

# Exploring Combat Models for Non-monotonicities and Remedies

Major William Vinyard, United States Marine Corps, wvinyar@bellsouth.net

Associate Professor Tom Lucas, Naval Postgraduate School, twlucas@nps.navy.mil



Prof Tom Lucas



Maj William Vinyard

## Introduction

Throughout the Department of Defense (DoD), analysts use Models and Simulations (M&S) to assist decision makers in choosing among various strategic courses of action, whether or not to purchase particular weapons systems, implement programs, and so on. The cost of these weapon systems and programs can run into billions of dollars, and there is always the possibility of lives lost. Therefore, ensuring that our models are logically coherent and provide believable results is of paramount importance. Models that contain non-monotonic output may be unbelievable to decision-makers. A model is non-monotonic if adding capability to one side, while holding everything else constant, results in a less favorable outcome for that side. When a model used to decide between alternatives exhibits non-monotonic behavior, the model's validity, as well as its usefulness for analysis, may come into question.

Non-monotonic behavior has been observed in large, well-established deterministic combat models. For example, **Saeger** and **Hinch**, in a force thinning study that was part of the Quadrennial Defense Review-1997, found that the model they were using suggested that more capability sometimes resulted in worse outcomes.<sup>1</sup> This article summarizes some of the findings in Vinyard (2001) of high-dimensional explorations of the relatively simple Dewar model.<sup>2</sup> In particular, two questions are addressed: (1) How widespread is non-monotonicity in the model? and (2) Can the response surface in non-monotonic regions be made more amenable to interpretation by decision-

makers without destroying the chaos that may be inherent to both real combat and the model?

## Background

In 1991, Dewar et al., released a study showing that even simple, deterministic combat models can be subject to non-monotonic behavior.<sup>3</sup> This finding was not new. However, the finding that the non-monotonicity was due to chaos inherent in the model was new. Perhaps this shouldn't have been surprising, as combat models are nonlinear systems with delayed feedback processes that can reinforce or dampen the system. That is, combat models are exactly the type of model likely to exhibit chaotic behavior. Of course, this can be troubling for modelers because a characteristic of chaotic systems is extreme sensitivity to initial conditions. Furthermore, it is an inescapable fact that no matter how careful our measurements, the data used in our analyses are subject to errors. In chaotic systems, even if the magnitude of these errors is extremely small, the uncertainty associated with the errors creates uncertainty about our knowledge of the system in the future. In fact, Sandmeyer, found that a computer's double precision round-off algorithm qualitatively affected a large, deterministic combat model's results.<sup>4</sup>

Dewar et al. cautioned that the chaotic behavior manifested in a combat model may or may not be an accurate reflection of chaos on the battlefield.<sup>3</sup> However, one of the essential aspects of combat may be chaos—for want of a nail ... the battle was lost. A combat model that exhibits chaotic behavior in an appropriate way seems, on an intuitive level, to be more realistic than a model that does not.

## Previous Research and Suggested Remedies

The original article by Dewar et al. spawned a series of articles. Many of these papers explored ways to explain, overcome, or attenuate the effects of chaos and the resulting non-monotonicities. Palmore offered multiple causes of instabilities and discussed several ways to address these

problems.<sup>5</sup> Louer advocated parametric variation of variables and stochastic decision thresholds.<sup>6</sup> He claimed that these methods cope with the underlying chaos, while providing results that are informative and trends that are monotonic in behavior. Louer also suggested the use of

experimental design . . . covering the full range of uncertainties . . . to develop distributions of the response functions. . . The skilled analyst then needs to . . . assess if they have any significant influence on the trends of the response function curves.<sup>7</sup>

Cooper urged "far-ranging sensitivity trials, to explore more of a model's domains."<sup>8</sup> Cooper also showed how a minor change to the Dewar model's decision rules results in purely monotonic behavior in the two-dimensional subspace he examined.<sup>9</sup> Huber and Tolk showed that "non-monotonic effects may largely be eliminated if dynamic *mission-oriented* decision rules are used rather than the static *state-oriented* decision thresholds [Dewar used]."<sup>10</sup> It has also been shown that careful stochastic manipulation of both decision thresholds and attrition coefficients can significantly smooth the non-monotonicities that arise due to dynamic instabilities inherent in combat models.<sup>11</sup>

Saeger and Hinch recently discovered non-monotonic behavior in a large, deterministic model.<sup>1</sup> In an attempt to get meaningful results out of the model, they defined a neighborhood of the phase space, randomly perturbed selected variables, and attempted to fit a probability curve to the output values of multiple runs. They concluded:

[T]he probability distribution of the [response curve] is a function of the variable that is perturbed. [But there is] no obvious way to determine, *a priori*, which variables to perturb [nor] by what magnitude to perturb them.

Their recommendation was to "perturb all variables [and] perform sensitivity analysis in perturbation magnitude."

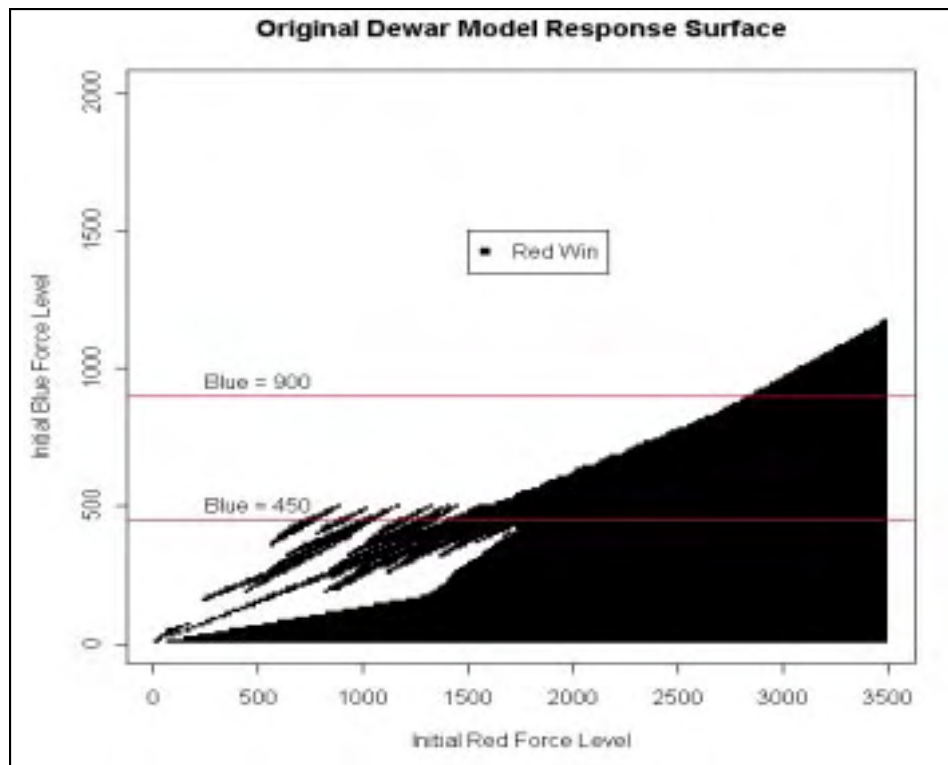
(See **COMBAT MODELS**, p. 36)

**Searching for Non-Monotonicity in the Dewar Model**

The previously published papers on the Dewar model examined only one or two of the model’s 18 dimensions. Thus, the question must be asked: Are these findings anomalies that occur only in the small portions of the model that have been examined, or do they generalize to other measures and dimensions? The combinatorial possibilities of main effects and interactions among the 18 dimensions are too great to examine en masse. Consequently, we use advanced statistical designs and billions of computational experiments to more fully explore the model.

The original Dewar model is a deterministic time-step simulation of a homogeneous Lanchester square law battle (see [3] for details). In addition to the attrition rate coefficients, the model has parameters for initial force sizes, reserves, reinforcement levels, reinforcement delays, and decision thresholds. At each time step, depending on the engaged force ratio and force levels, each side makes decisions on whether to withdraw or call in reinforcements. There is a natural symmetry in the parameters—i.e., for each Blue parameter there is a corresponding Red parameter. The previous research focused on the initial forces subspace (i.e., the two-dimensional space defined by initial Blue force level and initial Red force level) and the binary outcome measure ‘who wins.’ Figure 1 shows the non-monotonic output of this subspace in the Dewar model. In this two-dimensional graph, initial Red force levels vary from ten to 3500, in increments of ten. Initial Blue force levels vary from ten to 2000, also in increments of ten. Thus, the model was run 69,451 times to generate this surface. The black region represents those initial force levels that result in a Red win. Consider the following two scenarios:

- a. Initial Blue forces are fixed at 900 (this is indicated by a line on the graph), and initial Red forces vary from ten through 3500. As additional troops are added to the initial Red force, with all other variables held constant, the trend is monotonic. At an initial force level of about 2800 troops, Red starts to win. After that, adding troops to Red’s initial force level results in continued Red wins, just as one



**Figure 1. The Response Surface of the Original Dewar Model.**

would expect.

- b. The initial Blue force level is fixed at 450, while Red varies from 700 to 1800. Now, the response trend goes from Blue wins to Red wins, and back and forth many times. This non-monotonic trend seems to make it impossible for a decision-maker to decide whether or not adding more Red Forces is a good idea.

Note the broad region of extreme non-monotonicity in Figure 1. Also note that large portions of this subspace contain nice monotonic regions. If this graph is any indication of the subspaces that exist in larger models, then it is easy to see why extreme non-monotonicity might go unnoticed, even when it exists. In larger models, the dimensionality of the phase space is incomprehensibly vast. It is entirely possible that these large models are operating in purely monotonic regions. However, it is also possible that they are teetering on the edges of non-monotonic regions like the one pictured here.

In the Dewar model, there are  $\binom{18}{2} = 153$  pairs of variables. We chose to search the subspaces associated with the nine natural pairs of variables, with a natural pairing consisting of the same parameter for Red and Blue. For each of these pairs, we want

to see if the surface is monotonic over a range of settings for the other 16 parameters. To do this, we used random Latin Hypercube Sampling on the 16 parameters. From McKay et al., Latin Hypercube Sampling “can be viewed as a K-dimensional extension of Latin Square Sampling” and generates an efficient “space-filling” design.<sup>12</sup>

For each of the nine natural two-dimensional subspaces, 16 surfaces are generated and assessed for non-monotonicity. In total,  $9 \cdot 16 \cdot 69,451 = 10,000,944$  battles are simulated to generate 144 surfaces. In designing our sample, we adhere to the original Dewar model’s basic structure of a smaller, more efficient force opposing a larger, less effective force; or, if you prefer, a smaller Blue defensive force opposing a larger Red attacking force. To preserve the original model’s tension between opposing forces, we restrict the domain of the remaining variables to fairly thin hyperplanes, centered at the nominal values of the original model.

This exploration found that non-monotonicity, with respect to the measure ‘who wins,’ is prevalent in the model, with non-monotonicity for ‘who wins’ being found in seven of the nine two-dimensional subspaces explored (see Table 1). In fact, five of the subspaces exhibit pervasive non-

Two-dimensional Dewar Model Subspaces	Percent of Response Surfaces Containing Non-monotonic Behavior
Initial Force Levels	100%
Force Ratio Reinforcement Thresholds	81%
Percent of Remaining Forces Reinforcement Threshold	19%
Force Ratio Withdrawal Threshold	0%
Percent of Remaining Forces Withdrawal Threshold	0%
Reinforcement Blocks Available	6%
Reinforcement Block Delay	100%
Reinforcement Block Size	88%
Attrition Coefficients	94%

### Mitigating Non-Monotonicity in the Dewar Model

The fact that so many subspaces contain non-monotonic behavior is cause for concern. The Dewar model contains some of the same basic processes that many of the larger models use, such as decision thresholds and attrition processes. If the interaction of these processes in the Dewar model generates such widespread non-monotonic behavior, then the larger, more complex models may also be affected by similar non-monotonicities. What can be done to generate interpretable responses despite the non-monotonicities caused by the chaos in the battle trace? Much of the literature reviewed for this paper indicates that stochastic modeling can be a useful way to deal with non-monotonic behavior in both the simple Dewar model and other, more complex models.

To examine how making parameters stochastic affects non-monotonicity in the Dewar model, we ran a fractional factorial experiment to determine the effect of stochastic modeling on the trends of the response surface. The experiment varied nine factors consisting of the nine types of parameters in the model. Each factor had two levels, deterministic and stochastic. To efficiently search all nine factors simultaneously, a  $2^{9-3}$ , resolution V, fractional factorial design is used.<sup>13</sup> This requires 64 different input settings. Each surface consists of 69,451 points. Plus, each point in a stochastic model must be estimated. To get precise estimates, we use 1000 replications to estimate the probability that Red wins at each point. Thus, about 4.5 billion battles were simulated in the fractional factorial experiment.

It is not always easy to distinguish between bona fide non-monotonicities and random variation over an entire surface. Six parameters—measuring the number of non-monotonic jumps, trends, statistically significant jumps, statistically significant trends, and functions thereof—were used to assess the extent of the non-monotonicities. See [2] for how this was done and an empirical assessment. Following previous analyses, the attrition coefficients are modeled as normal random variables, with means equal to the nominal Dewar model values and standard deviations equal to ten percent of the square root of their means. All other parameters, when stochastic, are modeled as

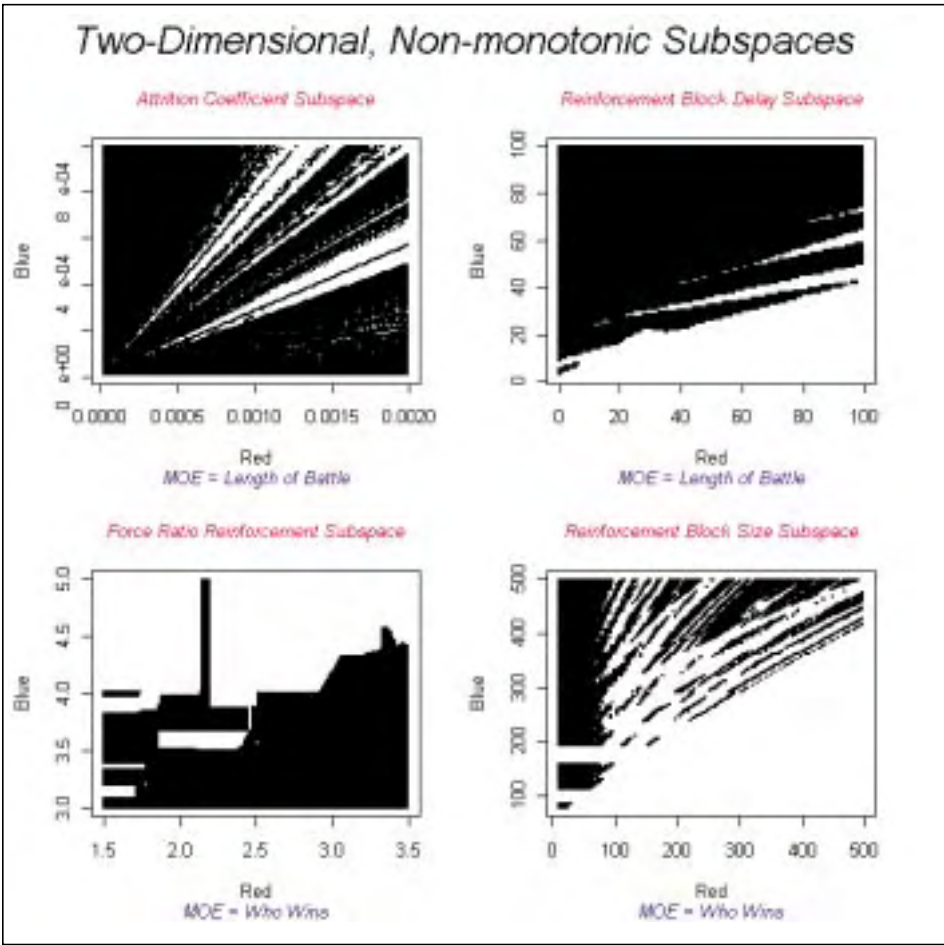
(See COMBAT MODELS, p. 38)

**Table 1. Latin Hypercube Sampling Results: 78 of the 144 response surfaces found during the Latin Hypercube Sampling exhibited non-monotonic behavior.**

monotonicity, with it showing up in over 80 percent of the surfaces checked. In total, 54% of the surfaces generated contain non-monotonic regions.

Figure 2 shows examples of the newly discovered widespread non-monotonicity in the Dewar model. The top row of the figure displays two of the many striking

examples we discovered using Latin Hypercube Sampling. These two graphs exhibit non-monotonicity with respect to the MOE ‘length of battle.’ The two bottom graphs in Figure 2 exhibit non-monotonicity with respect to the MOE ‘who wins’ in the force ratio reinforcement and reinforcement block size subspaces.



**Figure 2. Four of the Many Two-dimensional Subspaces Exhibiting Non-monotonic Behavior**

## COMBAT MODELS

(continued from p. 37)

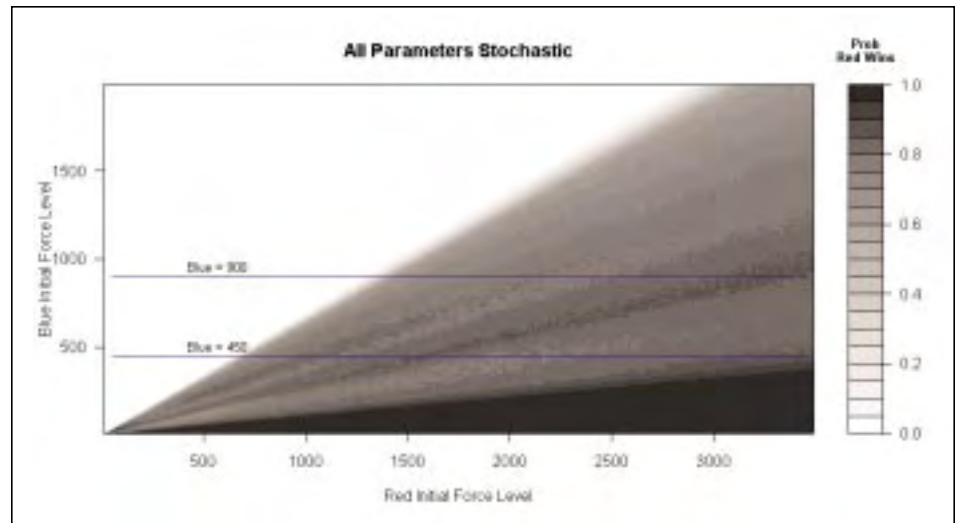
uniform random variables with intervals, centered on the nominal Dewar model values, ranging from plus to minus five percent of the nominal value. These small random deviations are within the uncertainties that would be present on a real battlefield.

The results are dramatic and convincing, and support previous work.<sup>1,6,7,11</sup> Stochastic perturbation usually dramatically reduces the non-monotonic behavior of the response surface, but can, by some measures, exacerbate it. The attrition coefficients are the model parameters, over the values we investigated, that have the greatest effect on the reduction of the non-monotonic behavior—see [2] for a detailed analysis of the effects and interactions between parameters. Figure 3 shows the same surface as Figure 1, with all of the parameters stochastic. Unlike the previous graph, Figure 3 appeals to our intuition (the outcome remains uncertain until one side or the other quits the field of battle).

## Conclusions

*Caveat actor et cavendo tutus*

Feigenbaum, Mandelbrot and other pioneers of chaos theory have shown that chaos is the rule rather than the exception in the real world.<sup>14</sup> The possibility that chaos and its consequent non-monotonicity are also the rule in large, complex combat models is very real. The Dewar model is relatively small compared to most of the larger, more complex models that DoD currently uses. Nonetheless, it includes many of the same processes present in these much larger models. Thus, it is reasonable to suppose that what we learn from studying the Dewar model will help us with our larger models. The bottom line is that non-monotonicity may be more pervasive in combat models than previously suspected. Stochastic modeling can be a viable method for dealing with non-monotonic response surfaces. However, stochastic modeling must be done carefully. When it is, the non-monotonic behavior of the model can be dramatically reduced while maintaining the chaos that may be inherent to combat, thereby making the trends of the response surface more useful for comparative analyses. Better analyses can help decision-makers save time, money, and other assets (and, perhaps, another soldier, sailor, airman or Marine will get to enjoy their pension).



**Figure 3. When all parameters are stochastic, the response surface represents the probability of a Red win. In this graph, the regions where Red or Blue wins are clearly delineated. The gray region represents those initial force levels where the outcome is uncertain.**

## References

- [1] Saeger and Hinch, "Understanding Instability in a Complex Deterministic Combat Simulation," not yet published paper, 24 August 2000.
- [2] Vinyard, "Reducing Non-monotonicity in Combat Models," Master's Thesis, Naval Postgraduate School, September 2001.
- [3] Dewar, Gillogly and Juncosa, "Non-monotonicity, Chaos, and Combat Models," RAND R-3995-RC, 1991.
- [4] Sandmeyer, "Simtech Project: Application of Supercomputers to Division/Corps Level Combat Simulation," AMSAA, Interim Note No. C-159, October 1990.
- [5] Palmore, "Research on the Causes of Dynamical Instability in Combat Models," USSACERL Technical Report 96/95, August 1996.
- [6] Louer, "More on Non-Linear Effects...", Phalanx, March 1993.
- [7] Louer, "Treatment of Non-monotonic Effects in Models for Analyses," PHALANX, March 1994.
- [8] Cooper, "Non-Monotonicity and Other Combat Modeling Ailments," PHALANX, June 1994.
- [9] Cooper, "Non-monotonicity Revisited," Phalanx, September 1994.
- [10] Huber and Tolk, "Non-monotonicity Effects in Combat Models and Tactical Decision Modeling," Phalanx, September 1994.
- [11] Lucas, "How One Randomizes Matters: A Study of Non-monotonicity and Randomness in Combat Analysis," Phalanx, March 1997.
- [12] McKay, Beckman and Conover, "A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code," Technometrics, vol. 21, no. 2, May 1979.
- [13] Box, Hunter and Hunter, Statistics For Experimenters, John Wiley and Sons, 1978.
- [14] Gleick, Chaos: Making a New Science, Penguin Books, 1987.

## Biographies

*Major William Vinyard has been on active duty since 1980, serving in a variety of billets throughout the Marine Corps. He recently graduated from the Naval Postgraduate School with an MS in Operations Research. He is currently stationed at the Marine Corps Logistics Base, Albany, Georgia, in the Logistics Operations Center working as an operations analyst.*

*Dr Tom Lucas is an Associate Professor of Operations Research at the Naval Postgraduate School. His primary research interests are combat analysis, design of experiments, and robust Bayesian statistics. Previously, he worked as a statistician at RAND and as a systems engineer at Hughes Aircraft Company. ☺*