REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*
01-11-2001

2. REPORT TYPE
Technical Document

3. DATES COVERED (From – To)
October 1996 – November 2001

4. TITLE AND SUBTITLE
Multiple Runs of a Deterministic Combat Simulation

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

6. AUTHOR(S)
Mr. Timothy J. Bailey

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
TRADOC Analysis Center
ATTN: ATRC-F
255 Sedgwick Avenue, Building 314
Fort Leavenworth, KS 66027-2345

8. PERFORMING ORGANIZATOIN REPORT NUMBER
Technical Document TRAC-F-TD-02-001

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

20020904 056

13. SUPPLEMENTARY NOTES

14. ABSTRACT
In the past, combat simulations of corps and division level operations required several days to model a few days of combat. Driven by the speed, size, and cost of the computers of that time, it was infeasible, if not impossible, to replicate the runs in a run design. Thus, the investigation of that scale of combat used deterministic models. Today, computer technology offers performance improvements of two to three orders of magnitude, all packaged to fit on a corner of a desk at less than a tenth of the cost of the computers ten years ago. We now have the ability to consider making significantly more runs of large combat simulations.

One reason to consider making multiple runs of a deterministic model may be the concept offered by Carl von Clausewitz, who identified chance events as explicit sources of "general friction," which can potentially turn success into failure and vice versa. This paper presents a deterministic method that produces multiple runs for analysis of the non-monotonic results of Clausewitzian war by varying the initial states of information systems within the model. The results of recent studies serve as case studies to explore the utility of this approach.

15. SUBJECT TERMS
Non-monotonic, Clausewitz, Information, Vector-In-Commander, Combat simulation, Deterministic, Multiple runs, Confidence Intervals, Sample size

16. SECURITY CLASSIFCATION:

b. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL	9

19a. NAME OF RESPONSIBLE PERSON
Mr. Timothy J. Bailey

19b. TELEPHONE NUMBER (include area code)
(913) 684-9217
DSN 552-9217

Multiple Runs of a Deterministic Combat Model

Timothy J. Bailey
255 Sedgwick Avenue
Fort Leavenworth, Kansas 66027
Phone: (913) 684-9217
FAX: (913) 684-9151
E-mail: baileyt@trac.army.mil

69th MORS Symposium
Working Group 12
12-14 June 2001
(1 November 2001)

ABSTRACT

In the past, combat simulations of corps and division level operations required several days to model a few days of combat. Driven by the speed, size, and cost of the computers of that time, it was infeasible, if not impossible, to replicate the runs in a run design. Thus, the investigation of that scale of combat used deterministic models. Today, computer technology offers performance improvements of two to three orders of magnitude, all packaged to fit on a corner of a desk at less than a tenth of the cost of the computers of ten years ago. We now have the ability to consider making significantly more runs of large combat simulations.

One reason to consider making multiple runs of a deterministic model may be the concept offered by Carl von Clausewitz, who identified chance events as explicit sources of *general friction*, which can potentially turn success into failure and vice versa. This paper presents a deterministic method that produces multiple runs for analysis of the non-monotonic results of Clausewitzian war by varying the initial states of information systems within the model. The results of recent studies serve as case studies to explore the utility of this approach.

ISSUE

Why are multiple deterministic model runs useful for some analyses? Most practitioners of military operations research with deterministic model experience will answer that *at times, deterministic models display non-monotonic tendencies.*

The definition of non-monotonic, used by Dewar et al (1996), is "adding more capabilities to one side only does not lead to as least as favorable combat outcome for that side." Because of this characteristic, some analysts question the utility of deterministic models to provide meaningful results, especially whenever their results are from a non-monotonic region of a solution space. The U.S. Army Training and Doctrine Command (TRADOC) Analysis Center (TRAC), located at Fort Leavenworth, Kansas, addresses this potential non-monotonic issue through an analytical method that allows multiple runs of the deterministic Vector-In-Commander (VIC) model.

BACKGROUND

In the 1980's and 1990's, TRAC used large (at the time), expensive, multi-user computers to run large (meaning greater than 250,000 lines of code) combat simulations to conduct detailed corps and division level analysis. It often took several days of execution time to simulate several days of combat. Accordingly, since it was infeasible, if not impossible, to replicate the runs of a run design, detailed corps and division combat models tended to gravitate to deterministic simulations.

The speed, size, and cost of the computers of that time drove this condition. An example of such a computer used in the early 1980's by TRAC to run VIC is the VAX 780-5; rated at 1.3 million instructions per second, it had 4 megabytes of memory and 352 megabytes of disk space. The VAX supported 15 to 20 people, cost about $1 million, and required 1 to 2 days to make a run of VIC. Today, personal computers (PCs) rated at 1 gigahertz with 500 megabytes of memory and 27 gigabytes of disk space run VIC. VIC run times are now less than 60 minutes. Each PC supports one user and cost $5 thousand. Technology now provides the capability to consider making multiple runs of large combat simulations.

Another thing to be touched on, but is not really a topic for this paper, is the eternal debate about stochastic and deterministic models. Although such a debate warrants revisiting from time to time, this paper does not discuss deterministic and stochastic pros and cons nor recommend whether combat simulations should be one or the other. This is because deterministic combat models exist now; they are in use today; and, since High Level Architecture versions are under development, they will be used in the future. Thus, a method, such as described in this paper, is needed to use deterministic models for current and future combat analysis.

CONTEXT

The Prussian theorist, Carl von Clausewitz offers one possible reason for non-monotonic behavior of combat outcomes in his concept that identifies chance events as an explicit source of *general friction*. According to Watts (1996), Clausewitz's *general friction* can propel small differences from what is expected or predicted and can potentially turn success into failure and vice versa. Clausewitz wrote of war that no other human activity "is so continuously or universally bound up with chance." Clausewitz attributes *general friction* to three main factors: the participation of human beings in war, the distribution of information in war, and the unpredictability (chance) of non-linear processes in war.

Thus, good combat models should display behavior consistent with Clausewitz's *general friction*, because they contain numerous non-linear expressions, and, in addition, they introduce non-linearity (chance) through feedback by modeling the flow of information across the battlespace and by making combat decisions on their perception of the state of the battle. Dewar et al (1996) found, when studying deterministic models, that "decisions based on the state of the battle can be seen to cause widespread non-monotonicities in the outcomes of the models."

In addition, since current emphasis on the examination of information, information systems, and information processes focuses analysis on one of the major contributors to *general friction*, chance should play an even more prominent role in combat outcomes of these information studies.

THEORY

In their work with the simple deterministic model detailed in Figure 1, Dewar et al (1996) established that even simple, attrition-based, deterministic models produce non-monotonic results. Figure 2 is a pictorial representation of these non-monotonic results. The axes represent starting Red troop strength (x-axis) and starting Blue troop strength (y-axis). White points of the graph represent Blue wins and black points represent Red wins.

One can observe non-monotonicity among the outcomes by selecting a constant Blue starting troop strength, which is represented by the red arrow. Red strength increases to the right along the arrow. Instead of a clearly defined Red strength above which Red wins every time, non-monotonic results in terms of reversal of outcomes are evident in several places along the arrow. These non-monotonic outcomes are the result of decision points within the simple deterministic model for the introduction of Blue and Red reinforcements. Each force introduces the reinforcements, whenever their respective force becomes reduced to a set troop strength level or the force ratios became unfavorable, as described in Figure 1.

Figure 2 clearly illustrates the concern about deterministic models, when their results fall in a non-monotonic portion of a solution space. However, instead of a single point estimate of the outcome as typically done with deterministic models, making

	Blue	**Red**
Initial Troop Strength	Variable	Variable
Combat Attrition Calculations	$B_{n+1} = B_n - \frac{R_n}{2048}$	$R_{n+1} = R_n - \frac{B_n}{512}$
Reinforcement Thresholds	$\frac{R_n}{B_n} \geq 4$ or $B_n < .8\,B_0$	$\frac{R_n}{B_n} \leq 2.5$ or $R_n < .8\,R_0$
Reinforcement Block Size	300	300
Maximum Allowable Reinforcement Blocks	5	5
Reinforcement Delay (time steps)	70	70
Withdrawal Thresholds	$\frac{R_n}{B_n} \geq 10$ or $B_n < .7\,B_0$	$\frac{R_n}{B_n} \leq 1.5$ or $R_n < .7\,R_0$

Figure 1. Simple deterministic model.

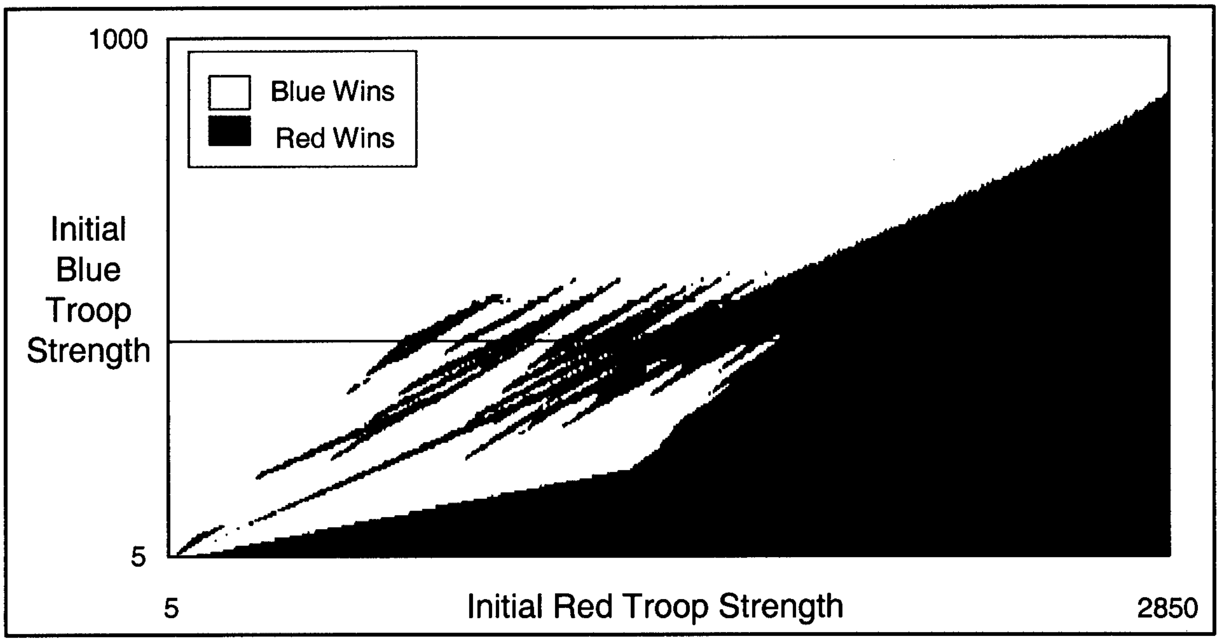

Figure 2. Battle outcomes.

multiple runs can provide a more thorough exploration of the solution space and will allow analysis of the non-monotonic results.

ANALYTICAL METHOD

Multiple runs provide a better estimate of the true population mean than a single run. This is especially true, if the runs fall within a non-monotonic portion of a solution space, because multiple runs allow for a more robust exploration of the solution space to account for chance variability that leads to the non-monotonic results. In addition, multiple runs provide a procedure to measure the variability of the results.

The set of criteria developed and employed by TRAC to perturb the model is presented in Figure 3. This perturbation scheme anchors the analytical method that enables multiple runs of the deterministic VIC combat model. The key criterion is affecting the many battlefield operating systems throughout the run, not just at the beginning.

Starting aerial sensors at random points along their orbits serves as the basis for the established perturbation scheme. Heretofore, orbiting aerial sensors typically started at one end of their orbit. This is a rather arbitrary synchronization of aerial sensors. It would be far more likely that they would not all arrive at the initial starting points of their orbits at the exact

- Must not directly affect the output being measured
- Must be mathematically sound
- Must not alter any performance data
- Must affect many battlefield operating systems
- Must continuously perturb the run – not just the initial conditions
- Should not require extensive code changes, and should be transparent to users who desire one run

Figure 3. Perturbation criteria.

instant of game initiation. Thus, having them at random points along their orbits at game initiation seems much more logical and realistic.

This perturbation scheme provides each run with a different set of detected units and, accordingly, different targets to engaged, and these differences continue throughout the entire run. Thus, since different targets are engaged and suffer attrition, decisions are made sooner in some runs, and later in other runs. Occasionally, entirely different decisions are made. Tests have demonstrated that this perturbation scheme provides the desired results. However, other perturbation schemes are possible, perhaps even desirable, depending on the specific study run design.

STATISTICAL BASIS

If the solution space (population) is the set of all possible solutions (in this case, combat outcomes), then each run is a sample of that population. Since the starting points of the aerial sensors are *random* positions on their orbits, multiple runs, even from a deterministic model, provide statistical samples from the population of the solution space.

The Central Limit Theorem (Devore 1982) states that, for a sufficiently large number of samples from a population, the samples will be approximately normally distributed (about the mean), their average mean will approximate the population mean, and, as the number of samples increases, the better the approximation. Therefore, the Central Limit Theorem tells us that an average of multiple runs typically is more representative of the average outcome of the entire population than a single run.

In addition, the Central Limit Theorem does not require that the population be normally distributed. If it is not, a conservative rule of thumb is that the sample size should be greater than or equal to 30. However, the closer the population is to a normal distribution, the fewer samples are required to approximate the mean.

The actual number of samples required can be determined using confidence intervals (Walpole and Myers 1985). A confidence interval to estimate the mean of an approximately bell-shaped population, when the true population variance is unknown and it is impossible to obtain a large sample size (greater than or equal to 30) is expressed as:

$$\bar{x} - t_{\alpha/2}\frac{s}{\sqrt{n}} < \mu < \bar{x} + t_{\alpha/2}\frac{s}{\sqrt{n}}$$

Where μ is the population mean, $\bar{x}$ is the sample mean, s is the sample standard deviation, $1\text{-}\alpha$ is the degree of confidence, n is the sample size, and t is the statistic for the Student's t-distribution.

If $\bar{x}$ is used as an estimate for μ, we can be $(1\text{-}\alpha)100\%$ confident that the error will not exceed a specified amount e, when the sample size is:

$$n = \left(\frac{t_{\alpha/2}\, s}{e}\right)^2$$

The method to determine the number of samples required is accomplished through an iterative application of this expression as presented in Figure 4.

- Chose an initial sample size, collect the samples, and compute the sample standard deviation.
- Calculate the required sample size using the expression above.
- If the required sample size is greater than the number of samples used, then additional samples have to be made until the required sample size is reached, then the method is repeated at the previous step.
- If, however, the required sample size is less than or equal to the sample size used to compute the sample statistics, then the sample size is sufficient.

Figure 4. Determining required sample size.

EXAMPLES

TRAC conducts a variety of studies, including attrition studies, in which the parameters being studied (e.g., weapon systems, munitions, force structures) can sometimes drive the results more than the chance events. Other TRAC studies(e.g., information studies) investigate parameters (e.g., Joint mission pass times, Joint mission priorities, interoperability) that may not drive the results more than the chance events.

Two studies recently conducted by TRAC will serve as case studies to examine the utility of using multiple runs. The *1st Infantry Division Limited Conversion to Division XXI Study* will be the case study for an attrition study. The *US/UK Sensor-To-Shooter Multinational C4 Interoperability Study* will be the case study for an *information study.* Both of these studies used the deterministic VIC combat simulation, as the evaluation tool to measure combat effectiveness.

In each of these studies, Red and Blue losses were one of the main measures used to examine combat effectiveness. It was interesting to note that, in both these attrition and information studies, Blue losses displayed much more statistical variance than did Red losses. The reason for this was that Red losses were used as one of the termination criteria for the battle. Accordingly, both of these examples focus on Blue losses. Nonetheless, the method described in Figure 4 should be applied to each measure of merit used in the analysis to determine the proper required sample size.

Figure 5 is a graph of the Blue losses for the attrition study, which consisted of ten runs each for a base case and three alternatives. Each alternative added more capability to the one before it, so one might expect that the alternative should do no worse than the previous alternative, and perhaps should do better. Thus, using Blue losses as the measure of merit, one would anticipate that the Blue losses decrease as we progressed through the alternatives. If they did not decrease, they should not do

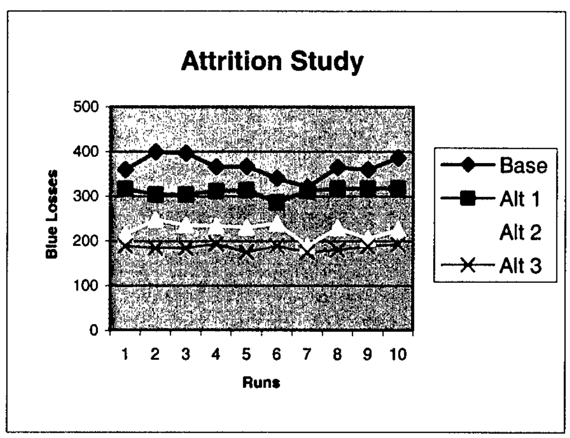

Figure 5. Attrition study Blue losses.

any worse, because we added more capability.

Inspection of Figure 5 shows that the Blue losses of the alternatives did indeed fall along expectations. Blue losses were the greatest in the base case; there were fewer Blue lost in alternative 1 than in the base case; there were fewer Blue lost in alternative 2 than in alternative 1; the fewest Blue losses were in alternative 3. Thus, the ordering of the alternatives is consistent, but one can still see Clausewitz's *general friction* at work, because there is some question about the magnitude of the results.

Figure 6 is a graph of the sample mean of the ten runs for each alternative. It displays the trend that is intuitive in the line plots of Figure 5. It is quite likely (except perhaps for run 7) that analysis from one run would yield findings consistent with multiple runs. Thus, the Blue results of this attrition study are reasonably well behaved and monotonic.

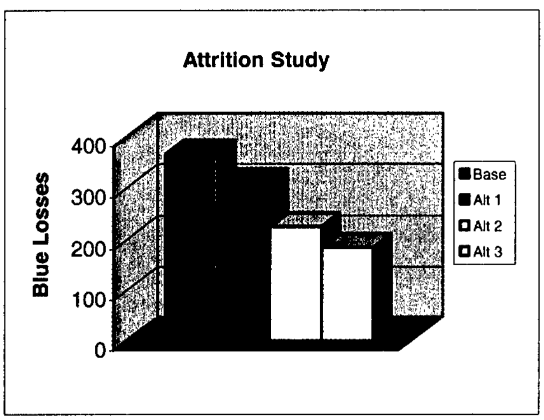

Figure 6. Mean attrition study Blue losses.

Figure 7 is a graph of the Blue losses for the information study, which consisted of ten runs each for a base case and two alternatives. As before, each alternative added more capability to the one before it, so one might again expect that the alternative should do no worse than the previous alternative. Again, using Blue losses as the measure of merit, one would anticipate that the Blue losses decrease as we progressed through the alternatives. If they do not decrease, they should not do any worse, because we added more capability.

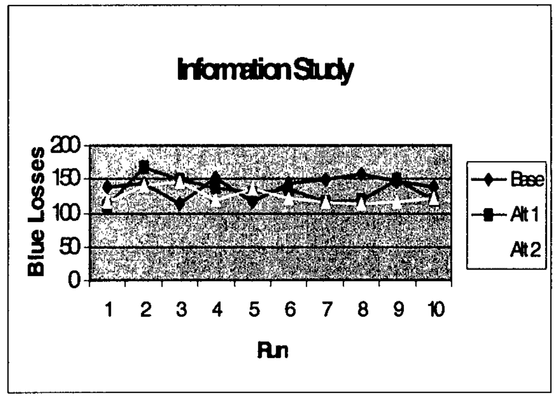

Figure 7. Information study Blue losses.

Inspection of Figure 7, however, shows that these expectations are not evident. Indeed, only three runs (4, 6, and 8) fall in the expected order, where Blue lost the most in the base case, less in alternative 1, and the least in alternative 2. One run (5) falls in inverse order meaning that Blue lost the least in the base case, more in alternative 1, and the most in alternative 2. Neither the order nor the magnitude of the results for the three alternatives is apparent from the chart. One can definitely see Clausewitz's *general friction* at work.

Figure 8 is a graph of the sample mean of the ten runs for each alternative. It displays the trend in ordering of the results that was expected, but not intuitive by inspection of the results displayed in Figure 7. Using the mean results, it can be seen the base case had the most Blue losses, alternative 1 had fewer Blue losses, and alternative 2 had the least Blue losses. The ability to make multiple runs allowed a meaningful analysis with a deterministic model, even when the results were from a non-monotonic region of the solution space.

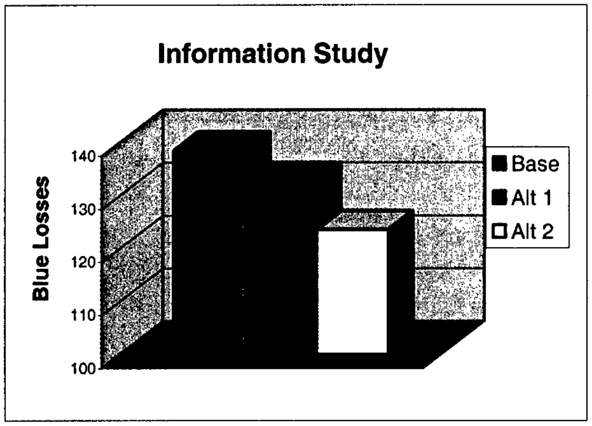

Figure 8. Mean information study Blue losses.

Although an initial sample size of 10 was used and it seemed to provide meaningful results, a look at this sample size and the confidence level calculations are in order. The steps in Figure 4 were performed for the information study and found that the base case had the greatest variance and accordingly required the largest sample size. Figure 9 is a plot of the required sample sizes for the information study base case, where the confidence level **1-**α and the error shift in the mean **e** are parameterized.

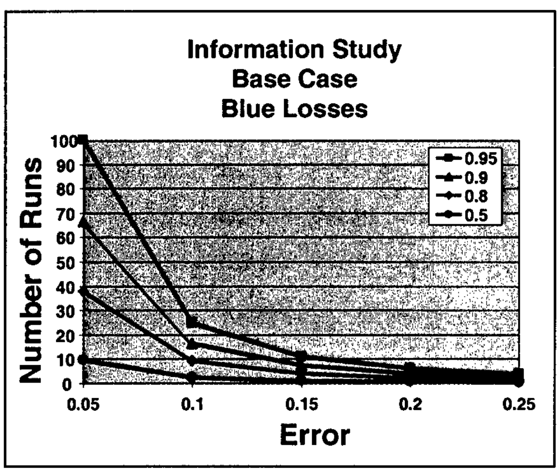

Figure 9. Required sample size.

Inspection of Figure 9 shows that 10 runs correspond to an 80% confidence level (green line) of detecting a 10% shift in the mean. The figure also shows that 17 runs corresponds to a 90% confident (blue line) of detecting a 10% shift, and 25 runs correspond to a 95% confidence (red line). Similarly, Figure 9 also provides the number of runs required for a smaller (5%) or larger (15%, 20%, and 25%) shifts in the mean. Again, one should perform this procedure on every measure of merit used in the analysis and the required number of runs will be the largest computed for any of the measures.

THE WAY AHEAD

The TRAC studies presented at Figure 10 successfully employed this multiple runs method using the deterministic VIC combat model to analyze a variety of Army attrition and information issues.

- *The Joint Sensor-To-Shooter Battle Management Study*
- *The 1st Infantry Division Limited Conversion to Division XXI*
- *The US/UK Sensor-To-Shooter Multinational C4 Interoperability Study*
- *Comanche Analysis of Alternatives*
- *Joint Contingency Force Advanced Warfighting Experiment Above Brigade Analysis*
- *TRACER/FSCS Affordability Review*

Figure 10. Successful multiple run studies.

MAJ Ross W. Snare III is exploring the mathematical foundation for this analytical method in his doctorial dissertation entitled *An Alternative Method of Conducting Analysis with Deterministic Combat Simulations.*

TRAC has adopted multiple runs of VIC as a standard operating procedure for studies whose results are in a non-monotonic region of the solution space.

SUMMARY

Clausewitz's concept of *general friction* tells us that chance events in war can turn success into failure and vice versa. Good deterministic combat models, like VIC, demonstrate that chance events can drive their results, even to the extent of chance reversal of outcomes (non-monotonic behavior). This is not a problem with these models that needs to be fixed; it is a characteristic of modeling Clausewitzian war. Since the distribution of information around the battlefield is one of the main causes of *general friction*, one should expect, especially with the current and anticipated future analytical emphasis on the impact of information on combat operations, an increase in the frequency of occurrences when chance events influence combat outcomes or, in other words, produce non-monotonic results.

TRAC successfully employs multiple runs of the deterministic VIC model to analyze information studies and other analyses, in which combat outcomes fall into non-monotonic regions of their solution spaces. The method employed at TRAC is anchored solidly in the foundations of mathematics and statistics. The key enabling technology, however, for the method is today's computer technology, which provides the speed and storage required to make multiple runs of proven, deterministic models. Technology now provides the means to handle the non-monotonic combat outcomes of Clausewitzian war using detailed deterministic division and corps levels combat simulations.

ACRONYM LIST

C4	Command and control, communications, and computers
FSCS	Future Scout and Cavalry System
PC	Personal computer
TRAC	TRADOC Analysis Center
TRACER	Tactical Reconnaissance Armored Combat Equipment Requirement
TRADOC	Training and Doctrine Command
UK	United Kingdom
US	United States
VIC	Vector-In-Commander

REFERENCE LIST

Devore, Jay L., 1982, *Probability & Statistics for Engineering and the Sciences*, Brooks/Cole Publishing Company, Monterey, California.

Dewar, J. A., J. J. Gillogly, and M. L. Juncosa, 1996, "Non-Monotonicity, Chaos and Combat Models," *Military Operations Research Journal,* Volume 2 Number 2.

Walpole, R. E. and R. H. Myers, 1985, *Probability and Statistics for Engineers and Scientists*, Third Edition, MacMillan Publishing Company, New York.

Watts, Barry D., "Clausewitzian Friction and Future War," 1996, *National Defense University,* McNair Paper Number 52.

DESCRIPTORS

Non-monotonic
Clausewitz
Combat simulation
Vector-In-Commander
Deterministic
Multiple runs
Confidence Intervals
Sample size