Multiresolution, multiperspective modeling (MRMPM) as an enabler of exploratory analysis

Paul K. Davis

RAND and the RAND Graduate School,
1700 Main St., Santa Monica, CA 90407-2138*

ABSTRACT

Exploratory analysis examines the consequences of uncertainty—not merely by standard sensitivity methods, but more comprehensively. It is particularly useful for gaining a broad understanding of a problem domain before dipping into details. Although exploratory analysis can be accomplished with models of many types, it is facilitated by multiresolution, multiperspective modeling (MRMPM) structures. Moreover, a knowledge of related design principles facilitates the characterization of more normal models in terms that permit exploratory analysis. This paper describes the connections and notes that, with current and emerging personal computer tools, MRMPM methods are becoming practical.[#]

Keywords: model abstraction, multiresolution models, exploratory analysis, sensitivity analysis, uncertainty analysis, policy analysis, variable resolution models

1. INTRODUCTION

Real-world strategy problems are typically beset with enormous uncertainties that should be central in assessment of alternative courses of action—although, to be sure, individuals and organizations often suppress those uncertainties and give a bizarre level of credence to wishful-thinking planning factors and other simplifications.[2,3] In the past, an excuse for downplaying uncertainty analysis—except for marginal sensitivity analysis around some "best-estimate" baseline of dubious validity—was the sheer difficulty of doing better. If analysis depended on the setup, run, and analysis times of computer programs, then extensive uncertainty work was often ruled out. Today, technology permits us to do extensive uncertainty analysis and to aspire even to doing it well with personal computers. Although better tools are still needed, much can be done with what is already widely available in the commercial world.

A key to treating uncertainty well is *exploratory analysis*. The objectives of exploratory analysis include (1) understanding the implications of uncertainty for the problem at hand and (2) informing the choice of strategy and subsequent modifications. In particular, *exploratory analysis can help identify strategies that are flexible, adaptive, and robust.*[2,3] This paper describes exploratory analysis (§2); puts it in context (§3); discusses enabling technology and theory (§4); points to companion papers applying the ideas, and concludes with some challenges for those involved in enabling technology for modeling and simulation (§5).

2. EXPLORATORY ANALYSIS

2.1 Definition

Exploratory analysis[1] is analysis that examines the consequences of uncertainty. In a sense, it can be thought of as sensitivity analysis done right, but because it is so different in practice from what most people think of as sensitivity analysis, it deserves a separate name. It is closely related to scenario space analysis[2] and "exploratory modeling."[4,5] It is particularly useful for gaining a broad understanding of a problem domain before dipping into details. That, in turn, can greatly assist in the development and choice of strategies. It can also enhance "capabilities-based planning" by clarifying *when*—i.e., in what circumstances and with what assumptions about all the other factors—a given capability such as an improved weapon system or enhanced command and control will likely be sufficient or effective.[3] This is in sharp contrast to establishing a base-case scenario and an organizationally blessed model and data base, and then asking "How does the outcome change if I have more of this capability?"

2.2 Types of Uncertainty

Uncertainty, of course, comes in many forms. It is useful to distinguish between input uncertainties, which I often refer to as parametric uncertainties, and structural uncertainty.[6] Input uncertainty relates to imprecise knowledge about what the correct

---

*Correspondence: Email: pdavis@rand.org; WWW: http://www.rand.org/personal/pdavis; Telephone: 310 451 6912.
#This paper draws extensively from a book now in review.[1]

In *Enabling Technology for Simulation Science IV*, Alex F. Sisti, Editor,
Proceedings of SPIE Vol. 4026 (2000) • 0277-786X/00/$15.00

values of a model's inputs are. In contrast, structural uncertainty relates to questions about the form of the model itself. Does it reflect all the variables on which the real-world phenomenon purportedly described by the model depends? Is the analytical form of the dependencies correct? Some uncertainties may be inherent because they represent stochastic processes. Some uncertainties may relate to fuzziness or imprecision, while others reflect discord among experts. Some relate to knowledge about the values of well-defined parameters, whereas others refer to future values that as yet have no true values.

It is convenient to express the uncertainties in parametric form. Even if we are unsure about the correct form of the model, we can describe this to some extent by having one or more parameters affect that form. For example, we may have parameters that control the relative size of quadratic, and exponential terms in an otherwise linear model. Or we may have a discrete parameter that is essentially a switch determining which of a set of distinct analytical forms applies. Some parameters may apply to the deterministic aspect of a model, whereas others pertain to a stochastic aspect; and so on. In taking this approach, we need to keep straight how the different uncertainties come into play. For example, we might have a model that describes the rate at which Red and Blue suffer attrition in combat according to a Lanchester square law:

$$\frac{d\tilde{R}}{dt} = -\tilde{K}_b \tilde{B}(t) \quad \frac{d\tilde{B}}{dt} = -\tilde{K}_r \tilde{R}(t),$$

where the attrition coefficients for Red and Blue have both deterministic and stochastic parts, each of which are subject to uncertainty:

$$\tilde{K}_b(t) = K_{bo}[1 + c_b \tilde{N}_b(t;\mu,\sigma_b)] \quad \tilde{K}_r(t) = K_{ro}[1 + c_r \tilde{N}_r(t;\mu,\sigma_r)].$$

Here the N terms are normal random variables with means of μ and standard deviations σ. hТy represent stochastic processes occurring within a particular simulated war, e.g., from one time period to the next. The means and standard deviations are ordinary deterministic parameters, as are the coefficients K_{bo},K_{ro}, c_r, and c_b. These have constant values within a particular war, but at what value they are constant is uncertain.

So far the equations have represented input uncertainty. However, suppose that we don't know whether the correct equations of combat—to the extent that such equations exist—are Lanchester linear, Lanchester square, or something in between. That is, suppose we are somewhat uncertain about the structural form of the phenomenon, but are confident for some reason that the form is Lanchesterian in some sense. We could represent this as input, or parametric, uncertainty by modifying the equation to read

$$\frac{d\tilde{R}}{dt} = -\tilde{K}_b \tilde{B}^e(t)\tilde{R}^f(t) \quad \frac{d\tilde{B}}{dt} = -\tilde{K}_r \tilde{B}^g(t)\tilde{R}^h(t)$$

Now, by treating the exponents e, f, g, and h as uncertain parameters we can change the very structure of the model. Thus, by varying parameter values, we can explore both input and structural uncertainties in the model.

There are limits to what can be accomplished. For example, suppose that the correct equations of combat are Lanchesterian except for also having factors that decay exponentially with time as combatants grow weary and less efficient, or as they begin to husband ammunition. We could not explore the consequences of different exponential decay times if we do not even recognize the phenomenon. This is not an idle comment, because we often do *not* know the true underlying form of the system model: we often recognize many aspects of the phenomena, but not others. And we may not observe them unless certain unique circumstances arise.

Despite this caveat, let us now consider what can be accomplished with exploratory analysis in understanding the consequences of uncertainty to the extent that we can characterize them all with parameter values.

2.3 Types of Exploratory Analysis

Exploratory analysis can be conducted in several ways. [1,3,4,5,6,7] Although all of the methods have been used in the past, their relationship appears often to be confusing or unappreciated.

2.3.1 Input exploration

Input exploration (or *parametric exploration*) involves conducting model runs across the space of cases defined by discrete values of the parameters within their plausible domains—not one-at-a-time as in normal sensitivity analysis, and not around some presumed base-case set of values, but rather for all the combinations of values defined by an experimental design (or a smaller sample). The results of such runs, which may number from dozens to hundreds of thousands or more, can be explored interactively with modern displays. Within perhaps a half-hour, a good analyst doing such exploration can often gain numerous important insights that were previously buried. He can understand not just which variables "matter," but *when*. For example, he may find that the outcome of the analysis may be rather insensitive to a given parameter for the so-called base case of assumptions, but quite sensitive for other plausible sets of assumptions. That is, he may identify when (i.e., in which cases) the parameter is important. For capabilities-based planning for complex systems, this can be distinctly nontrivial.

2.3.2 Probabilistic exploration

A complement to parametric exploration is *probabilistic exploration* in which uncertainty about the input parameters is represented by distribution functions representing the totality of one's so-called objective and subjective knowledge. Using analytical or Monte Carlo methods, the resulting distribution of outcomes can be calculated. This can quickly give a sense for whether—all things considered—uncertainty is particularly important. In contrast to displays of parametric exploration, the output of probabilistic exploration gives little visual weight to improbable cases in which various inputs all have unlikely values simultaneously. Probabilistic exploration can be very useful for a condensed "net assessment." Note that this use of probability methods is different from using them to describe the consequences of a stochastic process within a given simulation run. Indeed, one should be cautious about using probabilistic exploration because one can readily confuse variation across an ensemble of possible cases (e.g., different runs of a war simulation) with variation within a single case (e.g., fluctuation from day to day within a single simulated war). Also, consider that an unknown constant parameter for a given simulated war is no longer unknown once the simulation begins and simulation agents representing commanders should perhaps observe and act upon the correct values within a few simulated days. Despite these subtleties, probabilistic exploration can be quite helpful.

2.3.3 Hybrid exploration

After initial work with both parametric and probabilistic exploration, the preferred approach treats some uncertainties parametrically and others with uncertainty distributions. That is, it is *hybrid exploration*. It may be appropriate, for example, to parameterize a few key variables that are under one's own control (purchases, allocation of resources, and so on), while treating the uncertainty of other variables through uncertainty distributions. One may also want also to parameterize a few of the principal variables characterizing the future context in which strategy must operate. In miilitary affairs, for example, one might parameterize assumed warning time and size of threat. There is no general procedure here; instead, the procedure should be suitable to the problem at hand. In any case, the result of such exploratory analysis can be a comprehensible summary of how known classes of uncertainty affect the problem at hand.

2.4 Examples

Without elaboration, let me give a few examples of what exploratory analysis may look like. Figure 1 mimics a computer screen during a parametric exploration of what is required militarily to defend Kuwait against a future Iraqi invasion by interdicting the attacker's movement with aircraft and missiles.[8] Each square in the figure denotes the outcome of a particular model case (i.e., a specific choice of all the input values). The model being used depends on 10 variables–those on the x, y, and z axes, and seven listed to the side (the z–axis variable is also listed there, redundantly). The outcome of a given simulation is represented by the color (or, in this paper, by the pattern) of a given square. Thus, a white square represents a good case in which the attacker penetrates only a few tens of kilometers before being halted. A black square represents a bad case in which the attacker penetrates deep into the region that contains critical oil facilities. The other patterns represent in-between cases. The number in each square gives the penetration distance in km.

To display results in this way for a sizable scenario space, RAND has often used a program called Data View, developed at RAND in the mid 1990s by Stephen Bankes and James Gillogly. After running the thousands or hundreds of thousands of cases corresponding to an experimental design for parametric exploration, we explore the outcome space by working interactively at the computer. We can choose interactively which of the parameters to vary along the x, y, and z axes of the display. The other parameters then have the values shown along the right. However, they can be changed interactively by clicking on a given number and selecting from the resulting menu. For example, we could click on the 3600 and select another value from a menu of values for cases that have been run.

As mentioned above, in about a half an hour of such interactive work, one can develop a strong sense of how outcomes vary with *combinations* of parameter values. This is obviously much more than traditional sensitivity analysis. Moreover, one can search out and focus upon the "good" cases. Figure 1 is merely one schematic snapshot of the computer screen for choices of parameter values that show some successes, which is the reason for the figure's title. Most snapshots would be dominated by black squares because it is difficult to defend Kuwait against a large threat. Although Data View is not a commercial product, RAND has made it available to government clients and some other organizations (e.g., allied military staffs).

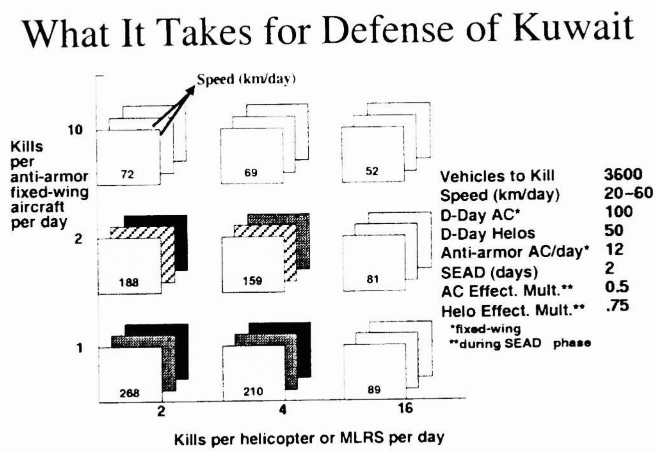

Figure 1—A Data View Display of Parametric Exploration

Other personal-computer tools can be used for the same purpose and the state of the art for such work is advancing rapidly. A much improved version of Data View called CAR™ is under development by Evolving Logic (www.evolvinglogic.com). For those who prefer modeling with Microsoft EXCEL,® there are plug-in programs that provide statistical capabilities and some means for exploratory analysis. Two such tools that I have experimented with are Crystal Ball® (www.decisioneering.com) and @Risk® (www.palisade.com/html/risk.html). For a number of reasons, however (visual modeling, array mathematics,...), I have most often used the Analytica modeling system (or a combination of it and Data View) in recent times. Analytica is an outgrowth of the Demos system developed by Max Henrion and Granger Morgan at Carnegie Mellon University.[8] It is marketed by Lumina, which has a web site at www.lumina.com.

Figure 2 shows a screen image from some recent work with Analytica on the same interdiction-of-invader-forces problem treated in Figure 1. In this case, we have a more traditional graphical display of results. Outcome is measured along the Y axis rather than by a color or pattern, and one of the independent variables is plotted along the X axis. A second variable (D-Day shooters) is reflected in the family of curves. The other independent variables appear in the rotation boxes at the top. As with Data View, we change parameter values by clicking on a value and selecting from a menu of values. Such interactive displays allow us to "fly through the outcome space" for many independent variables (parameters), in this case 9. More parameters could have been varied interactively, but the display was still quickly interactive for the given model and computer being used (a Macintosh PowerBook G3 with 256 MB of RAM).

So far, the examples have focused on parametric exploration. Figure 3 illustrates a hybrid exploration.[9] It shows the distribution of simulation outcomes resulting from having varied most parameter values "probabilistically" across an ensemble of possible wars, but with warning time and the delay in attacking armored columns left parametric. The probabilistic aspect of the calculation assumed, for example, that the enemy's movement rate had a triangular distribution across a particular range of values and that the suppression of air defenses would either be in the range of a few days or more like a week, depending on whether the enemy did or did not have air-defense systems and tactics that were not part of the best estimate. That is, if the enemy had some surprises up his sleeve, then suppression of air defenses would likely take considerably longer. We represented this possibility with a discrete distribution for the likelihood of such surprises. The two curves in Figure 3 differ in that the one with crosses for markers assumes that interdiction of moving columns waits for suppression of air defenses (SEAD). The other curve assumes that interdiction begins immediately because the aircraft are assumed stealthy.

This depiction of the problem shows in one display how widely the outcomes can vary and how the outcome distribution can be complex. The non-stealthy-aircraft case shows a considerable spike at the right end where many cases pile up because, in the simulation, the attacker halts once he has reached an objective at about 600km. Note that the mean is not a good metric: the "variance" is huge and the outcome may be bimodal or even multimodal.

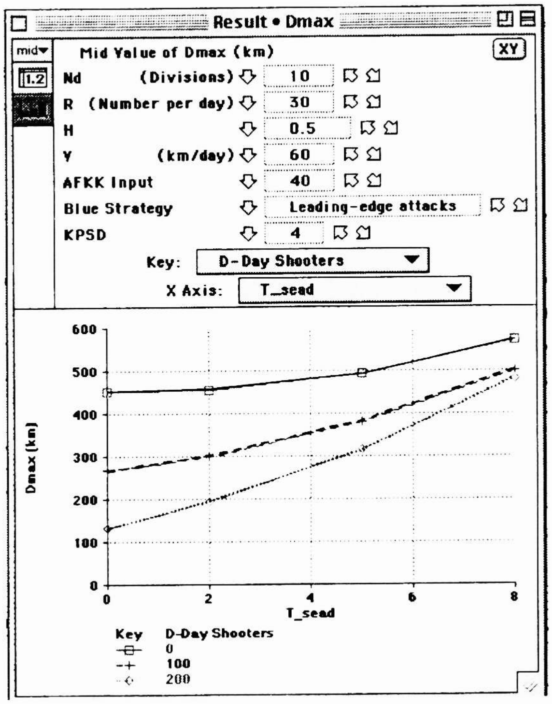

Figure 2—An Analytica Display of Parametric Exploration

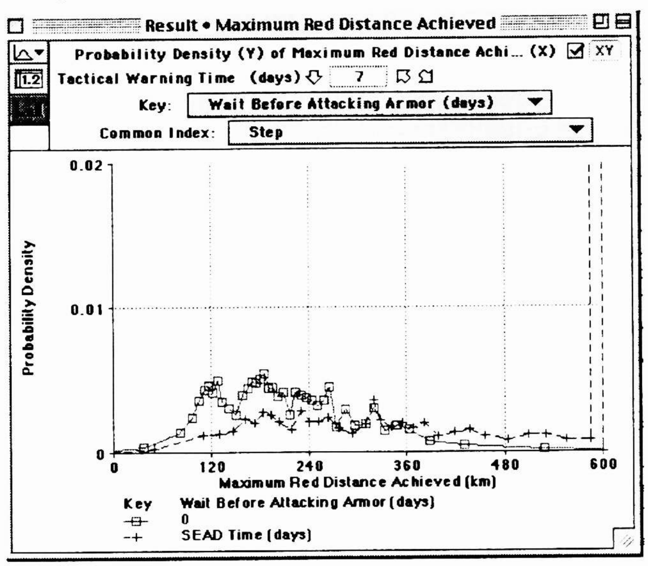

Figure 3—A Display of "Probabilistic" Exploration

2.5 Advanced Concepts

These results have been from analyses accomplished in recent years for the Department of Defense. As we look to the future, much more is possible with computational tools. Much better displays are possible for the same information and, even more exciting, computational tools can be used to aid in the search process of exploration. For example, instead of clicking through the regions of the outcome space, tools could automatically find portions of the space in which particular outcomes are found. One could then fine-tune one's insights by clicking around in that much more limited region of the outcome space. Or, if the model is itself driven by the exploration apparatus, then the apparatus could search for outcomes of interest and then focus exploration on those regions of the input space. That is, the experimental design could be an output of the search rather than an input of the analysis process. These methods are being championed and pursued by Steve Bankes and others at Evolving Logic (Bankes is also a consultant to RAND). At the core of much of their work is the evolving tool mentioned earlier called CAR.™

3. PUTTING EXPLORATORY ANALYSIS IN CONTEXT

3.1 Pursuing Analysis With Families of Models and Games

Exploratory analysis is an exciting development with a long history, including work in the 1980s and 1990s with RAND's RSAS and JICM models. Exploratory analysis, however, is only one part of a sound approach to analysis generally. It is worth pausing to emphasize this point. Figure 4 shows how different types of models and simulations (including human games) have distinct virtues. The figure is specialized to military applications, but a more generic version applies broadly to a wide class of analysis problems.

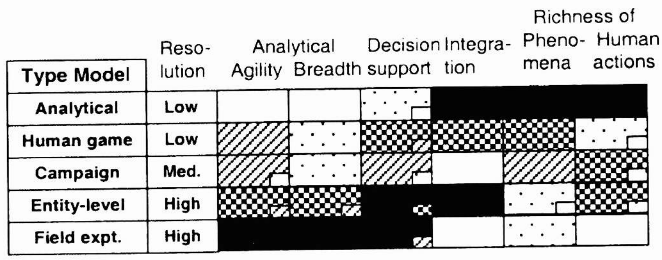

Figure 4—The Virtues of a Family of Models (Including Human Games)

White rectangles indicate "good;" that is, if a cell of the matrix is white, then the type model indicated in the left column is very effective with respect to the attribute indicated in the cell's column. In particular, analytical models (top left corner), which have low resolution, can be especially powerful with respect to their analytical agility and breadth. In contrast, they are very poor (black cells) with respect to recognizing or dealing with the richness of underlying phenomena, or with the consequences of both human decisions and behavior. In contrast, field experiments often have very high resolution (they may be using the real equipment and people), and may be good or very good for revealing phenomena and reflecting human issues. They are, however, unwieldy and inappropriate for studying issues in breadth. The small insets in some of the cells indicate that the value of the type model for the particular purpose can often be enhanced a notch or two if the models include sensible decision algorithms or knowledge-based models that might be in the form of expert systems or artificial-intelligence agents.

Figure 4 was developed as part of an exhortation to the Department of Defense regarding the need to have *families of models* and *families of analysis*.[10] Unfortunately, it is more usual for government agencies to depend more or less exclusively on a single model (e.g., such venerable models as TACWAR, BRAWLER, or JANUS). This is a serious shortcoming.[6]

3.2 The Niche of Exploratory Analysis

In the context of Figure 4 it is easy to recognize a niche for exploratory analysis: the top left hand corner of the matrix, which emphasizes analytical agility and *breadth* of analysis, rather than depth. However, the technique can be used hierarchically if one has a suitably modularized system model. That is, one can do top-level exploration first, and then zoom in to explore in more detail—but using the same techniques—the consequences of various details within particular modules. This is easier said than done, however, especially with traditional models. Specially designed models make things much easier, as discussed in what follows.

4. ENABLING EXPLORATORY ANALYSIS

4.1 The curse of dimensionality

In principle, exploratory analysis can be accomplished with any model. In practice, it becomes difficult with large and complex models. If F represents the model, it can be considered to be simply a complicated function of many variables. How can we run a computerized version of F to understand its character? If F has M inputs with uncertain values, then we could, of course, consider N values for each input, construct a full factorial design (or some subset, using an experimental design and sampling), run the cases, and thereby have a characterization. However, the number of such cases would grow rapidly (as N^M for full-factorial analysis), which quickly gets out of hand even with big computers. Quite aside from setup-and-run-time issues, comprehending and communicating the consequences becomes very difficult if M is large. Suppose someone asked "Under what conditions is F less than danger_point?" Given sufficiently powerful computers and enough time, we could create a data base of all the cases, after which we could respond to the question by spewing out lists of the cases in which F fell below danger_point. The list, however, might go on for many pages—perhaps thousands of pages. What would we do with the list? This is one manifestation of the curse of dimensionality.

4.2 The Need for Abstractions

It follows that, even if we have a perfect high-resolution model, we will need abstractions to use and understand it well. And, in the dominant case in which the high-resolution model is by no means perfect, we need abstractions that allow us to ponder the phenomena in meaningful ways, with relatively small numbers of cognitive chunks to deal with. People can reason with 3, 5, or perhaps even 10 such cognitive chunks at a time, but not with hundreds. If the problem is truly complex, it follows that we must find ways to organize our reasoning. That is, we must decompose the problem. We end up using principles of modularity and hierarchy. The need for an aspect of hierarchical organization is inescapable in most systems of interest—even though the system may be highly distributed and relatively nonhierarchical in an organizational sense.

A corollary of our need for abstractions is that we need models that use the various abstractions as inputs. It is not sufficient merely to display the abstracts as intermediate outputs (displays) of the ultimate detailed model. The reasons include the fact that when a decision maker asks a what-if question using abstractions, there is a 1:n mapping problem in translating his question into the inputs of a more detailed model. Although analysts can often respond quickly by tricking the model (if they already understand the mappings), the process can be cumbersome, slow, and treacherous. It is often better if the question can be answered by a model that accepts the abstractions as inputs.

4.3 Finding the Abstractions

Given the need for abstractions, how do we find them and how do we exploit them? A number of approaches suggest themselves, which fall into two groups.

4.3.1 When Conceiving New Models and Families

With new models, the issue is how to *design*. There are never-ending debates about how design can and should be accomplished, but several options here are as follows:

- Design the models and model families top down so that significant abstractions are built in from the start, but do so with enough understanding of the microscopics so that the top-down design is valid.[11]
- Design the models and model families bottom up, but with enough top-down insight so as to build in good intermediate-level abstractions from the start.
- Do either or both of the above, but with designs taken from different user or theoretical perspectives.

The list does not include a pure top-down or pure bottom-up design approach. Only seldom will either generate a good design of a complex system. Note also the idea of alternative perspectives. This recognizes that many abstractions are not unique. To the contrary, there are different ways of viewing what the key factors of the problem really are. For example, those in combat arms typically conceive military problems differently from logisticians, and even more differently than do those historians who seek explanation in terms of what they see as key macro drivers.

4.3.2 When Dealing With Existing Models and Families

Only sometimes do we have the opportunity to design from scratch. More typically, we must use (or adapt and use) existing models. Moreover, the model "families" we may have to work with are often families more on the basis of assertion or hope than lineage. What do we then do? Some possibilities here are:[11]

- Given existing models developed at high levels of resolution, study the model and the questions that users ask of the model to discover useful abstractions. For example, we may discover that inputs X, Y, and Z only enter the computations as the product XYZ. If so, then XYZ may be a natural abstraction. Or, to use a military example, if decision makers ask questions in terms of concepts like force strength or force ratio, then those are significant abstractions. For mature models, the obvious place to look is the list of displays that have been added over time to provide views into the internal workings of the model.
- Apply statistical machinery to search for useful abstractions. For example, if X, Y, and Z are inputs, such machinery might test to see whether the system's behavior correlates not just with X,Y, and Z, but with XY, XZ, YZ, and XYZ. If the computation in fact depends only on XYZ, that fact will show up from the statistical analysis.
- Idealize the system and develop "formal" mathematical representations (formal in the sense of being expressed symbolically without necessarily having the intention of computing the various terms and factors), a representation providing hints about the model's likely behavior. For example, such a representation might be much too complex to "solve," but it—coupled with some physical reasoning and a search for postulated simplifications—might highlight the likelihood of an overall exponential decay, or an inverse dependence on one input, or various other

nonlinearities that otherwise one might think to test for. It might suggest natural aggregation fragments such as the product XYZ mentioned above. In practice, this approach is most powerful if one considers the problem from different perspectives that suggest different but plausible simplifications. One might be that an integral is perhaps approximately equal to a representative value of the integrand times the effective width of the integration interval, and that this width "ought to be" proportional to a physical concept.

This list may seem straightforward, but it is nontrivial. The first approach is perhaps a natural activity for a smart modeler and programmer who begins to study an existing program—if, and only if, he is also a believer in higher-level depictions of the problem. If he doesn't really "believe" in the higher-level abstractions, except for the purpose of displays, he may be less useful here. The second approach seems to be favored by mathematically oriented individuals who lack enough class knowledge to take the first approach, or who believe—based sometimes on disciplinary faith—that such statistical procedures will prove successful and that those looking for more phenomenological abstractions are fooling themselves. The third approach is a hybrid. It argues that one should use one's understanding of phenomenology, and theories of system behavior, to gain insights about the likely or possible abstractions. Then, and only then, should one crank the statistical machinery. This, I believe is clearly superior to the second approach.

4.3.3 The Problem with Occam's Razor

The principle of Occam's razor requires that we prefer the simplest explanation and, thus, the simplest model. Some, particularly enthusiasts of statistical approaches, tend to interpret this principle to mean that one should minimize the number of variables in a model. They tend to focus on data (natural or simulation generated) and to avoid adding variables for the purpose of "explanation" or "phenomenology" if the variables are not needed to predict the data. In sharp contrast, subject-area phenomenologists may prefer to enrich the depiction–multiplying variables, providing a better picture of cause-effect chains, but going well beyond what can be supported with meager experimental data. My own predilection is that of the physical theorist rather than the statistician. I typically dislike repro-model and response-surface approaches unless they deal with conceptually meaningful variables and relationships—i.e., unless they stem from the last of the above approaches, rather than a more mindless application of statistical machinery.

The way out of violating the Occam's Razor principle here is to remember the longer form, which is sometimes said to be that we should adopt the simplest explanation that truly explains all there is to explain–but nothing simpler! This should include phenomena that one "knows about" even if they are not clearly visible in the limited data (e.g., historical data on who won various batttles with what overall attrition). I would add to this the admonition made decades ago by MIT's Jay Forrester to remember that to omit showing a variable one knows about may be equivalent to assuming its value to 1 (as a multiplier) or 0.

It is sometimes useful to have a competition among approaches. For example, phenomenologists working a problem may be utterly convinced that a problem must be described with complex computer programs having hundreds or thousands of data elements. In such a case, it is sometimes useful to study output behavior with a repro model approach. In one instance with which I am familiar, the result was to show that despite many man years of effort building and feeding a complex ground-force model, results were strongly dominated by a single higher-level abstraction, theater-level force ratio. The implication was not that real combat is dominated only by theater-level force ratio, but rather that various assumptions and compromises made in the development of the detailed model undercut any claims that its greater complexity and expense was adding predictive value relative to simpler models.

4.3.4 The connection between designing new models and working with existing ones

Although the discussion in §4.3.2 distinguished sharply between the case of new models and old ones, the reader may have noticed connections. In essence, working with existing models should often involve sketching what the models *should* be like and how models with different resolution *should* connect substantively. That is, working with existing models may require us to go back to design issues. If this seems suspicious, the reader might ask himself how often he has found it easier to rederive a model, and then decipher a program that he's been given, than to wade through the program on its own terms. Individuals differ but I, at least, often find it easier to engage the problem than to engage someone's else's idiosyncratically described solution. Furthermore, I then have a better understanding of assumptions and approximations.

With this background, let me now turn to the design of multiresolution, multiperspective models and families. Although this relates most directly to new models, it is relevant also to working with legacy models in preparing for exploratory analysis.

4.4 Multiresolution, Multiperspective Models (MRMPM) and Model Families

4.4.1 Definition

Multi-resolution modeling (MRM) is building a single model, a family of models, or both to describe the same phenomena at different levels of resolution,[11] *and* to allow users to input parameters at those different levels depending on their needs. Variables at level n are abstractions of variables at level n+1. MRM is sometimes called variable-or selectable-resolution modeling. Figure 5 illustrates MRM schematically. It indicates that a higher level model (Model A) itself has more than one level of resolution. It can be used with either two or four inputs. However, in addition to its own MRM features, it has input variables that can either be specified directly or determined from the outputs of separate higher-resolution models (models B and C, shown as "on the side," for use when needed. In principle, one could attach models B and C in the software itself—creating a bigger model. However, in practice there are tradeoffs between doing that or keeping the more detailed models separate. For larger models and simulations, a combination single-model/family-of-models approach is desirable. This balance needs for analytical agility and complexity management.

MRM is not enough because, as noted earlier, different applications require different abstractions even if the resolution is the same. That is, different "perspectives" are legitimate and important. Perspectives are distinguished by conception of the system and choice of variables. They are analogous to alternative "representations" in physics or engineering. Designing for both multiple resolution and multiple perspectives can be called MRMPM (pronounced Mr. MIPM).[1]

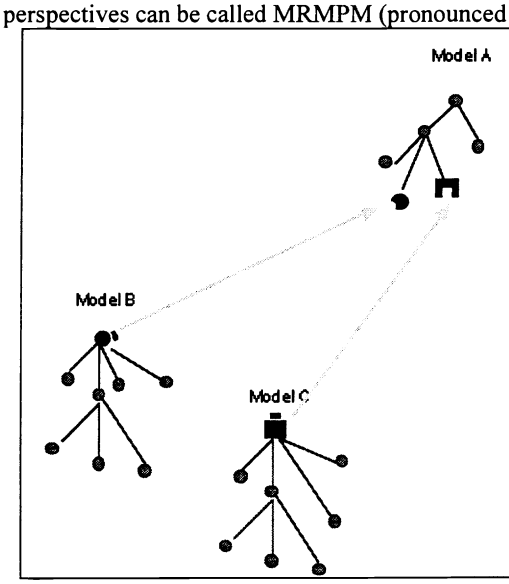

Figure 5—A Multiresolution Family of Models

4.4.2 Mutual calibration of levels or family members

Given MRMPM models (single models or families), we want to be able to reconcile the concepts and predictions among levels and perspectives. It is often assumed that the correct way to do this is to calibrate upward: treating the information of the most detailed model as correct and using it to calibrate the higher-level models. This, indeed, is often appropriate. The fact is, however, that the more detailed models almost always have omissions and shortcomings. Models at higher levels, and from different perspectives, address some of them explicitly. Further, the different models of a family draw upon different sources of information—ranging from doctrine or even "lore" on one extreme to physical measurements on a test range at the other. One class of information is not inherently better than another; it is simply different.

Figure 6 makes the point that members of a multiresolution model family should be *mutually* calibrated,[6,11] with information flows in both directions. In the military domain, for example, we may use low-resolution historical attrition or movement rates to help calibrate more detailed models predicting attrition and movement. This is not straightforward, because of the 1 to n mappings. It is often done crudely by applying an overall scaling factor (fudge factor), rather than correcting the more atomic features of the model, but it is likely to be something with which readers are familiar. On the other hand, much calibration is indeed upward. In a given study with detailed order-of-battle information, for example, inputs on the number of "equivalent divisions" or "equivalent F-15 aircraft" used in abstract models can be computed from the data feeding high-resolution models. Furthermore, at least in principle, the attrition coefficients' dependence on situation (e.g., open versus wooded terrain for ground forces) should be informed by high resolution work.

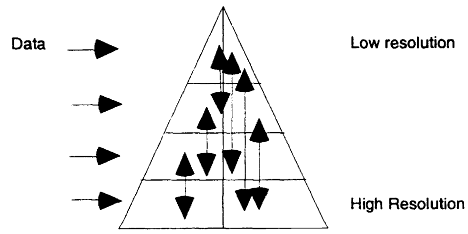

Figure 6—Mutual Calibration of a Family of Models

4.4.3 Design considerations

So, given their desirability, how do we build a family of models? Or, given pre-existing models, how do we sketch out how they "should" relate before connecting them as software or using them for mutual calibration? If we know what makes sense, then we may want to make changes in the existing models. Design issues have been treated elsewhere,[11] but I give some highlights here.

The first design principle may be to recognize that there are limits to what is feasible. In particular, there are limits to how well lower-resolution models can be consistent with high-resolution models. *Approximation is a central concept from the outset.* Several points are especially important in thinking about this:[11]

- Consistency between two models of differing resolution should be assessed in the context of how the models are being used. What matters is *not* whether they generate the same final state of the overall system, but whether they generate approximately the same results in the application. That may be something as specific as summary graphs, or rank ordering of alternatives.
- The implications for consistency of aggregation and disaggregation processes must also be judged in context. Some disaggregation assumptions represent aggregate-level knowledge not necessarily reflected in the most detailed model.
- Comprehensive MRM is very difficult or impossible for complex M&S, but having even some MRM can be far more useful than having none at all. MRM is not an all-or-nothing mater.
- The various members of an MRM family will typically be valid for only portions of the system's state space. As one moves from one region to another, valid description may require changing parameter values or even the structure of the model itself.
- Mechanisms are therefore needed to recognize different situations and shift models. In simulations, human intervention is one mechanism; agent-based modeling is another.
- Valid MRM will often require stochastic variables represented by probability distributions, not merely mean values. Further, valid aggregate models must sometimes reflect correlations among variables that might naively be seen as probabilistically independent.[6]

With these observations up front, the ideal for MRM is a hierarchical design for each MRM process, as indicated in Figure 5.

4.4.4 Desirable design attributes

From the considerations we have sketched above, it follows that models and analysis methodologies for exploratory analysis should have a number of characteristics. First, they should be able to reflect hierarchical decomposition through multiple levels of resolution and from alternative perspectives representing different "aspects" of a system.

Less obviously, they should also include realistic mechanisms for the natural entities of the system to act, react, adapt, mutate, and change. These mechanisms should reflect the relative "fitness" of the original and emerging entities for the environment in which they are operating. Many techniques are applicable here, including game-theoretic methods and others that may be relatively familiar to readers. However, the most fruitful new approaches are those typically associated with the term agent-based modeling. These include submodels that act "as the agents for" political leaders and military commanders or—at the other extreme— infantry privates on the battlefield or drivers of automobiles on the highway. In practice, such models need not be exotic: they may correspond to some relatively simple heuristic decision rules or to some well-known (though perhaps complex) operations-research algorithm. But to have such decision models is quite different from depending on scripts.

Because it is implausible that closed computer models will be able to meet the above challenge in the foreseeable future, the family of "models" should allow for human interaction—whether in human-only seminar games, small-scale model-supported human gaming, or distributed interactive simulation. This runs against the grain of much common practice.

4.4.5 When inputs to higher level models need to be stochastic

The last item in the bulleted list in §4.4.3 is often ignored in today's day-to-day work, even by good analysts who have a family of models. Often, when they seek to use models at different levels of resolution analytically, they decide upon a highest level model to be used for excursions—i.e., for examining sensitivities. They then "calibrate" this highest-level model by using one or more detailed models. For example, they might used the Brawler model of air-to-air combat between small groups of aircraft in different groupings; they would then use results of that work to calibrate the air-to-air model of a theater-level depiction such as in the TACWAR, JICM, Thunder, or START models. This is not easy. However, the analysts would sit down, talk, drew sketches, and so on, until they gained a sense of how to go about the calibration. Ultimately, for a particular study done on a limited budget and time scale (i.e., most studies), they might use expected-value outcomes of "representative" air to air engagements in Brawler to set attrition coefficients in the theater-level model. This might or might not be "correct," because the relationship between the engagement level and theater level is very complex: in a real air war, there may be thousands of engagements with a wide variety of characteristics and how to aggregate is not so clear. For example, one might imagine that 80% of engagements are "normal," but have little effect on relative force levels, while 20% of engagements are of a different character and lead to one side annihilating the other's aircraft with no losses of its own. The overall time dependence of relative force levels, then, might be dictated by the unusual, nonrepresentative, engagements. On the other hand, if one focused on them to do the calibration to theater level, then one might outrageously exaggerate one or both of the attrition rates. The "correct" way to go about the calibration would necessarily involve explicit, study-dependent, integration over classes of engagement.

Sometimes, the higher-level model inputs need to be stochastic. Figure 7 illustrates the concept schematically for a simple problem. Suppose that a process (e.g., one computing the losses to aircraft in air-to-air encounters) depends on five inputs, Q, X,Y, a, and b. But suppose that the outcome of ultimate interest involves many instances of that process with different values of X and Y (e.g., different per-engagement numbers of Red and Blue aircraft). An abstraction of the model might depend only on Q, a, and b (e.g., overall attrition might depend on only numbers of Red and Blue aircraft, their relative quality, and some command and control factor). If the abstraction shown is to be valid, the variable Z should be consistent with the higher-resolution results. However, if it does not depend explicitly on X and Y, then there are "hidden variables" in the problem and Z may appear to be a random variable, in which case so also would the predicted outcome be a random variable. One could ignore this randomness if the distribution were narrow enough, but it might not be.

To compute what $\tilde{Z}$ "should be," one would relate the probability density for $\tilde{Z}$ to a constrained integral over X and Y, appropriately weighting the on the basis of likelihood, and restricting the integration to regions where Z has the value of interest.

In the past, such calibrations have been rare—in significant part because analysts have lacked both theory and tools for doing things better. The "theory" part includes not having good descriptions of how the detailed model should relate to the simplified one. The tool part includes the problem of being able to define the set of runs that should be done (representing the integral of Figure 7) and then actually making those runs.

Ideally, such a calibration would be dynamic within a simulatiion. Moreover, it would be easy to adjust the calibration to represent different assumptions about command, control, communications, computers, intelligence, surveillance, and reconnaisance (C4ISR), as well as tactics. We are nowhere near that happy situation today,

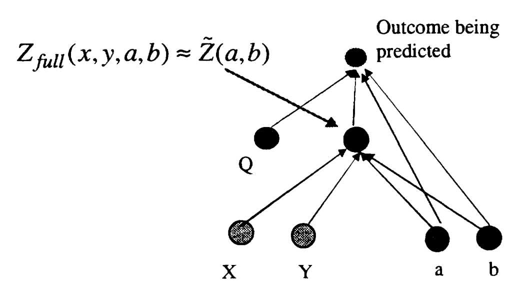

Figure 7—Input to Higher Level Model May Be Stochastic

5. RECENT EXPERIENCE AND CONCLUSIONS ABOUT NEEDS

Over the last several years, my colleagues and I have done considerable work related to the problem of halting an invading army using precision fires from aircraft and missiles.[8,10,11,12,13] The most recent aspects of that work included understanding in some detail how the effectiveness of such fires are affected by details of terrain, enemy maneuver tactics, certain aspects of command and control, and so on. This provided a good test bed for exploring numerous aspects of MRMPM theory.

A companion piece to this paper[14] describes how we developed a multiresolution personal-computer model (PEM), written in Analytica, to understand and extend to other circumstances the findings from entity-level simulation of ground maneuver and long-range precision fires. A major part of that work was learning how to inform and calibrate PEM to the entity-level work. There was no possibility, in this instance, of revising the entity-level model. Nor, in practice, did we have such a good understanding of the model as to allow us to construct a comprehensive calibration theory. Instead, we had to construct a new, more abstract, model and attempt to impose some of its abstractions on the data from runs of the entity-level simulation in prior work, plus some special runs made for our purposes. Had we had the intermediate-level PEM model several years earlier, we could have used it both to define adaptations of the entity-level model that would have generated some of the abstractions we needed and to better define the experiments conducted with the high-resolution model. Instead, we had to make do with the situation we found ourselves in. The result is a case history with what are probably some generic lessons learned.

Figure 8 illustrates one aspect of our multiresolution approach in PEM. The figure shows the data flow within a PEM module that generates the impact time (relative to the ideal impact time) for a salvo of precision weapons aimed at a packet of armored fighting vehicles observed by C^4ISR assets at an earlier time. Other parts of PEM combine information about packet location versus time and salvo effectiveness for targets that happen to be within the salvo's "footprint" at the time of impact, to estimate effectiveness of precision weapons. For the salvo-impact-time module, Figure 8 shows how PEM is designed to accept inputs as detailed as whether there is enroute retargeting of weapons, the RSTA latency time, and weapon flight time. However, it can also accept more aggregate inputs such as time from last update. If the input variable Resolution of Time of Last Update Calculation is set "low," then Time From Last Update is specified directly as input; if not, it is calculated from the lower-level inputs.

Being able to depict the problem as in Figure 8, and to provide users the option of what inputs to use, has proven very useful—both for analysis itself and for communicating insights to decision makers in different communities ranging from the C4ISR community to the programming and analysis community. In particular, the work clarified how the technology-intensive work of the C4ISR acquisition community relates to higher-level strategy problems and analysis of such problems at the theater level.

Another companion piece[15] describes how, in developing both PEM and a yet more abstract model (EXHALT) that we use for theater-level halt-problem analysis, we experimented with a variety of methods to deal with the multi-perspective problem. Perhaps the key conclusion of that work is that MRMPM work rather demands a building-block approach that empasizes study-specific assembly of the precise model needed. Although we had some success in developing a closed MRMPM model with alternative user modes representing different demands for resolution and perspective (e.g., the switches in Figure 8), it proved impossible to do very much in that regard: the number of interesting user modes and resolution combinations simply precludes being able to wire in all the relevant user modes. Moreover, that explosion of complexity occurs very quickly. Thus, despite the desire of many users to have a blackbox machine with can handle all the cases and perspectives of interest, it seems a fundamental reality that at-the-time-assembly from building blocks, not prior definition, is the stronger approach. This was as we expected, but even more so. As an aside, this finding maps directly into conclusions about how best to do adaptive operations planning (Ref. 2, Ch. 4).

The ultimate reason for our conclusion is that even in the relatively simple problem we examined, the real variable trees (akin to data-flow diagrams) are bushy rather than rigorously hierarchical. Furthermore, the different legitimate perspectives can simply not all be accommodated simultaneously without making the code itself very complicated to follow. In contrast, we found it easy to construct the model needed quickly—in hours rather than days or weeks—as the result of our building-block approach, visual modeling, use of array mathematics, and strong, modular, design.

Despite our overall good feeling about the experience, we also concluded that current personal computer tools—as powerful as they are in comparison with those in past years—are not yet up to the challenge of making the building-block/assembly approach rigorous, understandable, controllable, and reproducible without unrealistically high levels of modeler/analyst discipline. Thus, there are good challenges ahead for the enabling–technology community. Also, the search models for advanced exploratory analysis are not yet well developed.

In closing, I should note that there is a good deal of related work on MRM going on in the community (see, e.g., Refs. 16-18). This paper has not attempted to draw the connections, although the RAND publications on which this one draws do so.[1]

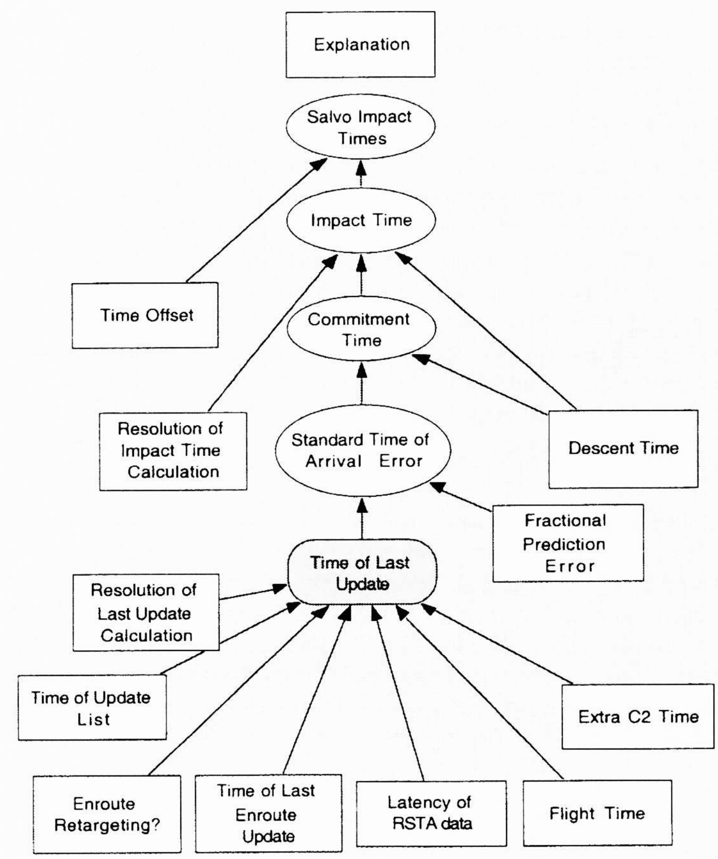

Figure 8—A Multiresolution, multiperspective design for the salvo-impact-time module

ACKNOWLEDGMENTS

This paper draws on work accomplished for the Office of the Secretary of Defense in several projects over the period 1997-2000. Reports on those projects are included in the references below. The paper benefited from discussions with James Bigelow and Jimmie McEver.

REFERENCES

1. Paul K. Davis and Richard Hillestad, *Exploratory Analysis for Strategy Problems With Massive Uncertainty*, RAND, Santa Monica, CA, draft book, 2000.

2. Paul K. Davis (ed.), *New Challenges in Defense Planning: Rethinking How Much Is Enough*, RAND, Santa Monica, CA, 1994

3. Paul K. Davis, David Gompert, and Richard Kugler, *Adaptiveness in National Defense: the Basis of a New Framework*, RAND Issue Paper, Santa Monica, CA, 1996.

4. Stephen C. Bankes, "Exploratory Modeling for Policy Analysis," *Operations Research*, Vol. 41, No. 3, 1993.

5. Robert Lempert, Michael E. Schlesinger, and Steven C. Bankes,"When We Don't Know the Costs or the Benefits: Adaptive Strategies for Abating Climate Change," *Climatic Change*, Vol. 33, No. 2, 1996.

6. National Research Council, *Modeling and Simulation*, Volume 9 *of Technology for the United States Navy and Marine Corps, 2000–2035*, National Academy Press, Washington, D.C., 1997.

7. Granger Morgan, and Max Henrion, *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis*, Cambridge University Press, Cambridge, Mass., 1992. Reprinted in 1998.

8. Paul K. Davis and Manuel Carrillo, *Exploratory Analysis of the Halt Problem: A Briefing on Methods and Initial Insights*, RAND, DB-232, Santa Monica, CA, 1997.

9. Paul K. Davis, David Gompert, Richard Hillestad, and Stuart Johnson, *Transforming the Force: Suggestions for DoD Strategy*, RAND Issue Paper, Santa Monica, CA, 1998.

10. Paul K. Davis, James A. Bigelow, and Jimmie McEver, *Analytic Methods for Studies and Experiments on Transforming the Force*, RAND, Santa Monica, CA, 1999.

11. Paul K. Davis and James A. Bigelow, *Experiments in Multiresolution Modeling*, RAND, Santa Monica, CA, 1998.

12. Jimmie McEver, Paul K. Davis, and James A. Bigelow, *EXHALT: an Interdiction Model for the Halt Phase of Armored Invasions* RAND, Santa Monica, CA, 2000.

13. Paul K. Davis, James A. Bigelow, and Jimmie McEver *Effects of Terrain, Maneuver, Tactics, and C4ISR on the Effectiveness of Long Range Precision Fires: a Stochastic Multiresolution Model (PEM) Calibrated to High-Resolution Simulation* (RAND, Santa Monica, CA, 2000.

14. James A. Bigelow, Paul K. Davis, and Jimmie McEver, "Case History of Using Entity-Level Simulation as Imperfect Experimental Data for Informing and Calibrating Simpler Analytical Models for Interdiction," *SPIE 2000*, Orlando Florida, April, 2000.

15. Jimmie McEver, Paul K. Davis, and James Bigelow, "Implementing Multiresolution Models and Families of Models: From Entity Level Simulation to Personal-Computer Stochastic Models and Simple 'Repro Models'", *SPIE 2000*, Orlando, Florida, April, 2000.

16. Paul Fishwick and Kangsung Lee, "Dynamic Model Abstraction," *Proceedings of the 1996 Winter Simulation Conference*, pp 764-771, San Diego, CA, 1996.

17. Douglas Popken, "Application of System Identification Techniques to Simulation Model Abstraction," (*Proceedings of SPIE*, Enabling Technology for Simulation Science III, Volume 3696, 1999

18. Bernard P. Zeigler, Tag Gon Kim, and Herbert Praehofer, *Theory of Modeling and Design* , Academic Press, San Diego, 1999.