*Proceedings of the 1989 Winter Simulation Conference*
E.A. MacNair, K.J. Musselman, P. Heidelberger (eds.)

# THE CONSTRAINT MODEL OF ATTRITION

Dean S. Hartley III
Oak Ridge National Laboratories
ORGDP, K1001
PO Box 2003, MS 7180
Oak Ridge, TN 37831-7180

ABSTRACT

Helmbold demonstrated a relationship between a ratio containing initial force sizes and casualties, herein called the Helmbold ratio, and the initial force ratio in a large number of historical battles. This paper examines some of the complexity of the Helmbold ratio using analytical and simulation techniques and demonstrates that a constraint model of attrition captures some aspects of historical data. The effect that the constraint model would have on warfare modeling is uncertain. However, some speculation has been attempted concerning its use in large scale simulations.

## 1. INTRODUCTION

Many computer models of warfare use an attrition model as the driver of the war model. Attrition levels are used to determine which side breaks off from a battle, which side wins, the rate of advance (or retreat) of the attacker, etc. Attrition algorithms range from simple percent per day calculations to complex heterogeneous Lanchester difference equations. Criticisms and defenses of various attrition algorithms range from discussions of face validity (military plausibility) to analyses of the sequence of events causing attrition (target acquisition, time to fire the first round, probability of hit, probability of kill given a hit, etc.) to historical validation attempts (Engel, 1954; Busse, 1969; Hartley, 1989).

The constraint model of attrition is proposed where an alternative to the uncertain validity of causal relationship attrition models is desired. Its value lies in a correlation with reality that does not rely on an explicit causal relationship. The constraint model also requires only a simple set of calculations to determine its predictions.

Helmbold's ratio is one of many possible parameters in the study of attrition rates in warfare. The motivation for examining this ratio arises from the apparent relationship of its natural logarithm with the natural logarithm of the force ratio ($x_0/y_0$), described by Helmbold (1961, 1964, 1971a). The Helmbold ratio is defined in Equation [1]. The meanings for the variables are given in Table 1.

$$(x_0^2 - (x_0 - casx)^2)/(y_0^2 - (y_0 - casy)^2) \ . \qquad [1]$$

Table 1: Variable definitions

| | |
|---|---|
| $x_0$ | initial force size for side x |
| $y_0$ | initial force size for side y |
| casx | casualties for side x |
| casy | casualties for side y |
| casratx | casx/$x_0$, casualty ratio for x |
| casraty | casy/$y_0$, casualty ratio for y |
| kilratx | casy/$x_0$, kill rate for side x |
| kilraty | casx/$y_0$, kill rate for side y |
| exchrat | casx/casy, casualty exchange ratio |
| forrat | $x_0/y_0$, force ratio |
| helmrat | the Helmbold ratio |
| lforrat | ln(forrat) |
| lhelmrat | ln(helmrat) |

The general form of Helmbold's relationship is shown in Eq. [2].

$$\ln(\text{helmrat}) = \alpha * \ln(x_0/y_0) + \beta \ . \qquad [2]$$

An earlier paper (Hartley and Kruse, 1989) investigates the role that $\alpha$ and $\beta$ have in describing the appropriate attrition algorithms for modeling warfare. The evidence points to a mixed linear-logarithmic law as one candidate; however, other, non-homogeneous Lanchester laws are not ruled out.

Helmbold space (the space in which the Helmbold relationship is graphed) is radially symmetric with respect to the assignment of x and y to the two sides. If a set of data were found which formed a line through the origin, random switching of x for y in the data would not change the equation for the line. However, the regression fit for any other set of data is, at least in part, a function of the choice as to which side in each battle is called x. Thus, the relationship Helmbold found is dependent on always assigning the attacker the role of x.

This paper examines the mathematical implications of Helmbold's ratio to aid in defining meaningful military constraints on force ratios and casualties that may lead to the historical results described by Helmbold. The immediate goal for the investigation is to discover boundary conditions (simply stated in terms of casx, casy, $x_0$, and $y_0$) that result in linearly (or nearly linearly) bounded regions in Helmbold space ($\ln(x_0/y_0)$ vs ln(helmrat)). The ultimate goal is to construct a constraint model of attrition and determine its plausibility.

## 2. THE SHAPE OF LN(HELMRAT)

Although the formulation of the Helmbold Ratio contains four variables, a complete analysis only requires three variables. This simplification is accomplished by choosing one variable ($y_0$, without loss of generality) as the basis variable: setting its value to be 1.0. The values of all other variables are denominated in terms of this variable. Thus, the value of $x_0$ is given as a fraction of $y_0$ (e.g., 0.5 or 1.5); casy ranges from 0.0 to 1.0; and casx ranges from 0.0 to $x_0$.

Figure 1 illustrates the shape of ln(helmrat) as a function of casx and casy, with $x_0$=1.0. The function rises asymptotically as casy approaches 0.0 and falls asymptotically as casx approaches 0.0. The function's value when casx=0.0 and casy=0.0 is ambiguous. The function increases with $x_0$. Because ln(helmrat) is not defined for casx = 0 or casy = 0, the figure can only indicate the direction of the behavior of the function near these values.

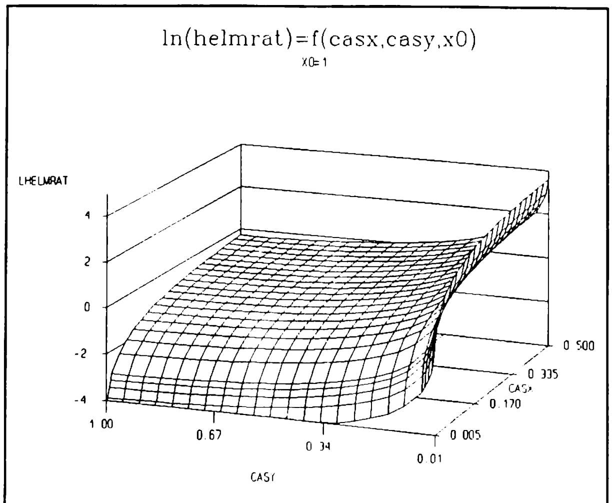

Figure 1: Ln(helmrat) at x0=1.0

Figure 2 is provided as an aid to the visualization of the behavior of ln(helmrat). The view is comparable to looking straight down on Fig. 1. The curved lines are contours of equal value of the function. These contours are of interest in restricting the range of ln(helmrat) by restricting its domain. A restriction based on the contours would produce an easily calculable range of values. Unfortunately, the function for the contours is more complex than is desirable. Desirable functions are linear relations among simple military values, such as those shown in Table 1.

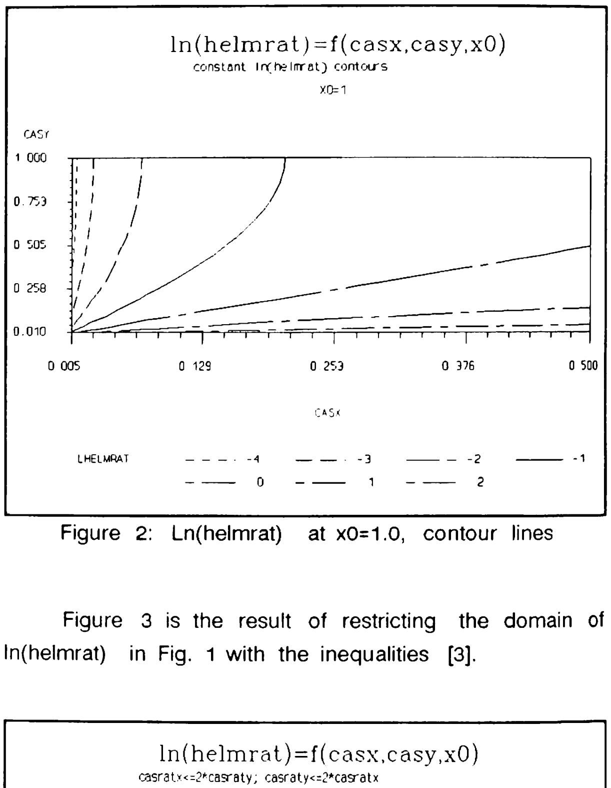

Figure 2: Ln(helmrat) at x0=1.0, contour lines

Figure 3 is the result of restricting the domain of ln(helmrat) in Fig. 1 with the inequalities [3].

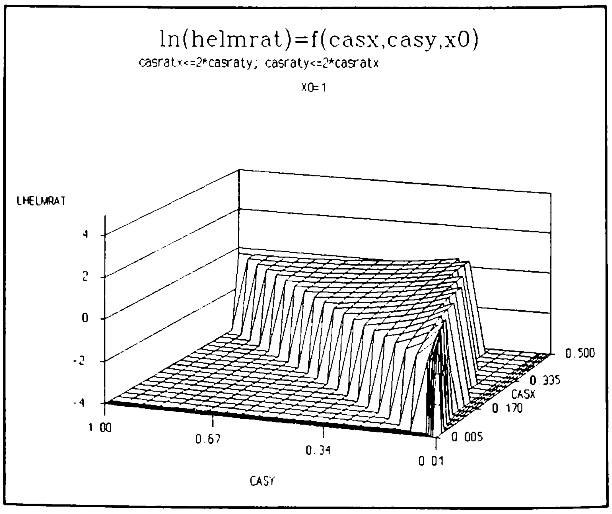

Figure 3: Constrained ln(helmrat)

casratx <= 2*casraty ,
casraty <= 2*casratx . [3]

Although the restrictions of the inequalities impose upper and lower limits on the allowed values of ln(helmrat), discovering the precise limits is more difficult than it would be if the constraints followed the constant valued contours.

## 3. CONSTRAINT DEFINITIONS

Several constraints are possible in addition to the pair described in inequalities [3]. The most frequently described constraint is that three times the number of attackers (compared to defenders) are needed (or sufficient) for a successful assault. Constant application of this rule would result in force ratios of 3 or larger (ln(forrat)>=1.1). This rule is commonly quoted, although not consistently observed (or justified) in practice (Helmbold, 1969). A second principle of war is called economy of force. This principle acts in the other direction: use no larger force than necessary. If a 7 to 1 force ratio is always sufficient, one might postulate a restriction that no battle will take place where this force ratio is exceeded in either direction (7:1 or 1:7, ln(forrat)=1.95 or -1.95). Hence the constraints [4] might be valid.

ln(forrat) <= 2 ,
ln(forrat) >= -2 . [4]

Constraints [4] yield a vertical set of constraints in Helmbold space, restricting the relative force sizes. Changing the values shifts the constraints left or right.

In actual engagements, there is often a difficulty in determining the appropriate force ratio. A local advantage may exist at the same time that there is rough parity along the front as a whole. For theoretical purposes, the force ratio will be defined as that of the actual engaged forces. Historically reported battles may or may not use this definition. This is one of the problems with the historical data. Historical reporting may cause other distortions. For instance, historical engagements with extremely unbalanced force ratios may never be reported as battles, resulting in an apparent constraint (such as [4]) where none exists.

Another potentially useful set of constraints would yield horizontal constraints. Constraints [5] produce horizontal constraints, effectively restricting the allowed Helmbold ratios for battles.

kilratx <= 2*kilraty ,
kilraty <= 2*kilratx . [5]

The first member of this pair may be taken to mean that side y is willing to accept situations no worse than when side x has a kill rate twice that of its own kill rate. The second member is the converse.

Another desirable constraint would be one that produces data points with a slope of 1.0. Replacing the kill rate variables in [5] with the raw casualty variables produces the desired result. These constraints are recast as exchange ratio constraints in [6] and the scatter plot is shown in Fig. 4 (region='in' shows points that satisfy [6]).

exchrat <= 2 ,
exchrat >= 0.5 . [6]

The first member of this pair may be taken to mean that side x is willing to accept no more than twice the casualties of side y, with the second member as the converse.

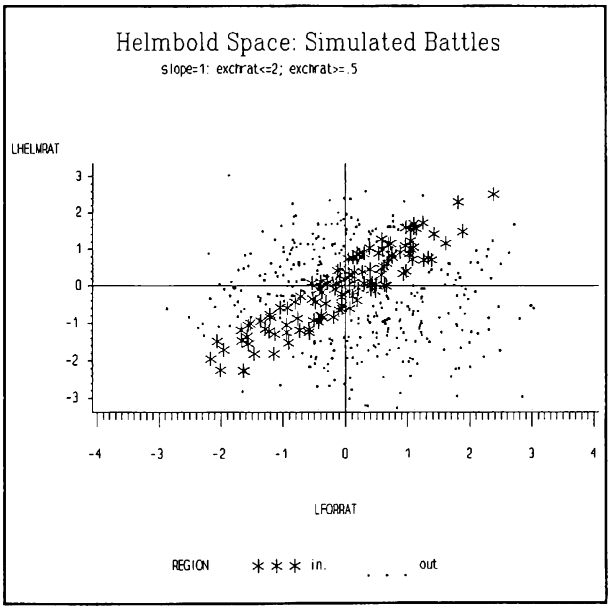


Figure 4: Constraints with slope=1

A fourth desirable constraint would yield a slope of 2.0. The constraints of [3] satisfy this need. In this case each side is concerned that its casualty ratio be no worse than twice its opponent's casualty ratio.

It should be noted that the definition of whether a given point in Helmbold space is included or excluded by [3], [5], or [6] is ambiguous. For instance, consider constraints [6]. Given $x_0$, and hence ln(forrat), there are multiple pairs (casx, casy) that produce a single ln(helmrat) value. Because the underlying equations do not precisely follow contours (in Fig. 2) of constant ln(helmrat), some of the pairs on a contour may meet the constraint and some may not. However, the constraint equations are close enough to the contours that the areas of fuzziness are small. Some of the points in Fig. 4 that are on the borders between inclusion and exclusion show dual symbols for this reason.

Constraints [3-6] supply a basis from which a wide variety of slopes of constrained data can be produced. Figure 5 illustrates the result of pairing slope 1.0 and slope 2.0 constraints [7] on the same Helmbold space data as in Fig. 4 to produce a set of data with slope 1.2. Other slopes are also possible.

$$\begin{aligned} exchrat &\le 4.5 , \\ exchrat &\ge 1.0 , \\ casraty &\le 2*casratx , \\ casratx &\le 5*casraty . \end{aligned} \qquad [7]$$

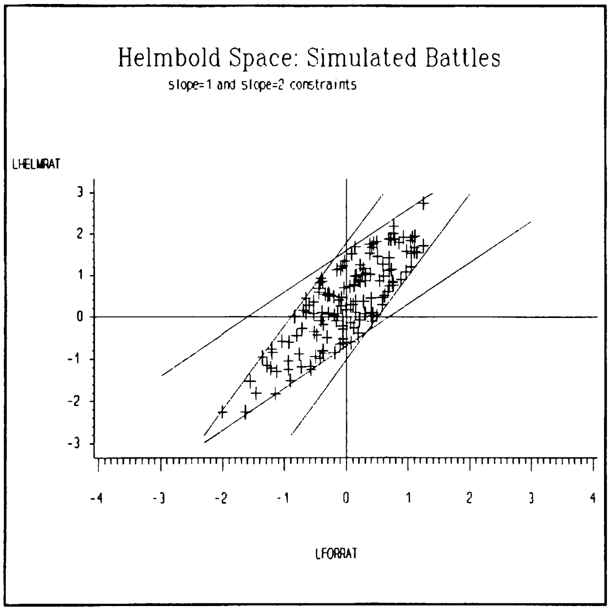

Figure 5: Four constraints producing slope=1.2 data

This slope approximates the slope of the data from 92 historical battles, mentioned in Helmbold (1961) and described more fully in Helmbold (1971a). The data are arranged with x as the attacker and y as the defender. Thus, a possible military interpretation-doctrine-verbalization of the formulae of the constraints [7] might be as follows:

1. The attacker will definitely break off the attack if his casualties reach a level that is 4.5 times that of his estimate of the defender's casualties.

2. The attacker will not consider breaking off the attack until his casualties are at least equal to his estimate of the defender's casualties.

3. The defender will certainly retreat if his casualty ratio exceeds twice his estimate of the attacker's ratio.

4. The attacker will certainly break off the attack if his casualty ratio exceeds five times his estimate of the defender's casualty ratio.

This interpretation has been structured with the defender basing his break contact decision on one comparison while the attacker bases his break contact decision on three comparisons. Other interpretations involving decisions to engage are possible. Note that in each case there is an estimate of the enemy's casualties (and, in the latter two constraints, initial force). (The size of one's own casualty figure may also be an estimate.) This means that the operations of such constraints in real battles would be imprecise. One would not expect to see sharp edges as in Fig. 5.

## 4. TESTING THE CONSTRAINT MODEL

A proposal that the four constraints of [7] provide a valid model for Helmbold's 92 historical battles implies that the appropriate corresponding military interpretations of the constraints are historically practiced principles of war! (Some fuzziness in the application of the constraints is allowed.) Moreover, Helmbold (1971b) has shown that a breakpoint hypothesis is incompatible with historical data.

The second consideration is dealt with first. Helmbold has provided a very closely reasoned demonstration of the implications of "breakpoint hypotheses." He tested these implications against historical data and found the data contradicted the implications. He clearly defined his assumptions and made some preliminary investigations into the effects of altering some of these assumptions. He did not, however, alter his first assumption that a side's breakpoint is solely determined by its own casualty ratio. The constraints [3, 4, 5, or 6] and their combinations are not breakpoint hypotheses under Helmbold's definition and are not affected by Helmbold's conclusions.

The first point is of greater concern, the military interpretations of constraints of types [3 - 6] may be too complex to be believable as considerations of a commander in the heat of battle. It is true that a commander rarely has a clear picture of the enemy's casualties. However, the amount of enemy fire may be a surrogate for an estimate of the enemy casualties, allowing this factor to be integrated into a commander's decision process. In any event, the historical evidence that certain regions in Helmbold space appear to be forbidden zones for battles may require a reexamination of the thought processes of military commanders.

Fifty replications (runs) of a simulation of the 92 historical battles were made. Each run consisted of 92 data points, satisfying constraints [7], randomly selected from Helmbold space. The $R^2$, $\alpha$, and $\beta$ values were calculated for each run. The $\alpha$ value represents the slope of the regression performed on the data and the $\beta$ value represents the intercept on the ln(helmrat) axis. Table 2 displays the minimum, maximum, mean, and standard deviation for these values over the set of 50 runs and the actual values for the historical sample.

Table 2: Historical vs constraint statistics

| Dataset | $\alpha$ | $\beta$ | $R^2$ |
|---|---|---|---|
| 92 Battles | 1.29 | 0.21 | 0.54 |
| 50 runs: means | 1.20 | 0.34 | 0.66 |
| 50 runs: minimums | 1.07 | 0.17 | 0.54 |
| 50 runs: maximums | 1.34 | 0.47 | 0.78 |
| 50 runs: std devs | 0.07 | 0.06 | 0.06 |

This set of statistics implies that the historical data could be produced by the action of these constraints. A random set of battles, constrained by [7], could yield the statistics generated by the regression of the actual historical data. The slopes, intercepts, and the spread (indicated by the $R^2$ values) of the runs are close enough to lend credibility to the supposition that constraints similar to [7] could be the operative factor in the battle outcomes.

Figure 6 superimposes the constraints on the historical data in Helmbold space. This picture is less convincing than the statistical evidence. Certainly, the definition of the constraints includes errors in judgment as to the true state of the battle. Such errors allow battles to fall outside of the constraint area. However, the shape of the distribution of the data does not appear diamond shaped.

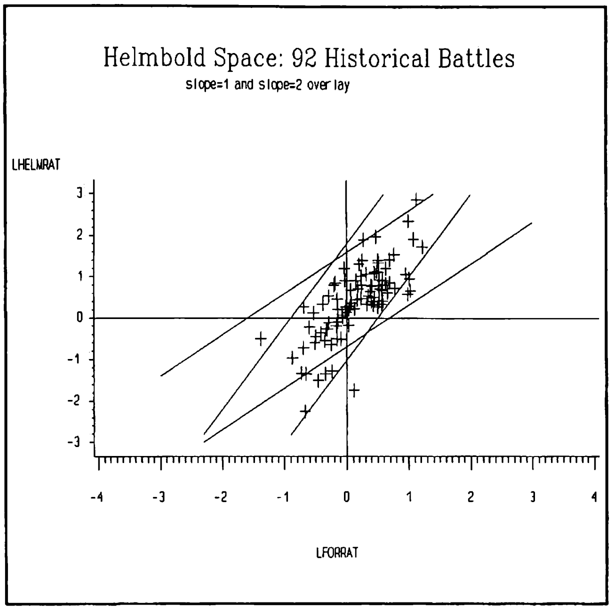


Figure 6: 92 battles, four constraints

Alternative constraints can be constructed that have slopes approximating that of the data. The constraints [8] have slopes of 1.5, defined so the 92 battles are contained within two constraints, rather than four. Figure 7 shows the

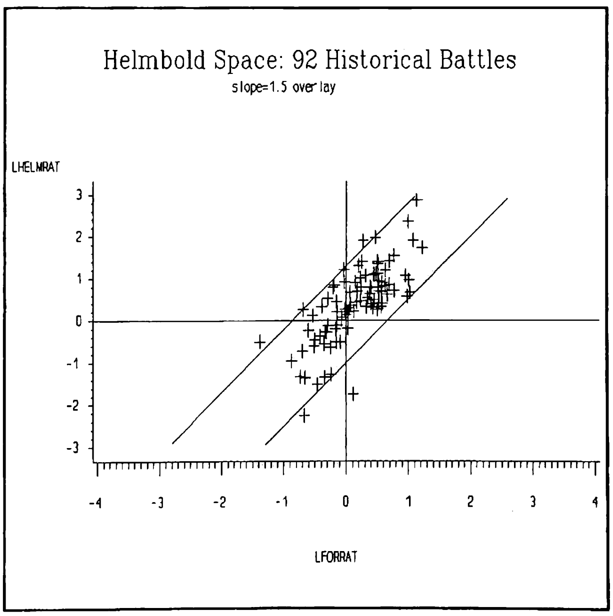


Figure 7: 92 battles, two constraints

battle data with these constraints superimposed. The visual results are more satisfactory than those of Fig. 6.

casraty*casy <= 2.8*casratx*casx ,
casratx*casx <= 3.7*casraty*casy . [8]

The problem with constraints [8] is not mathematical, but military. What is a reasonable military interpretation of these constraints that might conceivably be used in battle to determine the commander's actions? It does not take a great leap of faith to suppose that verbal precepts such as "massing forces" and "economy of force" may have mathematical expressions and that commanders drilled in such precepts would take actions that result in mathematical constraints on battles' outcomes. However, it seems unlikely that commanders through the ages performed the calculations suggested in [8] before deciding whether to break off an attack or to retreat from a defensive position.

Notwithstanding the arguments above, it is clear that there are "forbidden" regions in Helmbold space. The upper left and lower right portions of Helmbold space are not populated with historical battles. Whether or not historical commanders' though processes included the equivalent to constraints [8], those constraints appear to describe the results.

## 5. MODELING IMPLICATIONS

The modeling concept developed above may be formalized as the constraint model of attrition, defined by four assumptions.

- A1. The factors that influence the actual attrition results in battle are varied enough and contain sufficient random factors so that the overall effect is that the attrition value is a random variable.
- A2. The appropriate space for visualizing the attrition results is the two-dimensional space of the logarithm of the force ratio and the logarithm of the Helmbold ratio - Helmbold space.
- A3. Within Helmbold space, potential battles are normally distributed with regard to ln(force ratio) and normally distributed with regard to ln(Helmbold ratio).
- A4. Actual battles correspond to potential battles that remain after constraints eliminate portions of Helmbold space.

The constraint model does not assume that the generalizations of the four sets of constraints [3, 4, 5, 6] form a complete set of the militarily valid constraints (as the addition of [8] showed), or that those four (or five) are in fact militarily applied constraints. Their existence is proposed merely to provide a plausible set of constraints, having the property of allowing an approximation of observed historical data.

The constraint model should have no effect on very fine resolution modeling. Models that are based on firm physical and human factors, such as the effects of maneuver and fields of fire and the human use of maneuver and fields of fire, have no need of a statistical approximation to the effects of battle. Data of this sort might be gained through analysis of the National Training Center (NTC) battalion training activities. Company- and lower-sized models will continue to model explicitly many of the factors (e.g., individual weapon target acquisition) that contribute to the randomness of higher level results. Other factors contributing to higher level randomness do not apply to fine resolution models, e.g., the effects of Company and Platoon sized fire and maneuver tactics. Theater level models might use the constraint model profitably within a stochastic context. Mid-level models might pose the most difficult problem.

Using the constraint model in a theater level model would yield some interesting changes. Each engagement determines a unique force ratio. To determine the resulting attrition, information on the operative constraints is required. These constraints might be taken from historical data, with no attempt at rationalizing them, might be the result of the model builder's logical model, or might be the result of the model user's logical model. The model constraints might be fixed for the entire run or might be variable depending on the current state-space of the model. Given the constraints, the proper Helmbold ratio could be selected randomly. Because of the formulation of the Helmbold ratio, the casualties for one of the sides would be randomly selected with the casualties for the other side following from the formula. Another method of implementation model would consist of the superposition of the constraint model on top of an existing attrition model, e.g., a time-stepped heterogeneous Lanchester algorithm. The constraints would act as stopping rules, enforcing conformance to the boundaries. (The distribution of battles would not look like the historical distribution, as the model battles would all cluster on the boundaries, avoiding the center.)

The most interesting implication of this model of attrition to theater level models is the next step. The engagement of the forces is complete with the casualty determination; however, no resolution has occurred. The theater level model must deal explicitly with the logical problems of what happens next. Because the engagement is over, the forces must either physically disengage or another force must enter the engagement, creating a new engagement. Where another force is not available, rules must be developed to determine which force retreats. (If a time based attrition model were being used, the attrition would be calculated until it is stopped by a constraint. The particular stopping constraint encountered would determine the side that breaks, and thus loses.) If another force is available, rules must be developed to determine whether that force is employed. Further, the length of the engagement is not implicit in the constraint model and must be determined externally.

Current models have logic that perform the analogous movement and time functions; however, the rules are often hidden in the code and may not be subject to the scrutiny they deserve. Thus the constraint model might surface these as the important areas of modeling concern, rather than concern over the killing rates of weapons systems.

6. CONCLUSIONS

It has been shown that the constraint model of attrition can reproduce one historical parameter, position in Helmbold space. Plausible military interpretations for some constraints have been determined. The results derived from this research do not imply that the constraint model of attrition is valid in a causal sense; however, they do show that this validity is conceivable.

More important than questions of causality is the utility of the causality model as a simple method of generating results that relate to historical results. The constraint model of attrition is proposed where an alternative to the uncertain validity of causal relationship attrition models is desired. Its value lies in a correlation with reality that does not rely on an explicit causal relationship and a simple set of calculations to determine its predictions.

A non-causal interpretation of the constraint model leaves open the question of the causal mechanisms of attrition. The visual judgment that the two constraint model fits the 92 historical battles better than the four constraint model, together with the problem of military interpretations of the two constraint model supports the contention that the constraint model should not be causally interpreted. The mixed law Lanchestrian formulation of the earlier paper (4) may be superior, both in fitting the data and in providing an historically validated causal model.

The effect that the constraint model would have on warfare modeling is uncertain. However, it should be easy to implement in large scale simulations. It also has the advantages mentioned earlier of separating attrition from movement and victory.

ACKNOWLEDGEMENT

Research sponsored by the Joint Staff/J-8 under Interagency Agreement 1612-1612-A1 under Martin Marietta Energy Systems, Inc., Contract No. DE-AC05-84OR21400 with the U.S. Department of Energy.



REFERENCES

Busse, James J. (1969). An Attempt to Verify Lanchester's Equations. 3rd Israel Conference on Operations Research, Vol. 2., 587-597.

Engel, J. H. (1954). A Verification of Lanchester's Law. Operations Research 2, 2, 163-171.

Hartley, D. S. III (1989). Can the Square Law Be Validated?, Martin Marietta Energy Systems K/DSRD-57.

Hartley, D. S. III and Kruse, K. L. (1989). Historical Support for a Mixed Law Lanchestrian Attrition Model: Helmbold's Ratio, Martin Marietta Energy Systems K/DSRD-113.

Helmbold, Robert L. (1961). Historical Data and Lanchester's Theory of Combat, Combat Operations Research Group CORG-SP-128, DTIC AD 480 975.

Helmbold, Robert L. (1964). Historical Data and Lanchester's Theory of Combat Part II, Combat Operations Research Group CORG-SP-190, DTIC AD 480 109.

Helmbold, Robert L. (1969). Probability of Victory in Land Combat as Related to Force Ratio, The Rand Corporation P-4199, DTIC AD 696 489.

Helmbold, Robert L. (1971a). Air Battles and Ground Battles--A Common Pattern?, DTIC AD 718 975.

Helmbold, Robert L. (1971b). Decision in Battle: Breakpoint Hypotheses and Engagement Termination Data, The Rand Corporation R-772-PR, DTIC AD 729 769.

## AUTHOR'S BIOGRAPHY

DEAN S. HARTLEY III is a member of the research staff of Martin Marietta Energy Systems, Inc., the operator of the Department of Energy Facilities at Oak Ridge. He received his Ph.D. in mathematics (piecewise linear topology) from the University of Georgia and has worked in the field of operations research, both industrial and military for over 20 years. At Oak Ridge, he directs research in modeling, simulation, and gaming (MSG) and general military operations research. He is a joint ORSA/TIMS member, a member of MORS, and a Council Member of the Military Applications Section of ORSA.

Oak Ridge National Laboratories
ORGDP, K1001
PO Box 2003, MS 7180
Oak Ridge, TN 37831-7180
(615) 574-7670