Modeling and Validation Challenges for Complex Systems

Mikel D. Petty
University of Alabama in Huntsville
301 Sparkman Drive, Shelby Center 144, Huntsville, AL 35899 USA
pettym@uah.edu 256-824-4368

Keywords
Verification and validation, Complex systems, Sensitivity to initial conditions, Emergent behavior, Composition of component

Abstract: Many important systems, both natural and artificial, may be classified as complex, and the study of complex systems is ongoing. Such systems have special defining characteristics, including sensitivity to initial conditions, emergent behavior, and composition of components. Complex systems are increasingly prevalent as the subject of modeling efforts. There are at least two reasons for this; first, the systems that are of the greatest practical interest and thus most likely to be modeled tend to be complex, and second, because complex systems resist closed form analysis modeling is often the only way to study them. Unfortunately, the special characteristics of complex systems lead to additional challenges in both effectively modeling them and in validating the models. This paper, which takes the form of an introductory tutorial and literature survey, first defines complex systems in terms of their key characteristics and describes how validation risk applies to models of them. It then identifies a series of modeling and validation challenges that follow from the defining characteristics and suggests mitigation approaches for those challenges.

1. Introduction

“... our best equations for the weather differ from our best computer models based on those equations, and both of those systems differ from the real thing…” [1]

“... complexity lies somewhere between order and chaos.” [2]

Complex systems, where “complex” is meant in the sense of complexity theory as opposed to simply a synonym for “complicated”, are with increasing frequency the subject of modeling efforts. Among the reasons for this, two stand out. First, the systems of the greatest practical interest, and thus those most likely to be worth the effort and expense of being modeled, tend to be complex. Second, as a result of their special characteristics, complex systems generally resist closed form mathematical analysis, and so modeling is often the best or even the only way to study and experiment with them.

Complex systems have a number of special defining characteristics, including sensitivity to initial conditions, emergent behavior, and composition of components. Unfortunately for those involved in modeling complex systems, these special characteristics of complex systems lead to additional challenges beyond those encountered with non-complex systems in both modeling them accurately and effectively and in reliably and completely validating the models.

This paper, which is meant as an introductory tutorial and brief literature survey, has four main sections. The first describes complex systems and lists their defining characteristics, and motivates the interest in validating models of complex systems by discussing validation risk. Then, each of the following sections discusses one of three selected defining characteristics of complex systems (sensitivity to initial conditions, emergent behavior, and composition of components), explaining why the characteristic in question makes modeling and validation more difficult and offering some approaches to dealing with and mitigating the difficulties.

2. Complex systems

Complex systems were recognized as qualitatively distinct from non-complex systems at least as early as 1984, with the founding of the Santa Fe Institute, a research institute devoted to complexity theory [3]. Since then, a body of specialized knowledge has been developed on the subject, driven by both theoretical and experimental investigations [4].

2.1 Definition of complex systems

A range of definitions of complex system are available. Although they are far from as reassuringly consistent or precise as that of, say, an equivalence relation (e.g., see [5]), they are nevertheless informative.

“A system comprised of a (usually large) number of (usually strongly) interacting entities, processes, or agents, the understanding of which requires the development, or the use of, new scientific tools, nonlinear models, out-of equilibrium descriptions and computer simulations.” [6]

1 Quoted from [6], where it is attributed to [26].
There is general consensus that certain defining characteristics or properties are associated with complex systems. These characteristics are individually arguable, in that not every complex system necessarily exhibits every one of these characteristics, but they are collectively definitive; most complex systems will exhibit most of these characteristics. Taken together, they define the class of complex systems and serve to distinguish them from non-complex systems. A list of the defining characteristics with brief descriptions follows; the first three are described in detail in the following sections, whereas the others are briefly described here:

1. **Sensitivity to initial conditions.** Description to follow.
2. **Emergent behavior.** Description to follow.
3. **Composition of components.** Description to follow.
4. **Uncertain boundaries.** Determining the boundary between a complex system and the environment in which it is situated and with which it interacts can be difficult.
5. **Nesting.** Components of a complex system may themselves be complex systems.
6. **State memory.** Future states of a complex system often depend on past states in ways that are difficult to understand or model.
7. **Non-linear relationships.** Relationships between components of a complex system may be non-linear, which means a small cause may have a large effect.
8. **Feedback loops.** Negative (damping) and positive (amplifying) feedback loops exist between elements of complex system.

### 2.2 Validation risk in models of complex systems

Important systems, complex systems, and modeled systems overlap to a great extent. Systems that are important to their users, for reasons of safety, economics, or ubiquity, are often complex; the reverse is also true. For example, financial markets are important to those who participate in them, whether voluntarily or involuntarily, because of their potential impact on the participants’ quality of life and long-term security, and they exhibit all of defining characteristics listed earlier. Similarly, systems that are important are also often modeled, because their importance makes them more likely to be worth the effort and expense of being modeled; and again the reverse is also true. Finally, complex systems are often modeled, and once more the reverse is true. Because of their inherent structure, complex systems are often difficult to study using closed form mathematical analysis [2]. Consequently, modeling is often the best or even the only way to study or experiment with them.

---

2 Image acknowledgements for Figure 1: M. Peteron, GNU Free Documentation License, Wikipedia Commons (Air traffic control); National Aeronautics and Aerospace Administration, Public domain, Wikipedia Commons (Weather); National Institute for Standards and Technology, Public domain, Wikipedia Commons (Stock market).

3 An overlapping but somewhat different list is given in [15]; that list includes adaptivness and self-organization.

---

Figure 1. Examples of complex systems: air traffic control, weather, and the stock market.

“A complex system is one whose evolution is very sensitive to initial conditions or to small perturbations, one in which the number of independent interacting components is large, or one in which there are multiple pathways by which the system can evolve.” [7]

Both natural and artificial systems satisfy these definitions; examples of systems that are considered to be complex are illustrated in Figure 1.
Models are subject to validation risk. The general concept of validation risk is that validation that is improperly or incompletely performed can result in risk to the developers and/or the users of the model. This general notion has been refined into specific types of validation error and the type of validation risk that results from each. The validation errors are known as Type I, Type II, and Type III, and are defined in a manner that closely parallels the like-named error types in statistical hypothesis testing. Figure 2 summarizes these error types.

Whenever a model is used validation risk exists, and for a model of an important system, that risk is proportional to the importance of the system and to the model’s intended use. Obviously, a Type II validation error clearly has less potential consequences for a model of ant behavior being used for a video game than a model of metal fatigue being used to design the airframe of an airliner. Decisions made about important system using models can have major impact. As an example, consider the 2007 financial crisis in the United States. Some financial analysis have argued that that crisis was in significant part triggered by a financial model, namely the famous (or infamous) Gaussian copula, which is a model of the prices of collateralized debt obligations:

$$\Pr[T_A < 1, T_B < 1] = \Phi_d(\Phi^{-1}(F_A(1)), \Phi^{-1}(F_B(1)), \gamma)$$  \[8\]

The significant overlap of important systems, complex systems, and modeled systems means that our models are often of systems that are both important and complex; their importance magnifies validation risk, and their complexity complicates validation. Given the validation risk associated with models of important and complex systems, it is prudent to expend validation effort proportional to the risk, and to adapt or develop validation methods suitable for complex systems.

---

4 The figure is from [17]; it is adapted from a flowchart that shows how the different types of errors might arise found in [14]. Definitions of Type I and Type II validation errors analogous to the statistical errors of the same name appear in [27] and in subsequent editions of this source, e.g., [11].
3. Sensitivity to initial conditions

“This is only true when small variations in the initial circumstances produce only small variations in the final state of the system. In a great many physical phenomena this condition is satisfied; but there are other cases in which small initial variation may produce a very great change in the final state of the system, as when the displacement of the ‘points’ causes a railway train to run into another instead of keeping its proper course.” [9]

“Small differences can build upon themselves and create large differences, making precise prediction difficult.” [2]

The first of the three defining characteristics of complex systems to be examined for its effect on modeling and validation is sensitivity to initial conditions. Here the phrase “initial conditions” refers, of course, to either the starting state of the system (e.g., a rocket motor at ignition), or if the system has an effectively continuous existence (e.g., the weather), the state of the system at the beginning of the time period being studied or modeled. The state evolution of complex systems can be highly sensitive to its initial conditions, with the result that small differences in initial state can become magnified over time into large differences in future state [1]. Figure 3 illustrates this; in the figure, the horizontal axis represents time, advancing from left to right, and the vertical axis represents system state, notionally simplified to a single variable. States that are only slightly different at some initial time \( t_0 \) can evolve away from each other, becoming arbitrarily different at some future time \( t_1 \).

Models of complex systems, if they accurately represent the system’s characteristics, can be similarly sensitive to initial conditions. From two model starting states that are quite similar, the execution of a model of a complex system can produce widely divergent end states.

3.1 Modeling

Sensitivity to initial conditions can introduce modeling challenges in these ways:

1. Implementation side effects. Technical aspects of the model that are purely implementation details and do not correspond to any aspect of the simuland\(^5\) can have significant side effects that influence, or even overwhelm, the results. A well-known example is the effect of the numerical precision of the implementation language on numerical integration of differential equations in physical models [10].\(^6\) In stochastic models that rely on random number generators, the seed and cycle length of the random number stream can, through the magnifying effect of sensitivity, significantly affect the model’s results [11].

2. Sensitivity consistency. If a complex system is sensitive to initial conditions, the modeler may seek similarly sensitivity in a model of that system. However, given the nature of the sensitivity, it can be quite difficult accurately to match the model’s sensitivity to that of its simuland. Even if both the complex system and a model of it are sensitive, small differences between simuland sensitivity and model sensitivity can lead to large differences in outcomes.

---

\(^5\) The simuland is the system, phenomenon, or process that is the subject of a model, i.e., the modeled system [17].

\(^6\) [10] provides an example; fourth-order Runge-Kutta integration with a fixed time step used to calculate an orbit in a two-body (sun and planet) gravitational system completely breaks down in the vicinity of the sun due to numerical precision issues, with the result that the simulated planet incorrectly “flies off completely into space”.

![Figure 3. Sensitivity to initial conditions; system state diverges over time.](image-url)
3. Input imprecision. Because sensitivity magnifies small differences in initial conditions, a small difference between the simuland’s true initial state and the values of the model input data describing that state can again lead to large differences in outcomes. Consider, for example, a weather model that uses a three-dimensional array of air temperature, pressure, and humidity values to define the initial state of the atmosphere. Small errors in measuring those values can be magnified, as the model executes, into large discrepancies between the model’s prediction and the actual weather. The input data precision needed by the model to accurately predict the simuland’s future may exceed that obtainable due to limits in instrumentation accuracy or observation availability. This observational uncertainty is one reason that the useful predictive power of current weather models is currently limited to a few days, and the maximum achievable limit, even with perfect models, is considered to be “about two weeks” [12].

These methods can mitigate the modeling challenges associated with sensitivity to initial conditions:
1. Selective abstraction. During conceptual modeling, identify simuland features and state variables that are not required for the model to satisfy its intended purposes. Omit them in the implemented model, thereby eliminating them as possible sources of sensitivity.
2. Ensemble forecasting. The core idea of ensemble forecasting is to execute multiple runs of a model, each of which was initialized with slightly different initial states, and then develop a prediction based on the multiple results. The differences in the inputs are intended to reflect the uncertainty in the knowledge of the initial state. The multiple results may be aggregated or averaged, and the variation and divergences between them analyzed; the details of aggregation and analysis depend on the application, but statistical methods are often employed. In some forms and contexts this is a familiar idea; modelers using a discrete event simulation to study a queueing system often conduct multiple trials, each beginning with a different random number seed. In the case of weather models, different values for the initial conditions of the atmosphere may be used, with the differences generated based on the noise or uncertainty in the observations upon which the input is based [1]. The uncertainty of the forecast may be estimated based on the variation in the different forecasts generated.

3.2 Validation
Sensitivity to initial conditions can introduce validation challenges in these ways:
1. Results distributions. Broad distributions (i.e., large variance) in both simuland observations and model results can reduce the power of statistical comparisons of the two [13].
2. Sensitivity analysis. The potential for widely divergent outcomes from closely similar initial conditions can complicate conventional sensitivity analysis by requiring more closely spaced sampling of the response surface to capture the response variation.
3. Input imprecision. Matching model initialization data to simuland observation data precisely enough to compare simulation outcome and model results can be problematic.

These methods can mitigate the validation challenges associated with sensitivity to initial conditions:
1. Increased trials. Increasing the number of trials (i.e., executions of the model) can regain some statistical power through larger sample sizes.
2. Sensitivity analysis. Sensitivity analysis can be used as a validation method by statistically comparing the magnitude and variability in the simuland observations to the magnitude and variability in the model results, in effect using sensitivity as a metric for validation comparison [14].
3. Precision awareness. Understand the precision available in simuland observation data, and based on that precision, use an appropriate comparison threshold when comparing simuland observations and model results. For example, it is a mistake to expect the model to match the simuland within one unit when the observations are only accurate to within five units.

7 In addition to multiple runs of a single model, ensemble forecasting may also refer to an aggregating or merging of the results of multiple models. This approach is used to predict hurricane tracks.
4. Emergent behavior

“The behavior of many complex systems emerges from the activities of lower-level components.” [2]

“Much of the focus of complex systems is how ... interacting agents can lead to emergent phenomena. ... individual, localized behavior aggregates into global behavior that is, in some sense, disconnected from its origins.” [2]

The second of the three defining characteristics of complex systems to be examined for its effect on modeling and validation is emergent behavior. Emergent behavior is behavior that is not explicitly encoded in the agents or components that make up the model; rather, it emerges during a simulation from the interaction of agents or components with each other and the simulated environment [15].

An important aspect of emergent behavior is that it is not directly predictable or anticipatable from the individual agents’ or components’ behaviors, even if they are known completely. Figure 4 illustrates a form of natural emergent behavior that exhibits this. Emergent behavior is, in some intuitive sense, unexpected; it produces “surprise” in the observer [2]. There is the possibility of multiple levels of emergence, with mesoscale behavior that emerges from microscale interactions itself contributing to the emergence of even higher level macroscale behaviors [2].

4.1 Modeling

Emergent behavior can introduce modeling challenges in these ways:

1. Incomplete observations. Because emergent behavior is potentially unpredictable, available observations of simuland may not include all possible simuland emergent behavior. Indeed, the modeler may not even be aware of some potential simuland emergent behaviors.

2. Indirect representation. Because emergent behavior is not, in general, predictable from the individual behavior of agents or components within the complex system, those aspects or characteristics of it that produce emergent behavior can be difficult to identify and include in the model.

3. Overabstraction risk. Because emergent behavior is produced indirectly from potentially non-obvious aspects of simuland, modeler may unintentionally abstract away those aspects, eliminating the possibility of the model generating interesting or important emergent behavior.

These methods can mitigate the modeling challenges associated with emergent behavior:

1. Additional observations. Increasing the number or duration of simuland observations, and broadening the range of conditions under which the simuland is observed, can increase the likelihood of observing and detecting the full repertoire of emergent behaviors.

2. Conceptual modeling focus. When developing the conceptual model of a complex system, give explicit attention to the inclusion of emergent behaviors, or aspects of the complex system that may give rise to emergent behaviors (such as inter-agent interactions).

---

8 Image acknowledgement for Figure 4: C. A. Rasmussen, Public domain, Wikipedia Commons.
4.2 Validation

Emergent behavior can introduce validation challenges in these ways:

1. *Incomplete observations.* Emergent behavior is, by its nature, difficult to predict, observe, measure in the simuland; this was already noted as a modeling challenge. It is also a validation challenge, as some emergent behavior observed in the model results may not have been observed in the simuland, thus leaving gaps in the data for use in validating the model’s behavior.

2. *Incomplete results.* Conversely, emergent behavior observed in the simuland can be similarly difficult to generate in the model results. Of course, if the behavior is not in the model results, it can not be validated beyond noting that it is missing.

3. *Face validation unreliability.* Because of emergent behavior is unpredictable, face validation based on subject matter experts is less reliable. The experts may overestimate or underestimate the likelihood of occurrence of emergent behavior, or they may have little direct knowledge of it.

4. *Test case uncertainty.* Because emergent behavior is not directly predictable, designing model validation test cases (trials) which will generate specific emergent behaviors for validation can be difficult.

These methods can mitigate the validation challenges associated with emergent behavior:

1. *Additional observations.* Increasing the number or duration of simuland observations, and broadening the range of conditions under which the simuland is observed, increases the likelihood of acquiring the data needed to validate emergent behavior.

2. *Structured face validation.* To overcome deficiencies in the knowledge of any particular subject matter expert, use teams of experts and conduct organized face validation assessments. The latter may be based pre-planned validation scenarios designed to cover the full range of simuland behaviors [16] [17] and employ Delphi methods, wherein panels of experts make forecasts and examine the model’s results over multiple rounds, eventually converging on a consensus assessment of validity [18].

3. *Scenario space search.* Generate validation test cases automatically via heuristic search in scenario space, i.e., generating new test cases based on

---

9 Experts often underestimate the probability of an unlikely event, implicitly assuming a normal probability distribution when a “fatter tailed” distribution would be more appropriate. Examples of such distributions and their asserted applications include power laws for city sizes [28] and deaths in warfare [29], and Lévy stable laws for stock market price changes [12].
5. Composition of components

“We would, however, like to make a distinction between complicated worlds and complex ones. In a complicated world, the various elements that make up the system maintain a degree of independence from one another. ... Complexity arises when the dependencies among the elements become important.” [2]

The third of the three defining characteristics of complex systems to be examined for its effect on modeling and validation is composition of components. Complex systems are, by definition, composed of interacting components. Similarly, models of complex systems are often composed of submodels, and those submodels are most typically organized in a structure that reflects the structure of the complex system itself. For example, a spacecraft model may be composed of power system and thermal submodels, with the thermal submodel providing input to power system model to predict power loading.

5.1 Modeling

Composition of components can introduce modeling challenges in these ways:

1. **Interface compliance.** The existence of multiple submodels, and thus the need for interfaces between them, adds new opportunities for modeling errors, such as mismatches in data types, measurement units, and execution sequence.

2. **Architecture selection.** The appropriate software architecture framework for organizing and connecting the component models (such as hierarchy, blackboard, or agent-based) may not be obvious, and it may have unintended effects on the model results [19].

3. **Model correlation.** Different component models may have differences (such as underlying assumptions, representational granularity, or level of fidelity) that negatively affect the overall model’s results [20].

These methods can mitigate the modeling challenges associated with composition of components:

1. **Interface analysis.** Specifically examine submodel-to-submodel interfaces to determine if interface structures are consistent and accurate [14].

---

10 As discussed earlier, interactions between those components can lead to emergent behavior.

11 Arguably, the entire subject of simulation interoperability is embedded in this modeling challenge. Clearly, this is no small matter.
2. Known problem review. Review available lists of known interoperability problems typically encountered to see if they apply [21].

3. Architecture reuse. Reuse and revise known model architectures when appropriate, and exploit available architecture-based systems engineering processes (e.g., the Distributed Simulation Engineering and Execution Process [22]).


5.2 Validation

Composition of components can introduce validation challenges in these ways:

1. Weakest link validity. The overall validity of a model assembled as a composition of component models may be limited by the lowest fidelity component model. For example, a high fidelity ground vehicle movement model composed with a low fidelity terrain model will likely not produce accurate movement speeds.

2. Error location ambiguity. Errors in model results detected during model validation may be difficult to associate with correct component model; indeed, they may result from an interface error, rather than one of the component models.

3. Statistical method unsuitability. The statistical methods used most often in validation typically compare single variables, e.g., the Student t test compares the means of two populations, or the Mann-Whitney U test determines whether two independent samples of observations come from the same distribution. Models of complex systems have states represented by multiple non-linear variables related non-linearly, requiring the of use multivariate methods that accommodate non-linear effects [14].

4. Noncomposability of validity. In a model assembled as a composition of components, i.e., from submodels, the submodels are typically validated individually. Unfortunately, submodel validity does not ensure composite model validity; even if the submodels are separately valid, the composite models may not be. It has been mathematically proven that for non-trivial models separately valid component models can not be assumed to be valid when composed [23].

These methods can mitigate the validation challenges associated with composition of components:

1. Uncertainty estimation. Determine or estimate the possible error range for key model results variables for each component model. Then propagate and accumulate those errors to find the overall error range for the same variables for the composite model [24]. If the overall error is too large, revise the model.

2. Non-linear multivariate statistics. Apply multivariate statistical methods to validation of non-complex systems models. For example, Hotelling T^2-statistic, which is a generalization of Student’s t statistic that is used in multivariate hypothesis testing, can be used for constructing ellipsoidal joint confidence intervals in validation [25].

3. Composition validation. During validation of a composite model, validate both the component models individually and overall composite model. This is directly analogous to conventional unit and system testing in software engineering practice.

6. Summary

Complex systems, which are increasingly often the subject of modeling efforts, have certain defining characteristics that make them more difficult to model and make models of them more difficult to validate. The specific modeling and validation challenges can be associated with the complex system characteristic that causes them. Although these challenges can be problematic, and in some cases are in principle impossible to overcome entirely, they can often be mitigated through informed application of appropriate methods.

7. References


8. Author’s biography

Mikel D. Petty is Director of the University of Alabama in Huntsville’s Center for Modeling, Simulation, and Analysis. He is also an Associate Professor of Computer Science and a Research Professor of Industrial and Systems Engineering and Engineering Management. Prior to joining UAH, he was Chief Scientist at Old Dominion University’s Virginia Modeling, Analysis, and Simulation Center. He received a Ph.D. in Computer Science from the University of Central Florida in 1997. Dr. Petty has worked in modeling and simulation research and development since 1990 in areas that include human behavior modeling, simulation interoperability and composability, multi-resolution simulation, and applications of theory to simulation. He has published over 165 research papers and has been awarded over $14 million in research funding. He served on a National Research Council committee on modeling and simulation, is a Certified Modeling and Simulation Professional, and is an editor of the journals SIMULATION and Journal of Defense Modeling and Simulation. He was the dissertation advisor to the first two students to receive Ph.D.s in Modeling and Simulation at Old Dominion University.