Proceedings of the 1993 Winter Simulation Conference
G.W. Evans, M. Mollaghasemi, E.C. Russell, W.E. Biles (eds.)

FAMILIES OF MODELS THAT CROSS LEVELS OF RESOLUTION: ISSUES FOR DESIGN, CALIBRATION AND MANAGEMENT

Paul K. Davis
Richard Hillestad

RAND and the RAND Graduate School
1700 Main Street; Santa Monica, CA 90407-2138

ABSTRACT

This review paper summarizes an ARPA-sponsored project to study variable-resolution modeling (VRM) and the connection of models across levels of resolution. We describe work introducing basic concepts, highlighting a design approach called integrated hierarchical variable-resolution modeling (IHVR), exploring mathematically some long-standing issues of aggregation in ground-combat modeling, and experimenting with cross-organizational efforts to develop and compare models created with different techniques and perspectives. We also describe highlights of a conference held in May, 1992 and touch briefly on some subsequent work. Despite the substantial progress, we conclude that the VRM problem is central, difficult, and greatly under invested. It should be approached as a matter of military science rather than technology, although we discuss new software tools that can be quite valuable.

1 INTRODUCTION

Advances in technology are now providing enormous computer power and the capability to link computer programs running in distant locations. Research on virtual-reality systems is beginning to bear fruit even though we are still on the frontiers of this domain. For example, as discussed in Defense Modeling and Simulation Office (1992) and Defense Science Board (1993), a revolution is occurring in the way in which military training and exercises are conceived and conducted and there are major changes occurring also in analysis in support of R&D and acquisition decisions.

Unfortunately, technological progress has now gotten out ahead of our understanding of underlying phenomena and of how to represent those phenomena in models and computer programs. Based on concerns expressed about this in earlier work (Bankes, 1992, Davis and Blumenthal, 1992), ARPA funded a small project at RAND to begin investigating one of the central challenges currently limiting progress in advanced simulation, the problem of variable-resolution modeling (VRM) and the related problem of how to connect models across levels of resolution. Great strides have been made in the "software challenge" of connecting dissimilar simulations (e.g., simulations differing in language, time step, and the representation of common elements) (Weatherly, et. al 1991, Brendley and Marti, 1993, and Cohen et. al, 1993), but these do not address the "correctness" of the information provided from one model to another, nor whether model assumptions and operations can be understood. In fact, connecting models across levels of resolution raises complex issues of phenomenology and mathematics, many of which relate to aggregation and disaggregation. In this paper we review work presented at a conference (Davis and Hillestad, 1993a) sponsored in May 1992 by ARPA and the Defense Modeling and Simulation Office. Portions of the material here were published first in Davis and Hillestad (1993b). Other references relevant to the subject include Courtois (1985), Innis and Rexstad (1983), and Sevinc (1990).

2 BASIC CONCEPTS AND DEFINITIONS

To introduce variable-resolution modeling (VRM) and the related issue of developing integrated families of models, we will address four key questions (see Davis, 1993): (1) What *is* variable-resolution modeling and when can it be valid? (2) Why might one want it? (3) What forms can it take and how does it relate to "families of models"? and (4) How should one go about it?

2.1 Definitions

Variable-resolution modeling is building new models or model families so that users can change readily the resolution at which phenomena are treated. *Cross-resolution model connection* is linking *existing* models with different resolutions. *Seamless design* is design that permits changing resolution with (a) smooth consistency of representation (description) and (b) consistency of prediction.

"Seamlessness" in this context means that when one changes resolution, either by "zooming" within a single model or by moving from one model to another within a family, one can do so without mental disruptions and with some confidence that the results are consistent in a sense to be discussed later. We cannot aspire to continuously variable resolution analogous to the zooming of a camera, but we *can* aspire to making discrete changes in resolution that are easy to follow and, within limits, valid.

The reason for discussing variable-resolution modeling and cross-resolution model connection together is that cross-resolution model connection can often be accomplished best by standing back and pretending to have the luxury of starting over again. Knowing what one would *like* to have makes it easier to make a set of coherent model adaptations rather than a series of patches.

2.2 What is Resolution?

Resolution is a multifaceted concept, as indicated in Figure 1. Using a military example to make the distinctions, higher *entity resolution* might mean following units as small as battalions rather than divisions; higher *attribute resolution* might mean following the number of various weapons held by each battalion rather than merely assigning the battalion a net "strength;" higher *logical-dependency resolution* might mean including constraints on the attributes and their interrelationships (e.g., the sum of the men in the units comprising a division should equal the number of men in the division). Higher *process resolution* might mean computing combat attrition at the battalion level, rather than computing it at the division level and then spreading the attrition equally across the division's battalions. Higher *spatial and temporal resolution* means using finer scales for space and time. Other mathematical facets of resolution are described in Palmore (1993) (e.g., rounding from 3.14 to 3 is an aggregation operation).

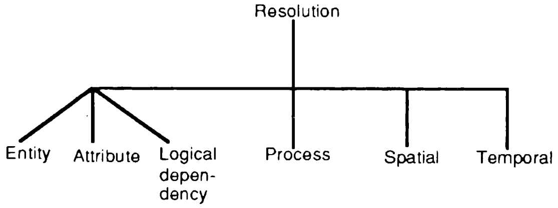

Figure 1: Aspects of Resolution

It is typically meaningless to talk, then, about "the" relative resolution of two models, because unless they have been designed by VRM principles, the resolution relationships between them are likely to be complex and confusing, with one model's resolution being higher in some respects, lower in others, and the same in still others (see also Harshberger, et. al, 1993).

Resolution is closely related to the issues of aggregation and disaggregation, although the relationship is imperfect since not all low-resolution models are aggregations of higher resolution models. In discussing VRM, however, we are specifically interested in models that *are* coherently related to each other by aggregation/disaggregation relationships, which correspond to the logical relationships "has a" or "is a component of," respectively (Eddy, 1993, Palmore, 1993).

2.3 What is a "Valid" or "Consistent" Aggregation?

When we change resolutions, replacing a detailed model by a more aggregate one, we must ask whether results are "consistent." Figure 2 illustrates the issues. Consider a system in state S(T1) at time T1. Assume a detailed model represented by an abstract time generation operator denoted **G(T2;T1)** and such that S(T2) = **G(T2;T1)**S(T1) (i.e., the state at time T2 is "generated" from the state at time T1). Suppose s(T1) is an aggregate state of the system denoted, where s(T)=**AGG** S(T). For example, a division's state might be computed by summing appropriately the states of the division's component units. Assume a disaggregation operator **DISAGG**, which purportedly generates the detailed state from the aggregate state (i.e., S(T)=**DISAGG** s(T)). Finally, assume an aggregate model represented by the operator **g(T2;T1)**, which purports to describe the time dependence of the aggregate state. Are the detailed and aggregate models "consistent?"

We consider the models to be *consistent in the aggregate* if one can move clockwise or counterclockwise from S(T1) (top left) in getting to the final aggregate state s(T2) (bottom right), and still get the same result. Formally, this requires:

$$\mathrm{AGG}\ G(T2;T1) = g(T2;T1)\ \mathrm{AGG}.$$

If, in addition, we can obtain the correct detailed state S(T2) by moving clockwise or counterclockwise around the diagram, then we say there is *complete consistency.* This requires:

$$G(T2;T1) = \mathrm{DISAGG}\ g(T2;T1)\ \mathrm{AGG}.$$

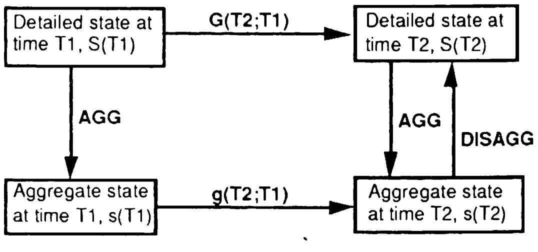

Figure 2: Defining Consistency

As an example, suppose the detailed state describes strengths of a division's battalions. The aggregate state might be division strength computed as an appropriate sum over battalion strengths. Given a division's strength, one might approximate the detailed state by assuming the component battalions have equal shares of division strength. That might or might not be reasonable.

Similarly, the detailed model might assess attrition of each separate battalion, while the aggregate model might estimate division-level attrition from other features of the aggregate state (e.g., the ratio of attacking and defending divisions). Summing battalion attritions might or might not agree with attrition generated by the aggregate model.

By and large it is hard enough to achieve consistency in the aggregate without demanding complete consistency, especially since aggregation eliminates information. Nonetheless, complete consistency is sometimes possible because the information discarded by aggregation is irrelevant to details of the subsequent state. For example, the lay-down of forces in tomorrow's battle depends more on the circumstances of that battle and doctrine than on the lay-down of forces in last week's battle in a different location.

Finally, a caveat. Different definitions are needed when attempting to understand the validity of aggregation for describing averages over similar systems (e.g., a range of future battles with somewhat different scenario assumptions). An aggregation useful for acquisition decisions might be deplorable for a commander's operations planning in a given war.

2.3 Why Do We Want VRM?

Why do we even need variable resolution? Why not just "do everything right" (at high resolution)? In fact, we need low-resolution modeling to: (a) make initial cuts at problems; (b) "comprehend the whole;"(c) conduct analysis under uncertainty; (d) enhance flexibility; (d) reduce costs and speed analysis; and (e) make use of low-resolution knowledge and calibrate higher resolution models.

We need high-resolution models: (a) to understand underlying phenomena; (b) represent detailed knowledge; (c) simulate "reality," and (d) use high-resolution knowledge and to calibrate lower-resolution models.

Importantly, we will need both low- and high-resolution models (actually, models of many resolutions) even as computing power becomes seemingly infinite. This is well understood by analysts who have dealt with higher level issues amidst the massive uncertainty characteristic of those issues (e.g., for a general, what strategy to employ without knowing that of his enemy). It is part of our *bounded rationality* (Simon, 1981). Also, much of our knowledge of the world comes not in the form of detailed "data" of the form needed by high-resolution models, but in more aggregate forms that require aggregate models for interpretation. In general, wise analysts work with models of varied resolution, moving up and down among them to achieve a best overall consistent understanding of the phenomena.

2.5 How Might One Go About It (VRM)?

Given an interest in VRM, how can we achieve it? There are three principal approaches: (a) *selected viewing*, (b) alternative sub models (or model families), and (c) *integrated hierarchical variable resolution (IHVR).* The first is not really variable-resolution *modeling*, because it depends on a single high resolution model, but it includes some "zooming" and "unzooming" capability by providing displays of varied resolution. Consistency is guaranteed, because the low-resolution variables are literally aggregations of high-level state variables.

This approach has many shortcomings, including the complexity of the underlying model and the inability to "understand" the model simply. There can also be peculiar phenomena at the aggregated level because of "hidden-variable effects." And, as anyone who has tried to use this approach analytically knows, it is difficult to respond well to "What if?" questions by policy makers, because the questions are conceived in aggregate terms. To translate the question into something a complex model can deal with requires many non-unique and dubious assumptions about how to disaggregate.

The alternative sub model approach is relatively common and consists of providing two or more versions of a given sub model so that users can decide at the time how much resolution to use. In a military problem one might, for example, have a sub model estimating attrition of aircraft that consists of nothing more than a series of probabilities of penetrating various defense zones. The individual probabilities would be input data. An alternative sub model might describe the geographical location of defensive systems and airbases and simulate the interactions of an individual aircraft with individual defense batteries as it flies through its mission.

This approach is often very useful, but in practice it seldom lives up to the hopes expressed for it by designers because: the models often reflect different views of the world with no clear-cut relationship between them, much less an ability to zoom or calibrate; and the data for the alternative models may not be consistent (even if there is an understanding about what consistency entails). We shall describe the third method in what follows. When it is feasible, it is a superior approach.

2.6 Integrated Hierarchical Variable Resolution Modeling (IHVR)

At this point we need to make an important distinction that appears to be seldom recognized. If one wants to have variable resolution in entities, attributes, and logical-dependency, then object-oriented methods can help greatly (see, e.g., Rumbaugh, et. al, 1991; Zeigler, 1991, 1993). This is quite useful in military modeling where, for example, an army corps breaks down hierarchically into divisions, brigades, battalions, companies, and platoons. By contrast, there are few general methods and tools available to help deal with variable resolution in hierarchies of *processes* (e.g., in combat modeling, the processes of attrition and movement). Many software-engineering methods are quite relevant (e.g., hierarchical data-flow diagrams), but they do little to help cope with the deep phenomenological and mathematical issues of variable process resolution (Davis, 1993; Harshberger, et. al, 1993). The techniques we describe here appear to be unusual and stem from early work with political models (Davis, Bankes, and Kahan, 1986) and more recent applications of the same methods to combat models (Davis and Huber, 1992) and theories of decision making in crisis (Davis and Arquilla, 1992). Perry, Schrader and Allen (1993) also use hierarchical methods.

To illustrate IHVR, consider two ways to model one of the simplest of military problems, calculation of the "damage expectancy" DE to a target (essentially the probability of "destruction" by some criterion) due to attacking it with n weapons from a launch-platform at a range R. The weapons have an accuracy described by a range-dependent CEP, a reliability r, an explosive yield Y, and a height of burst HOB. The target's response to weapon detonations is described by a "VN number," a more sophisticated concept than simple target hardness. Figure 3a (a data-flow diagram) shows diagrammatically how one might model the problem non hierarchically; here DE is given by a somewhat complex function of the many variables. To calculate DE one needs to know all of them, as well as the function F and its parameters.

By contrast, with the design of Figure 3b one can do an "aggregated" calculation of the damage process using number of weapons and the single-shot kill probability SSPK as inputs, or a higher resolution calculation in which one assesses the weapon's CEP and lethal radius and then uses those to assess the SSPK. In the more detailed model one has high-resolution input variables such as CEP(R), Range, Yield, VN, and height of burst. The asterisks indicate which variables can be specified as inputs or calculated from below (if there are lower-level variables). Note that "at a glance" one understands the relationship between the high and low-resolution processes. Obviously, if one specifies SSPK as an input, perhaps having looked it up in a table of weapon characteristics, it "should be" an average value of SSPK calculated over an appropriate distribution of possible CEPs and lethal radii, which, in turn, should be generated from appropriate distributions of CEP(R), range, yield, VN, and height of burst. It should not merely be the value for a so-called "representative case." What distributions are "appropriate" would depend on the application.

This type of model provides true variable-resolution capability. Further, it can be "seamless" in the sense defined earlier. If the calibrations are accomplished properly and the variable names chosen wisely, and especially if there are diagrams such as this to help explain, the relationship between one level of resolution and another is obvious and can be made rigorous.

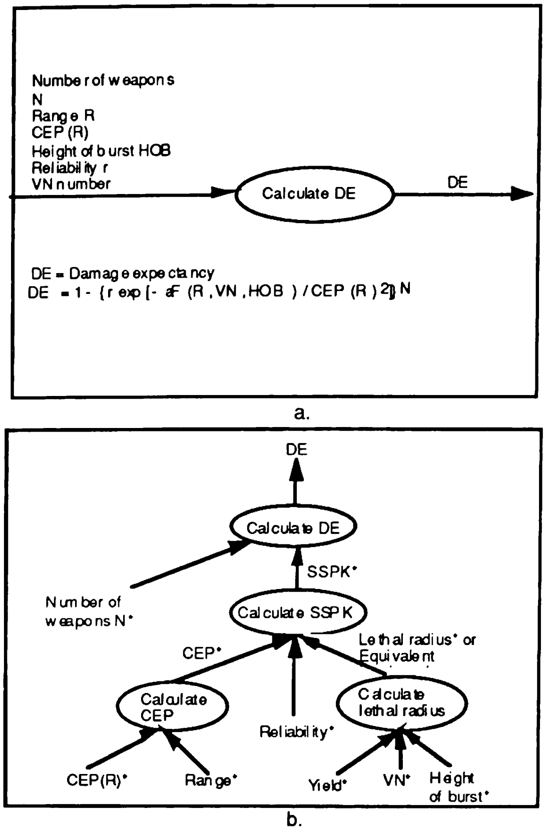

Figure 3: Non hierarchical and Hierarchical Designs of a Damage-Process Model

To be sure, Figure 3 describes an exceedingly simple case, but there are instances throughout nature and through man-dominated phenomena such as combat in which hierarchies of processes can be constructed. They may not be exact, because of interactions among different branches of what is supposed to be a hierarchical tree, but if the interactions are weak, the hierarchical depiction can be very good (Simon, 1981).

Figure 4 shows a combat model (Huber, 1992) for assessing arms control proposals in the German government. The highest level variable is a force ratio in the main sector of an attack. That "force ratio" is really a measure of relative capability, taking into account the attacker's and defender's ground forces and air forces. Indeed, this model computes ground-force and air-force potential for the two sides and then combines them before taking a ratio. Each of those potentials is, in turn, computed from higher resolution variables down to the level of range-payload relationships for aircraft (bottom right-hand corner). As before, asterisks indicate where the variable in question can be specified directly rather than computed from below. Thus, one can select resolution level and, if one wants to calibrate a given parameter value, one knows that to do so one must compute an appropriate average using lower-level (higher resolution variables) in a distribution of cases.

Models designed in this way lend themselves well to rigor, maintenance, and certain types of flexibility. Note also that it makes no difference theoretically whether there is literally one model with multiple levels, or whether the lower levels are actually separate models of a multi-resolution *family*. To put it differently, when one can draw diagrams such as Figures 3 and 4, and write down the corresponding details, one is entitled to claim to have an *integrated* design of a single model or a family of models. Although there has been some excellent work with families of models (e.g., Scheckeler, 1993 and work by Air Force Studies and Analysis on air-to-air combat), we are aware of very few families that were designed as we describe and recommend.

Unfortunately, IHVR is no panacea, even when generalized to have many "top-level variables." First, the real-world processes of interest are not always related hierarchically: the relationship may be more like a network than a tree. Second, choosing a particular process hierarchy is choosing a particular *perspective* on the problem and one cannot always switch perspective readily once the model is constructed. For example, a model that conceives movement rate as due to an effective force ratio cannot be connected easily to a perspective that sees movement as a sequence of volitional moves after winning, losing, or withdrawing from particular battles. See also Harshberger, et. al (1993) for good air force examples. Third, the aggregations may be poor approximations in any single scenario or battle, even though useful for describing averages over outcomes of many scenarios. As an example, when an attacker engages a well prepared defender at a 3:1 force ratio, the attacker will typically win or lose, but not break even. By contrast, the *deterministic* aggregated model will predict a break-even outcome. A stochastic aggregated model would be bimodal. In a similar way, some aggregate models are good representations of system behavior over time, but not during a particular "short" period.

Despite these cautionary notes, we recommend IHVR methods whenever possible, and suggest active efforts to find methods in which they will be good approximations of reality (e.g., methods that exploit natural separations of time scale and phases of phenomena). This implies increasing the dosage of "theory" in modeling, and backing away from working "linearly" at the computer. We expect that many models will use combinations of IHVR and alternative sub models. It's not an all-or-nothing choice.

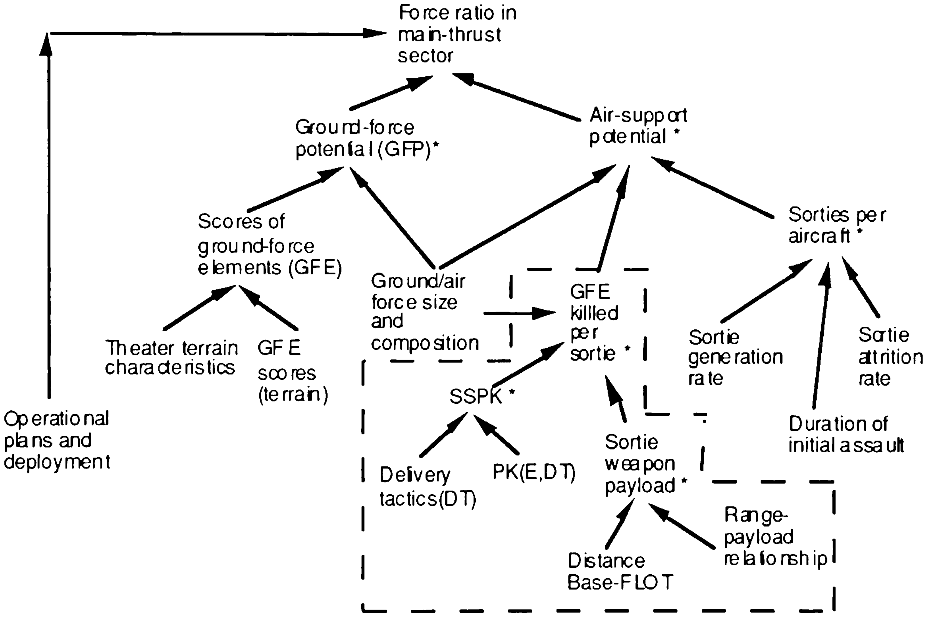

Figure 4: An IHVR Model Used in Arms Control Assessments of Capability

2.7 Why Is Linking Models So Troublesome?

One purpose of our work has been to clarify the generic reasons that workers encounter so many problems when they connect existing models of varied resolution. Davis (1993) uses a simple ground-combat problem to describe in some detail, at the level of equations and reproducible spreadsheet programs, a hypothetical case history of two models developed independently and then connected with the expectation that one could serve as the high resolution version of the other and thereby provide calibration. It illustrates why comparing and relating two models often proves to be confusing and error-prone, even though they appear initially to be compatible. Some of the problems involve deeply buried definitional issues (e.g., variables with a given name may mean substantially different things in the two models), conceptual issues (e.g., whether the "mountainous terrain" referred to in the model is a description of average terrain over a front of 100 km or over the 10-20 kilometers of front on which battle actually focuses, which might be gently rolling terrain in a narrow valley of very mountainous terrain), failure to define calibration relationships as appropriate statistical *averages* over cases sampling large domains of the more detailed variables, working with ill-conceived aggregation relationships, and attempting to use deterministic aggregated models when stochastic phenomena are fundamental.

The conclusion of this work on connecting models was that models should not be connected without a prior design study by people capable of understanding the theoretical and mathematical issues involved. Further, it will often prove to be the case that the models should not be connected without substantial redesign and that in some instances it would be better to "start over." With modern model-building technology, "starting over" may be much less expensive than managers with intuitions developed twenty years ago might expect.

3 THEORETICAL ISSUES IN AGGREGATION

In this section we touch briefly on a number of theoretical efforts to better understand aggregation issues.

3.1 Aggregation in Lanchester Equations

Having concluded that one aggregation-related problem was that modelers too often leap to programming before doing adequate mathematical analysis of the phenomenology, we decided to do an experiment to see what a dose of theory might buy in a long-standing and bitterly fractious dispute within the combat modeling community. Briefly, the dispute is between those who use relatively high resolution ground-combat models that describe combat as a matrix interaction of various categories of weapons (e.g., TACWAR, CEM, TLC/NLC) and those who use much simpler models that characterize the antagonists' forces as having scalar "strengths" or scores (e.g., ATLAS, MASTER, TACSAGE). Some of the former have argued vituperously for years that the aggregated approaches are obviously and fatally flawed; the latter have tended to argue that the simpler models are easier to understand and manipulate, and that they are arguably just as valid as the detailed models, given uncertainties about the "data" of the detailed models.

The problem we posed is as follows. Let **A** and **D** denote vectors of attacker and defender strength respectively. Each component of the vectors represents capability for a given class of weapons such as tanks or artillery. The "detailed model" is a set of "heterogeneous Lanchester equations" as shown below. The ith component of the attacker force strength suffers attrition due to the effects of the 1st, 2nd...and nth components of the defender force through the elements of the killer-victim matrices **K** and **K'**. Thus, K_{ij} is the rate at which a unit of the defender's ith class of weapons kills the jth class of attacker weapons.

$$\begin{aligned} &\mathbf{A} = \{A_1, A_2, \ldots A_n\} &\quad& \mathbf{D} = \{D_1, D_2, \ldots D_n\} \\ &d\mathbf{A}/dt = -\mathbf{KD} && d\mathbf{D}/dt = -\mathbf{K'A} \\ &\mathbf{K} = \{K_{ij}\} && \mathbf{K'} = \{K_{ij}\} \end{aligned}$$

If we assume for the sake of discussion that these equations are valid, then the controversial issue is whether a more aggregated model can also be valid. That is, is there a linear aggregation of **A**, α, which obeys a Lanchester equation, and similarly for the defender:

$$\begin{aligned} &\alpha = \sum_i f_i A_i &\quad& \delta = \sum_i g_i D_i \\ &d\alpha / dt = -k\delta && d\delta / dt = -k'\alpha \end{aligned}$$

This problem, simple as it may seem to be, can be solved only with rather sophisticated matrix algebra (Hillestad and Juncosa, 1993). The theoretical effort was worthwhile, however, because the conclusions were not obvious to us or to others to whom we presented them in 1992. The principal conclusions were:

- The models can be "consistent in the aggregate," but if and only if the aggregations are defined in terms of the eigenvectors of the matrix **KK'**. This means that the weighting factors f_1...and g_1...must be particular functions of the matrix elements of both attacker and defender.

- Partial aggregations in which one reduces the dimensionality of the problem are also possible, but only in special cases when the effectiveness of successive weapon classes is a multiple of the effectiveness of the first weapon class.

- Complete consistency is possible (i.e., one can disaggregate to obtain the detailed state) if, in addition, the vulnerability of successive weapon classes are simple multiples of the vulnerability of the first weapon class. An example would be situations in which attrition is spread more or less evenly among components of a unit.

These results are mathematically exact and say nothing about how precisely the conditions must be met for the aggregation approximations to be "reasonably" accurate, but they are interesting in themselves because: (a) they refute the notion that the aggregation is "obviously wrong," (b) provide some insight about the kinds of circumstances in which the aggregation can be correct, and

(c) encourage us to find regimes (e.g., short periods of time with constant conditions) in which the conditions are met approximately. This is crucial, because aggregations are essential in all military simulations, even those thought of as detailed (e.g., in physics-level models, all the tanks may be treated as identical, even though in the real world there might be a mix of somewhat different tanks). At the same time, those who expect an "intuitive" aggregation to be valid are often going to be disappointed. In particular, the common aggregation approach based on assigning strengths or scores to forces without considering the particular opponent will fail unless actual opponents are close to the "average opponents" those postulating the weighting factors had in mind.

3.2 The Ubiquitous Configuration Problem

One of the most interesting challenges to variable-resolution modeling is the difficulty in finding accurate aggregations in problems in which the spatial configuration of objects is important (e.g., the disposition of personnel, equipment, and units on the battlefield). Horrigan (1993) describes factor-of-five-and-ten errors in estimates of weapon-system capability when superficially plausible aggregate models are used instead of more microscopic stochastic models. Horrigan refers to this as the *configuration problem.* Significantly, Horrigan's work does not demonstrate that aggregation is impossible or bad. Rather, it illustrates forcefully how many common forms of aggregation (e.g., deterministic Lanchester equations, which omit any reference to configuration) can introduce large systematic errors. In Horrigan's words:

"Configuration...determines which targets can be encountered by which weapons, in what order, at what ranges, and with what velocities and orientations. The detection probabilities, acquisition probabilities, hit probabilities and damage probabilities associated with a particular target-weapon encounter, of course, are each a function of the kind, state, and position of the target and the kind, state, and position of the weapon (and other variables)...Numerous stochastic processes and the associated random variables that in combat are probabilistically dependent ...are [all too often] treated as probabilistically independent..."

Horrigan's examples relate, for example, to penetrating aircraft attempting to survive passage through a region of air defenses by a variety of concentration tactics; to the common assumption of assigning the same small-arms proficiency to all members of a heterogeneous group; and to efforts to represent ground combat through Lanchester equations. Upon thinking deeply about Horrigan's points, one begins to recognize how ubiquitous the configuration problem really is. Indeed, we speculate that it should be considered to be a subset of a more general class of problems having to do with allocation of resources, not only to positions, but, e.g., to missions.

3.3 Interactions Among Entities on a Network

Yet another aggregation-related difficulty is specifically troublesome in the context of distributed simulation on a network such as SIMNET (Gilmer, 1993). Suppose, for example, that one group of workers has described an army platoon as a collection of item-level objects such as tanks. Now suppose that in the context of a networked war game, that platoon is to encounter an enemy battalion developed, owned, and described by another group. And suppose that the battalion is characterized as a single object with various attributes such as a strength of weapons in each of various weapon classes. How do the platoon and battalion engage? Who builds the "methods" describing the attrition processes and what should the person doing so assume about the kinds of objects that will be encountered? The attrition processes could aggregate the platoon's objects into a single aggregate object or could disaggregate the battalion into item-level objects, but in either case a great many assumptions are necessary and one can readily see that the usual tendency will be to depend heavily on allegedly "representative" configurations—a dangerous business.

3.4 Need for Stochastic Aggregate Models

Another common problem discussed at the 1992 conference is that microscopic stochastic phenomena are often represented in the aggregate by deterministic so-called expected-value models. This can be pernicious. For example, a deterministic model may describe results of ground combat at a 3:1 force ratio as a break-even battle in which both sides have the same fractional loss rate. By contrast, a stochastic representation would show a bimodal distribution of results, with equal probabilities of attacker and defender victories, but with the outcome in any given case being clear-cut. Clearly, using the deterministic model in the context of training and exercises could be teaching precisely the wrong lessons.

Stochastic models, however, are no panacea. Indeed, many of them offer far less than advertised, treating some parameters as random variables, but holding other critical parameters as fixed (e.g., attrition rates and movement rates) even though they may be highly uncertain. Nonetheless, one of the clear lessons from the variable-resolution work done so far is that more future aggregate models should be appropriately stochastic. The Army's Concepts Analysis Agency (CAA) has begun moving along this path. Let us also mention that if one does use a deterministic low-resolution model, care must be taken in connecting it to stochastic higher resolution models (Allen, 1993).

3.5 Partial Differential Equation Approaches

Fowler (1993) describes variable-resolution issues in the context of partial differential equation models, drawing on well known equations in physics for insights that could be used in combat modeling. Two features of his paper are especially interesting: first, he explicitly considers conservation laws and other boundary conditions when examining resolution shifts (something seldom done in combat modeling); and, second, he demonstrates conceptually how combat modeling is arguably akin mathematically to "transport modeling," in gas laws, especially if one distinguishes among different phases of combat. Although his methods were presented more for insight and metaphor than for immediate practical application, they will prove illuminating to those with backgrounds to appreciate them.

4 EXPERIMENTS TO COMPARE AND UNDERSTAND AGGREGATION METHODS

Our efforts with Lanchester equations were an experiment to see whether inserting some theory could affect relatively basic resolution-related arguments in the military modeling community. The answer was yes, although we were looking only at a specific case. In a similar vein and at the suggestion of Donald Blumenthal, we conducted a cross-organization experiment in which we and colleague John Owen (on loan to RAND from Britain's Defence Operational Analysis Establishment) developed relatively detailed physics-level models of a simple combat problem along with some Lanchester-equation aggregations, and then compared results and methodology with those of a similar effort at Lawrence Livermore National Laboratory by Blumenthal and Ralph Toms (Hillestad, Owen, and Blumenthal, 1993).

The problem considered envisioned nothing more complex than a group of tanks and infantry fighting vehicles approaching a comparable but smaller defender group across a flat plain. The model simulated movement and firing of each system (i.e., each tank) over time. It assumed that the defender's single-shot kill probability was higher than that of the attacker's, but that both sides' kill probabilities increased with decreasing range so that the defender's advantage vanished at zero range. This was a very simple way to reflect the common situation in which the defender has an advantage due to a combination of having found and prepared good firing positions.

The problem was so simple that one might reasonably expect two teams of comparable skill and knowledge to get identical results. And, even if both sides entertained approximations about how to simplify the problem (e.g., by treating groups of weapons as a single highly effective weapon or by postulating and calibrating some Lanchester equations), one might expect results to be quite similar from group to group. To be sure, this was an experiment done almost in "spare time," without the benefit of checking and rechecking, peer review, and so on. At the same time the work was done about as carefully as a great deal of modeling. The results were sobering.

While the most basic simulations agreed well, major differences arose as soon as we compared more approximate results. Further, in comparing deterministic and stochastic outcomes, it was apparent that even in the aggregate, deterministic simulations are quite misleading in situations close to break-even circumstances (the very circumstances in which most combat modeling is done!). This was not really a "surprise" to anyone in the group, all of whom understood stochastic theories, but it certainly was a reminder of problems long since swept under the proverbial rug, by ourselves and nearly all workers involved in aggregated modeling. Also, while the inconsistencies were readily understandable in retrospect, they were due to precisely the kinds of errors that occur routinely in simulation modeling. In particular, even expert intuition is a remarkably poor basis for assuming aggregation relationships.

Even more sobering perhaps, a modicum of theoretical effort is sometimes not enough, because the aggregation issues are so subtle. As one example of this, the first attempt to construct a Lanchester-equation description of the problem failed miserably because no one had thought to consider the finite speed of the tank rounds. Even if they had, they probably would have convinced themselves that they were so large compared with all other speeds as to be effectively infinite. In fact, however, the finite speed of the shells meant that for a roughly three-second interval shooters lacked information on whether targets already shot at would be killed. As a result, they shot at some targets more than once, reducing their effectiveness substantially. The Lanchester model failed to account for this, until, upon comparing results, the need to do so was recognized. In this case, the comparison to more detailed simulations was crucial: the problem would not have been uncovered from a theoretical analysis unless it were exceptionally well done and rigorous.

The significance of these experiments was not in their particular conclusions or the methodology employed. Rather, it was in sensitizing us—and others to whom we briefed our experiences—to the importance of the modeling community having forums in which to conduct *detailed* comparisons of both theory and results. This is the analog in modeling to what scientists have long believed in passionately: the importance of being able to reproduce experiments. Progress comes not only from talent and hard work, but also from the iterations that come as the result of comparisons, arguments, and new insights. Other modeling groups have made comparable observations from time to time (see, e.g., Greenberger, et. al., 1976 on the value of "counter modeling"), but the lessons need to be constantly relearned because, unlike in the experimental sciences, there is no well established mechanism for military modelers to routinely "publish" their work in a form allowing and encouraging others to review it and argue with it.

5. MODELING ENVIRONMENTS

Object-oriented modeling and programming has a great deal to offer and is very helpful in dealing with variable-resolution modeling and cross-resolution model connection (Zeigler, 1991 and Rumbaugh, et. al, 1991, 1993). This goes beyond the benefits of modularity alone. It encourages and even "forces" hierarchical representation of objects and attributes where appropriate, it highlights important distinctions such as between aggregation and generalization, and can support some types of variable resolution in process through state generalization (sub states), state aggregation (concurrent submachines), and event generalization (subevents) (Eddy, 1993).

Unfortunately, object-oriented methods do not necessarily provide models and programs that are easy to understand. In particular, by taking an object-focused approach, they often leave fairly obscure how the various processes interact. Rumbaugh, et. al attempt to deal with this through strong design that includes multiple diagrammatic perspectives (object hierarchies, state-transition diagrams, and data flow diagrams). We and colleagues at RAND are pursuing object-oriented models (Hillestad, Moore, and Larson) and object-oriented environments that include a readable computer language (Anabel, conceived by colleague Ed Hall), a hypermedia approach to on-line documentation, and exploitation of commercial tools for input and output (e.g., the graphics of Microsoft EXCEL™ on the Macintosh computer) (see Anderson, Bankes, Davis, Hall, and Shapiro, 1993 for ambitions; a working version of the language and major aspects of the environment were emerging as of July, 1993). Some of these methods can be quite valuable also for verification and validation of models (Davis, 1992).

Another ambitious effort to develop appropriate modeling environments for good software engineering and variable resolution is described in Landauer and Bellman (1993), which describes *wrapper* approaches in some detail and surveys larger issues involved in defining a good environment.

6 CONCLUSIONS

Our conclusions, then, include a recommended design approach to variable-resolution modeling and cross-resolution model connection; suggestions on how to avoid a set of common and generic pitfalls; a call for theoretical work to precede plunging into modeling that postulates aggregation and disaggregation relationships; and a plea for routine in-depth cross-organizational comparisons and discussions of both modeling assumptions and results. We also suggest that managers not assume that connecting existing models will save time and money. The results may be incomprehensible and dangerous. Further, the advances in modeling technology have been so profound that redesigning and recoding, in object-oriented languages, will often be far less expensive than traditional estimates suggest. Finally, we urge sponsors to support extensive theoretical and experimental efforts, in many organizations, to develop sound aggregation and disaggregation relationships, design methods, and related tools for VRM work. The subject is seriously under invested.

REFERENCES

Allen, Patrick, "Combining Deterministic and Stochastic Elements in Variable Resolution Models," in Davis and Hillestad (1993), abbreviated henceforth as DH93.

Bankes, Steven, *Issues in Developing the Potential of Distributed Warfare Simulation*, RAND R-3141, 1992.

Brendley, Keith and Jed Marti, "A Distributed Network Approach to Variable Resolution Modeling," in DH93.

Cohen, Danny, David Mills, Richard Nance, Paul Reynolds, Mike Sovereign, Patrick Cheatham, and Robert Weber, *Panel Review of Aggregate Level Linkage Technology, Aerospace Corp., 1993* .

Courtois, P.J., "On Time and Space Decomposition of Complex Structures," *Communications of the ACM*, **28**, No. 6, pp. 590-603, 1985.

Davis, Paul K. and Donald Blumenthal, *The Base of Sand Problem: a White Paper on the State of Military Combat Modeling*, RAND N-3148, 1992.

Davis, Paul K. and John Arquilla, *Deterring or Coercing Opponents in Crisis: Lessons from the War With Saddam Hussein*, RAND, R-4111, 1991.

Davis, Paul K. and Reiner Huber, *Variable-Resolution Combat Modeling: Motivations, Issues, and Principles*, RAND N-3400 1992.

Davis, Paul K. and Richard Hillestad (ed.), *Proceedings of Conference on Variable Resolution Modeling*, RAND, CF-103, 1993. Abbreviated DH93.

Davis, Paul K. and Richard Hillestad, "Aggregation, Disaggregation, and the Challenge of Crossing Levels of Resolution When Designing and Connecting Models," Proceedings of the 4th Annual Conference on Artificial Intelligence, Simulation, and Planning in High Autonomy Systems, September 20-22, 1993. IEEE Computer Society Publications.

Davis, Paul K., *An Introduction to Variable Resolution Modeling and Cross-Resolution Model Connection*, RAND, R-4252, 1993.

Davis, Paul K., *Generalizing Concepts of Verification, Validation, and Accreditation in Combat Models*, RAND, R-4249, 1992.

Defense Modeling and Simulation Office (DMSO), *Defense Modeling and Simulation Initiative*, May 1992.

Defense Science Board, *Impact of Advanced Distributed Simulation on Readiness, Training, and Prototyping*, 1993.

Eddy, Frederick, "Object Oriented Modeling: Holistic and Variable-Resolution Views," in DH93.

Fowler, Bruce, "Resolution Changes and Renormalization for Partial Differential Equation Combat Models," in DH93.

Gilmer, John "An Examination of Two Modeling Choices and Their Implications for Distributed Simulation," unpublished presentation at the annual meeting of the Military Operations Research Society , June, 1993.

Greenberger, Martin, Matthew Crenson, and Brian Crissey, *Models in the Policy Process*, Russell Sage Foundation, New York, 1976.

Harshberger, Edward, Bart Bennett, and David Frelinger, "An Approach to Hierarchies of Models: Process Independence," in DH93.

Hillestad, Richard and Mario Juncosa, *Cutting Some Trees To See the Forest: On Aggregation and Disaggregation in Combat Models*, RAND R-4250, 1993.

Hillestad, Richard, John Owen, and Donald Blumenthal, *Experiments in Variable Resolution Combat Modeling*, RAND N-3632, 1993.

Hillestad, Richard, Louis Moore, and Eric Larson, *The Theater Level Campaign Model/Nonlinear Combat Model Tools Kit: An Overview*, RAND, forthcoming.

Horrigan, Timothy, "The Configuration Problem" and Challenges for Aggregation," in DH93.

Huber, Reiner, "Crossing Levels of Resolution in Defense Analysis: a Requirement of the New Strategic Environment," in DH93.

Innis, G. and E. Rexstad, "Simulation Model Simplification Techniques," *Simulation*, July 1983, pp. 7-15.

Landauer, Christopher and Kirstie Bellman, "Integrated Simulation Environments," in DH93.

Palmore, Julian, "Variable Resolution Modeling in Mathematics," in DH93.

Perry, Walter L., John Schrader, and Barry Wilson, "A Variable Resolution Approach to Modeling Command and Control in Disaster Relief Operations," in DH93.

Rumbaugh, James, Michael Blaha, William Premeriani, Frederick Eddy, and William Lorensen, *Object-Oriented Modeling and Design*, Prentice-Hall, 1991.

Scheckeler, Günther, "A Hierarchy of Models at Different Resolution Levels," in DH93.

Sevinc, Suleyman, "Automation of Simplification in Discrete Event Modeling and Simulation," in *Int. Journal of General Systems*, Vol. 18, pp. 125-145, 1990.

Simon, Herbert, *Sciences of the Artificial*, MIT Press, Cambridge, Mass, 1981.

Weatherly, Richard, David Seidel, and Jon Weissman, "Aggregate Level Simulation Protocol," *Proceedings of the Summer Computer Simulation Conference.*, 1991.

Zeigler, Bernard, "A Systems Methodology for Structuring Families of Models at Multiple Levels of Resolution," in DH93.

Zeigler, Bernard, "Object Oriented Modeling and Discrete-Event Simulation," Advances *in Computers, Vol. 33, 1991.* Reprinted in DH93.

AUTHOR BIOGRAPHIES

PAUL DAVIS is RAND's Corporate Research Manager for Defense Planning and Analysis. He holds a B.S. from the University of Michigan and a Ph.D. from MIT in Chemical Physics. Before joining RAND he was a senior executive in the Office of the Secretary of Defense. His research involves defense planning, modeling and simulation, and decision making theory. He is on the faculty of the RAND Graduate School.

RICHARD HILLESTAD is RAND's Resident Scholar for Systems Analysis. He holds a B.S. in electrical engineering from the University of Wisconsin and a Ph.D. from UCLA in operations research. His current research involves object-oriented modeling and tools, force structure modernization, airport safety, and transportation. He is on the faculty of the RAND Graduate School.