magnitude of the spatial grid convergence error. Some of these deal with the errors associated with temporal convergence, iterative convergence, and round-off error. Even though grid convergence error is fairly well understood, it is our view that it is commonly the largest contributor to error in numerical simulations. The reason for this paradox is simple: cost. The grid size used for a numerical solution is usually at the limit of computer time or budgetary constraints; sometimes the grid used is simply considered "good enough" for the simulation at hand. If modeling and simulation is to achieve the level of credibility it is capable of, the lack of careful attention to grid convergence must be corrected.

3.8 Results Interpretation Errors

Figure 5 gives our categorization of types of results interpretation errors: post processor input errors, programming errors, data misrepresentation errors, and data interpretation errors. Post processor input errors and programming errors are the same type of unintentional errors, i.e. mistakes, as pointed out earlier under programming errors. Data misrepresentation errors and data interpretation errors, however, should be considered as intentional errors. By data misrepresentation errors we mean inaccurate or inappropriate construction of continuous functions from the discrete solution in the post-processor. Examples of these are; oscillations of the continuous function in between discrete solution points due to the use of a high order polynomial function in the post-processor; extrapolation of solution variables outside the discrete solution domain of independent variables; and inappropriate interpolation of the discrete solution between multiblock grids. We believe that these should be called intentional errors based on the question: "What is the mathematically correct reconstruction of the continuum functions from the PDE's using the discrete solution points?" When viewed from this perspective, one becomes concerned about the issue because this is not the perspective taken in modern visualization packages. The view of these general purpose packages is that there is no connection between the two. Reconstruction is done based on speed, convenience, and robustness of the package.

By data interpretation errors we mean errors made by the interpreter, i. e., the user, based on observation of the results. In other words, an error made by the user in interpreting the results. By this we *do not* refer to errors in decision made by the user based on the results, such as incorrect design choices or inappropriate policy decisions based on the data. An example of this type of error is the conclusion that a predicted solution is chaotic when it is not (and vice versa).

3.9 Modeling and Simulation Dubiety

The last issue to address in this paper is a final recommendation on terminology. We have defined and pointed out a large number of uncertainties and errors that occur in different phases of modeling and simulation. The distinction between an uncertainty and an error is, we believe, crucial for correct representation and propagation, and for possible reduction or elimination in modeling and simulation. This distinction, however, is not useful when the effect of all of the uncertainties and errors are combined into a measure of total simulation result. We propose the term *dubiety* to mean the total level of doubt or variability in the simulation caused by all sources of uncertainty and error in the simulation.

The quantification of dubiety, since it is a mixture of uncertainty and error, should have the following three features. First, it will present a plausible interval consistent with the available information, in which the predicted value is believed to lie. It will include the characteristic of lack of knowledge which is normally represented as probability distributions. Probability distributions are used to describe uncertainty in code input parameters, and multiple solutions from the simulation also generate probability distributions. As the number of solutions is increased, or when techniques such as Latin Hypercube is used, the confidence in the probability distributions will improve. Second, dubiety will also include the effects of intentional errors, i. e., mathematical approximations, as measured by the exact solutions to PDE's. Use of high quality experimental data for code validation and the use of specialized exact analytical solutions for code verification will build confidence in the numerical solutions. Sensitivity-uncertainty studies with input and numerical parameters will produce some indication of total simulation credibility. This type error will produce an error with some similarities to precision, or random. errors in experimental uncertainty. Third, dubiety can detect unintentional errors, i. e., mistakes, only when independent information is used for comparisons. For example, comparisons with other independent numerical solutions, by verification testing, and by experimental validation data. Strategies for detecting unintentional errors will be analogous to those for detecting bias errors in experimental data.

4. Summary and Conclusions

We have presented a framework for the phases of modeling and simulation in which the physical system, or proposed system, is described by partial differential equations. We have carefully defined the meaning of, and distinguished between, uncertainty and error. These comprehensive definitions are required to categorize the broad range of deficiencies that can exist in modeling and simulation. Using these definitions, we defined a taxonomy for classes of sources of uncertainty and error that are appropriate to each of the phases identified in modeling and simulation. Our framework applies regardless of whether the discretization procedure is based on finite elements, finite volumes, or finite differences.

With this structure we believe more comprehensive procedures should be developed for representing, combining, and propagating individual sources of uncertainty and error through the entire modeling and simulation process. We believe the present work shows that the traditional probabilistic representations and propagation procedures will not be sufficient to account for error sources. Non-probabilistic mathematical representations, such as possibility theory, fuzzy sets, and Dempster-Shafer theory, may be more appropriate for error sources and specialized types of uncertainty. Although the advantages of these representations are speculative, it is hoped that statisticians and information theorists will become interested in applying these to modeling and simulation. From the heat transfer and fluid dynamics analysts view point, most of these approaches will appear alien. We believe, however, these new type approaches may be needed to more confidently assess the dubiety of modeling and simulation for complex engineered systems. For the prediction of high consequence events, particularly those that have little or no experimental data, these methods may prove to be critical.

References

1. *American Nuclear Society: Guidelines for the Verification and Validation of Scientific and Engineering Computer Programs for the Nuclear Industry*, . 1987.

2. Bradley, R.G. *CFD Validation Philosophy*. in *AGARD Symposium Validation of Computational Fluid Dynamics*. 1988. Lisbon, Portugal: North Atlantic Treaty Organization.

3. *International Standard ISO 9000-3: Quality Management and Quality Assurance Standards - Part 3: Guidelines for the application of ISO 9001 to the development, supply and maintenance of software*, . 1991.

4. Blottner, F.G., *Accurate Navier-Stokes Results for the Hypersonic Flow Over a Spherical Nosetip*. Journal of Spacecraft and Rockets, 1990. **27**(2): p. 113-122.

5. Mehta, U.B., *Some Aspects of Uncertainty in Computational Fluid Dynamics Results*. Journal of Fluids Engineering, 1991. **113**: p. 538-543.

6. Oberkampf, W.L. *A Proposed Framework for Computational Fluid Dynamics Code Calibration/Validation*. in *18th AIAA Aerospace Ground Testing Conference*. 1994. Colorado Springs, CO: American Institute of Aeronautics and Astronautics.

7. Mehta, U.B., *Guide to Credible Computer Simulations of Fluid Flows*. Journal of Propulsion and Power, 1996. **12**(5): p. 940-948.

8. Roache, P.J., *Need for Control of Numerical Accuracy*. Journal of Spacecraft and Rockets, 1990. **27**(2): p. 98-102.

9. Roache, P.J., *Quantification of Uncertainty in Computational Fluid Dynamics*, . 1997.

10. Strouboulis, T. and J.T. Oden, *A Posteriori Error Estimation of Finite Element Approximations in Fluid Mechanics*. Computational Methods in Applied Mechanics and Engineering, 1990. **78**: p. 201-242.

11. Celik, I. and W.-M. Zhang. *Application of Richardson Extrapolation to Some Simple Turbulent Flow Calculations*. in *The Fluids Engineering Conference*. 1993. Washington, DC: The Fluids Engineering Division, ASME.

12. Babuska, I., T. Strouboulis, and C.S. Upadhyay, *A Model Study of the Quality of a Posteriori Error Estimators for Linear Elliptic Problems: Error Estimation in the Interior of Patchwise Uniform Grid of Triangles*. Computational Methods in Applied Mechanics and Mechanical Engineering, 1994. **114**: p. 307-378.

13. Ferziger, J.H. and M. Peric', *Further Discussion of Numerical Error in CFD*. International Journal for Numerical Methods in Fluids, 1996. **23**: p. 1263-1274.

14. Mehta, U.B. *Guide to Credible Computational Fluid Dynamics Simulations*. in *26th AIAA Fluid Dynamics Conference*. 1995. San Diego, CA: AIAA.

15. Roache, P.J. *Verification of Codes and Calculations*. in *26th AIAA Fluid Dynamics Conference*. 1995. San Diego, CA: AIAA.

16. Draper, D., *Assessment and Propagation of Model Uncertainty*. Journal R. Statistics Soc. B, 1995. **57**(1): p. 45-97.

17. Lasky, K.B., *Model Uncertainty: Theory and Practical Implications*. IEEE Transactions on Systems, Man and Cybernetics-Part A: Systems and Humans, 1996. **26**(3).

18. Neelamkavil, F., *Computer Simulation and Modelling*. 1st ed. 1987, New York: John Wiley & Sons.

19. Zeigler, B.P., *Multifacetted Modelling and Discrete Event Simulation*. 1st ed. 1984, Orlando: Academic Press.

20. Zeigler, B.P., *Theory of Modelling and Simulation*. 1st ed. 1976, New York: John Wiley & Sons.

21. Morgan, M.G. and M. Henrion, *UNCERTAINTY A guide to Dealing with Uncertainty*

in Quantitative Risk and Policy Analysis. 1st ed. 1990, New York: Cambridge University Press.
22. Bossel, H., *Modeling and Simulation*. 1st ed. 1994, Wellesley, MA: A. K. Peters, Ltd.
23. Schlesinger, D.S. and e.a. Dr. Roy E. Crosbie, *Terminology for Model Credibility*. Simulation, 1979. **32**(3): p. 103-104.
24. Sargent, R.G., *Simulation Model Validation*, in *Simulation and Model-Based Methodologies: An Integrative View*, T.I. Oren, B.P. Zeigler, and M.S. Elzas, Editors. 1984, Springer-Verlag Berlin Heidelberg: Syracuse. p. 537-555.
25. Sargent, R.G. *An Expository on Verification and Validation of Simulation Models*. in *1985 Winter Simulation Conference*. 1985. Sacramento, California.
26. Nance, R.E., *Model Representation in Discrete Event Simulation: The Conical Methodology*, . 1981, Virginia Polytechnic Inst. and State University.
27. Balci, O. *Guidelines for Successful Simulation Studies*. in *Proceedings of the 1990 Winter Simulation Conf.* 1990.
28. Box, G.E.P., *Sampling and Bayes' Inference in Scientific Modeling and Robustness*. Journal Statist. Soc. A, 1980. **143**(A): p. 383-430.
29. Oberkampf, W.L. and F.G. Blottner, *Issues in Computational Fluid Dynamics Code Verification and Validation*, . 1997, Sandia National Laboratories.
30. Lewis, R.O., *Independent Verification and Validation*. 1st ed. 1992, New York: John Wiley & Sons, Inc.
31. Knepell, P.L. and D.C. Arangno, *Simulation Validation A Confidence Assessment*. 1st ed. 1993, Washington: IEEE Computer Society Press.
32. Klir, G.J. and T.A. Folger, *Fuzzy Sets, Uncertainty, and Information*. 1st ed. 1988, Englewood Cliffs, NJ: Prentice Hall.
33. Yee, H.C., P.K. Sweby, and D.F. Griffiths, *Dynamical Approach Study of Spurious Steady-State Numerical Solutions of Nonlinear Differential Equations I. The Dynamics of Time Discretization and its Implications for Algroithm Development in Computational Fluid Dynamics*. Journal of Computational Physics, 1991: p. 249-310.
34. Yee, H.C. and P.K. Sweby, *Nonlinear Dynamics & Numerical Uncertainties in CFD*, . 1996, NASA/Ames Research Center: Moffett Field, CA.
35. Yee, H.C., *et al. On Spurious Behavior of CFD Simulations*. in *13th AIAA Computational Fluid Dynamics Conference*. 1997. Snowmass, CO: AIAA.
36. Yee, H.C. and P.K. Sweby, *Dynamical Approach Study of Spurious Steady-State Numerical Solutions of Nonlinear Differential Equations II. Global Asymptotic Behavior of Time Discretizations*. Computational Fluid Dynamics, 1995. **4**: p. 219-283.
37. Yee, H.C. and P.K. Sweby, *Global Asymptotic Behavior of Iterative Impllicit Schemes*. International Journal of Bifurcation and Chaos, 1994. **4**(6): p. 1579-1611.
38. Yee, H.C. and P.K. Sweby, *Some Aspects of Numerical Uncertainties in in Time-Marching to Steady-State Numerical Solutions*. 27th AIAA Fluid Dynamics Conference, 1996: p. 47.
39. Richardson, L.F., *The Approximate Arithemetical Solution by Finite Differences of Physical Problems Involving Differential Equations, with an Application to the Stresses in a Masonry Dam*. Transaction of the Royal Society of London. 1910. **Series A, Vol. 210**: p. 307-357.

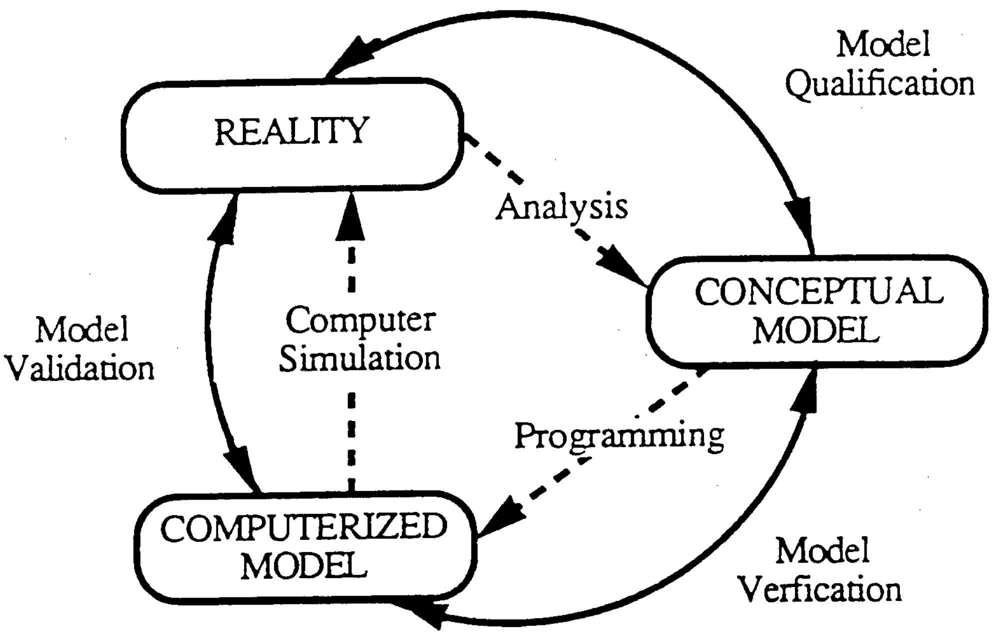


Figure 1

Early View of Modeling
and Simulation (from Ref. 23)

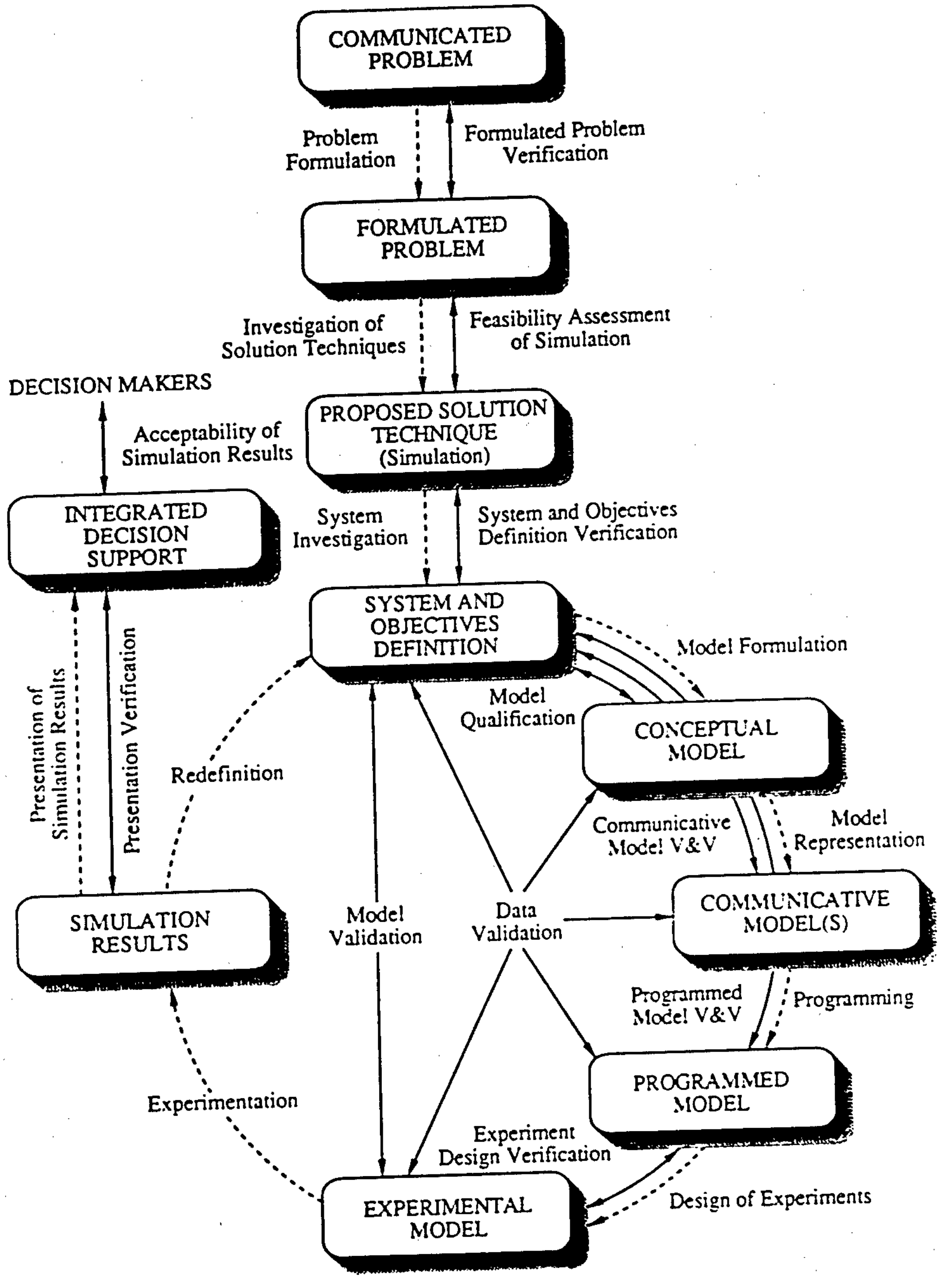


Figure 2

View of Modeling and Simulation from Nance (Ref. 26) and Balci (Ref. 27)

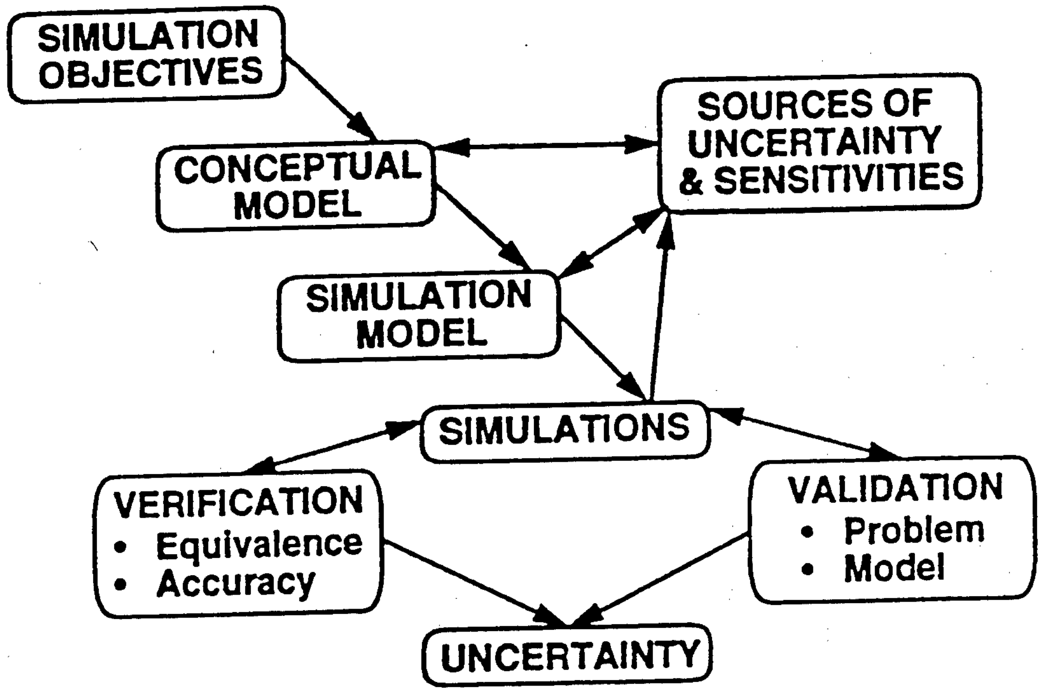
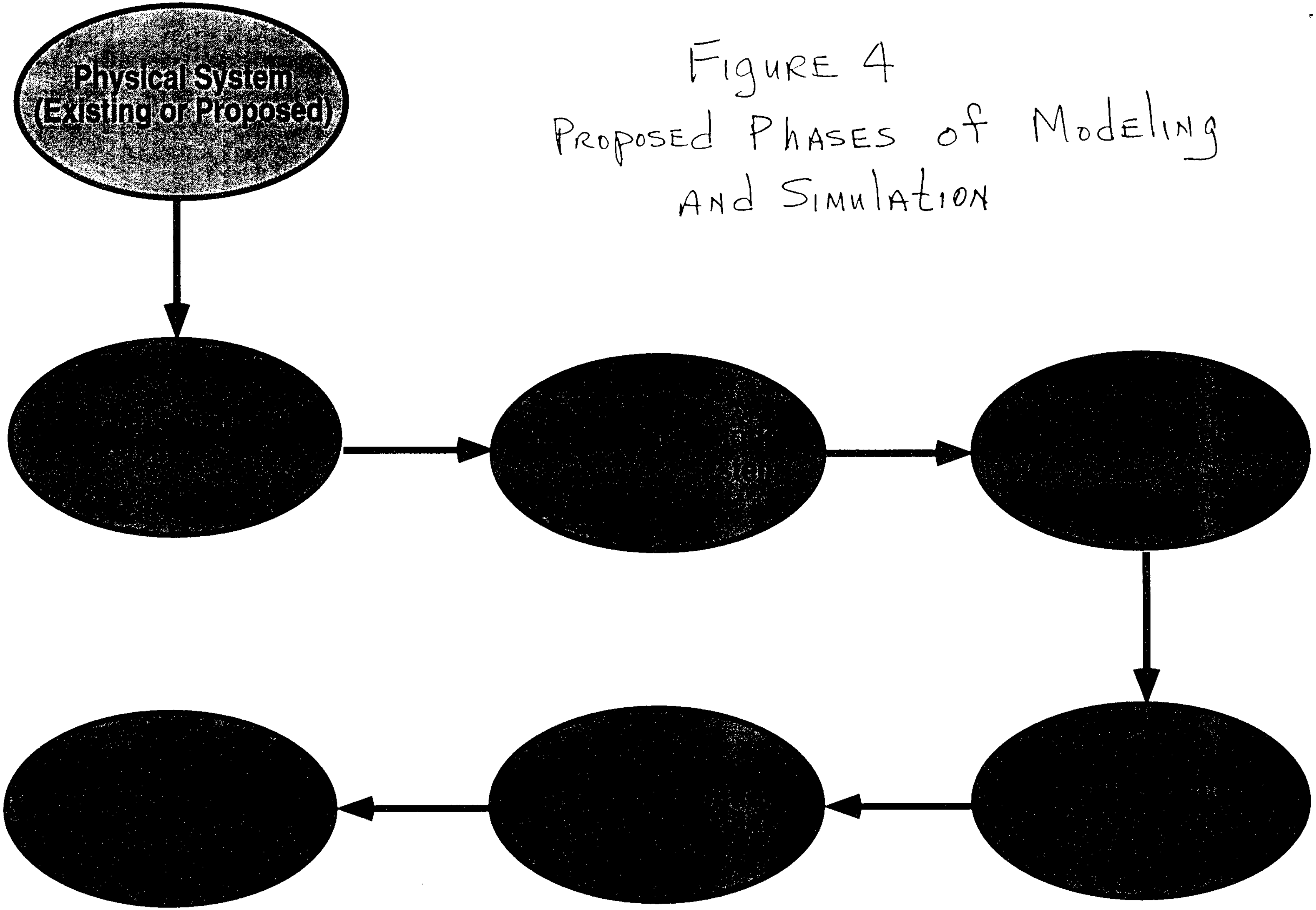


Figure 3

View of Modeling and Simulation from Mehta (Ref. 7)

Figure 4
Proposed Phases of Modeling
And Simulation



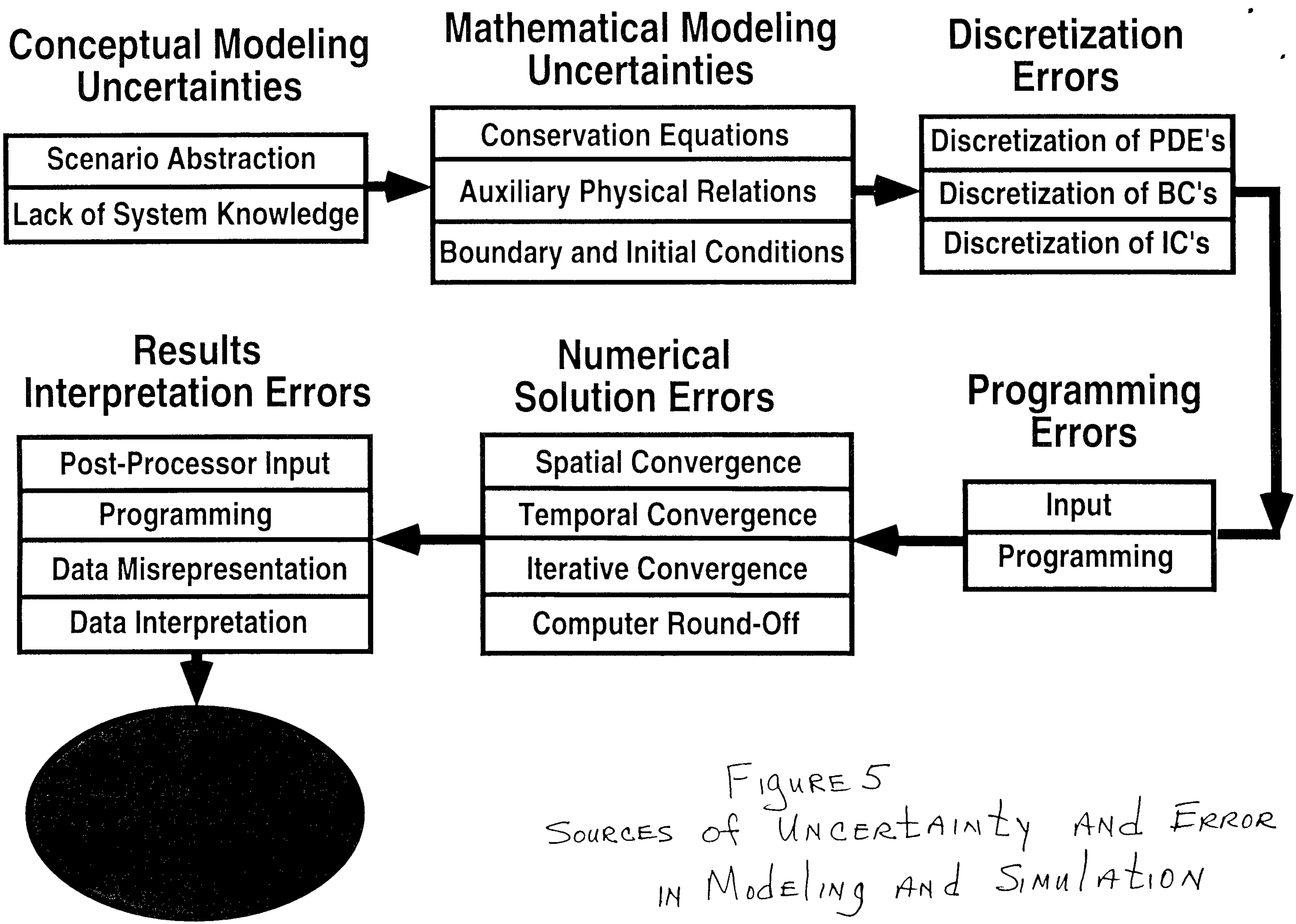


Figure 5
Sources of Uncertainty and Error in Modeling and Simulation

M98000927

Report Number (14) SAND--97-2607C
CONF-980610--

Publ. Date (11) 199710

Sponsor Code (18) DOE/DP ; DOE/MA , XF

UC Category (19) UC-705 ; UC-905 , DOE/ER

DOE