Entropy in Combat Data

F. Carvalho-Rodrigues
Laboratório Nacional de Engenharia a Technologia Industrial
1699 Lisbon Codex, Portugal

John T. Dockery
Joint Staff, J6, Pentagon
Washington, DC 20318-6000, USA
Alexander E. R. Woodcock
S1 nectics Corporation, 10400 Eaton Place
Fairfax, Virginia 22030, USA

Abstract

We present arguments that g_ Shannon based entropy, computed from combat casualty data, can be used as a predictor 2f combat outcome Data is taken from historical records, field exercises and from. mathematical simulation. It ***is further*** *argued that the selected entropy is a "true" system parameter in that it is nowhere localized in the components but is resident only in the system interactions among components. This contribution reprises and anticipates the results of a more comprehensive article submitted for journal publication.*

Introduction

This paper and a companion contribution to this symposium explore a different way of looking at combat data. Combat .is, viewed as a complex system interaction mediated by the enveloping environment.. This is a viewpoint not unlike the cybernetic based approaches to combat. Viewed as a system, we inquire after descriptive parameters, which reflect only these systems aspects of combat. We adopt a rigorous definition that a true system parameter is nowhere localized in the components, but is resident ón1:y in the, interaction among and between the components. In short, if the system goes away, so does the parameter.

We are, for example, looking for system analogues to such basic component parameters as *mass* and *charge*. Such parameters could be expected to be no respecters of system structure details. Examples that are being explored include entropy, resilience, cohesion, fractal measure, and complexity. We do not exclude the very real possibility that some of these parametric measures may be perceptually dependent, e.g. complexity. In this particular paper the concept of entropy is explored within the foregoing context.

The paper first briefly discusses entropy. A Shannon based entropy is proposed as the basis of a predictor of combat outcome. The paper then moves to a consideration of the historical record followed by an examination of field exercise experience. Finally, some Lanchester equation based simulations are presented in support of the concept.

The Concept of Entropy

At last count we had collected a total some fifteen situation specific entropy definitions, e.g. separate entropies for spatial and temporal aspects of a cellular automata. All purport, in one manner or another, to measure the "disorder" in a system. Two of the most well known are the thermodynamic entropy and the Boltzmann state entropy. Both of these grow without limit until the "death" of the system, and are not useful predictors. Such entropies were explored by J. Lawson in some OT his early and informal lectures on the possibilities of computing an entropy for combat. [1]

Carvalho-Rodrigues experimented in 1989 with an entropy based on Shannon's communication entropy as a measure of combat activity. [2] This entropy may be expressed as in equation 1.

$$HS = p.\ln(1/p) \qquad (1)$$

where $p = C/N$ with C-number of casualties in unit time and N-number of combatants

The attractive feature of equation 1 can be seen in Figure I below which shows that this particular form of entropy goes through a maximum value at around a value of p=37%

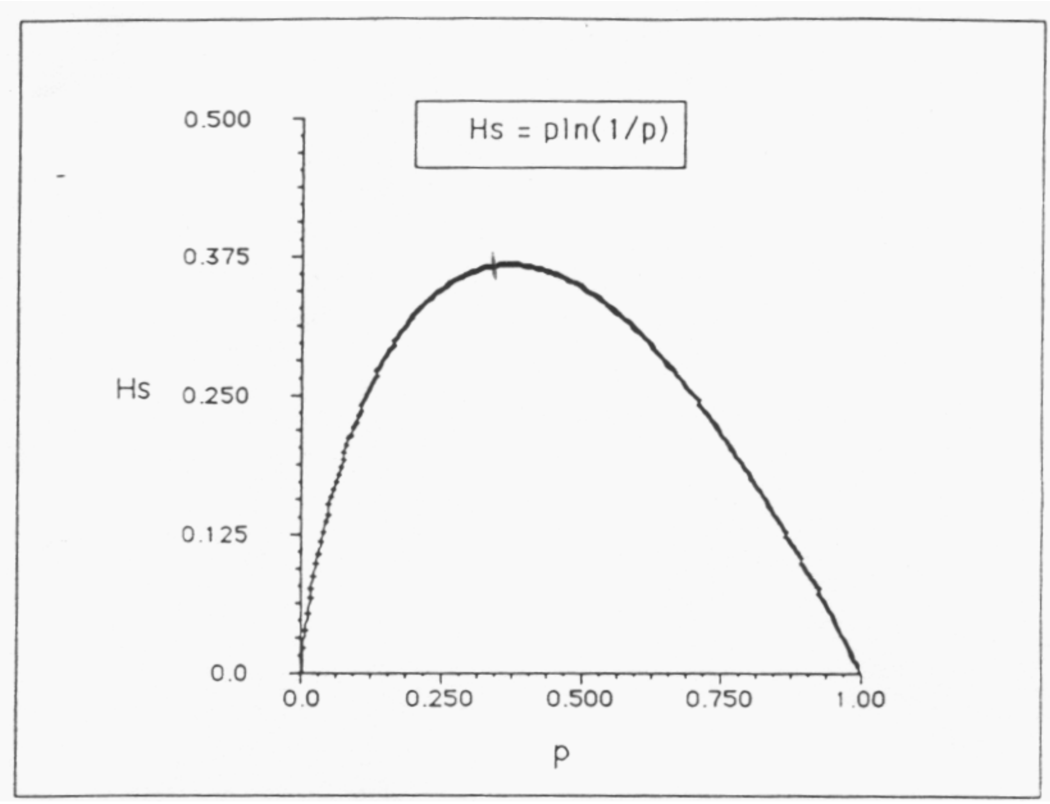

Figure 1: Curve of Hg = p.ln(1/p)

What is also interesting is that for low values of p, the value of HS attains high values as a percentage of its maximum. Thus, 10 % at 60% of peak; 20% at 84% of peak; and 25% at 92% of peak.

The working hypothesis is that adoption of the curve in Figure I as a model would help explain why combat forces are considered to be seriously degraded seemingly low values of percentage total loss. In any case equation 1 is advanced as a possible model. The remainder of this presentation consists of a series of comparisons with data from various sources beginning with the historical record.

A Comparison with Historical Data

Case 1: The Source

In order to test the overall predictability historical battles assembled by Dupuy for of Hi (i=Red and Blue) a collection of Helmbold of the US Army Concepts

Analysis Agency was consulted. (4,5]. Helmbold [5] himself has considered the question of a link between casualties and victory. He was also the source of many insightful comments. [6].

The basic data set contains detailed historical descriptions of some 601 battles from circa 1600 until circa 1970. Data on straightforward conflict without excessive maneuver, and preferably in a single assault or meeting engagement, *was* desired. It was hypothesized that if the Hi was a predictor then it would have the best chance to manifest itself under the least complicated ground combat conditions. Historical battle without complex C2 were presumed to tip such a description. Therefore, not wishing to introduce additional complexity, the following selection criteria were set, which yielded 59 battles satisfying these criteria:

• Under 10,000 combatants per side.
• Battle last ten hours or less.

The data set gives only the final casualty tallies. A few excerpts from the set are displayed in Table I following.

Table 1: Sample battles

Battle	Duration (Hrs)	Casualty Ratio-Attack	Casualty Ratio-Defense	Outcome
Culloden - 1746	0.67	1558/5400	309/9000	Defense
San Jacinto-1836	0.30-	39/743	1600/1600	Attack
Hill 272-1918	2.5	109/2950	250/2563	Attack
Rawiyeh-1967	4.0	150/5350	300/4350	Attack

The subset favored attackers by a ratio in draws were arbitrarily assigned to the of better than two to one. Three case that nominal defender. In summary: resulted

Mode	Number of Cases
Attack	38
Defense	18
Draw	3

The battles by century are shown below with a further breakout for World Wars I and II. The absence of battles from WWII is a consequence of our selection criteria.

1600's	1700's	1800's	1900's	WWI	WWII
4	13	12	30	23	1

As regards duration of combat we have

0-1 Hr	>1-2 Hr	>2-4 Hr	>4 Hr
13	10	18	18

Case 1: **The Analysis**

From the computed entropies the predictor of the combat outcome. Results quantity &I _ (Hd - Ha) was selected as the were as follows in Table II

Table: II: Results of Computation of δ_i Normalized to Unit Time

Mode	Correct	Incorrect
Attack	35	3
Defense	15	6
Totals	50	9

The results in Table II were compared with the hypothesis that the figures could have been generated at random. The chi-square was computed with one degree of freedom for a 2x2 contingency table with a value of X2 = 25.75.

By comparison the significance -level for 95% confidence is 3.84.and for 99% is 6.64.

The-results for attacker and defender entropies are displayed respectively in Figures 2 and 3.We examined the distribution of points in both figures and correlated them with outcomes. Losses for both the defender (d) and attacker (a) are predominantly beyond a value of Ci/Ni = 20-25% (i = d, a). One notes that the selected battles for the most part did not exceed the peak of the curve.

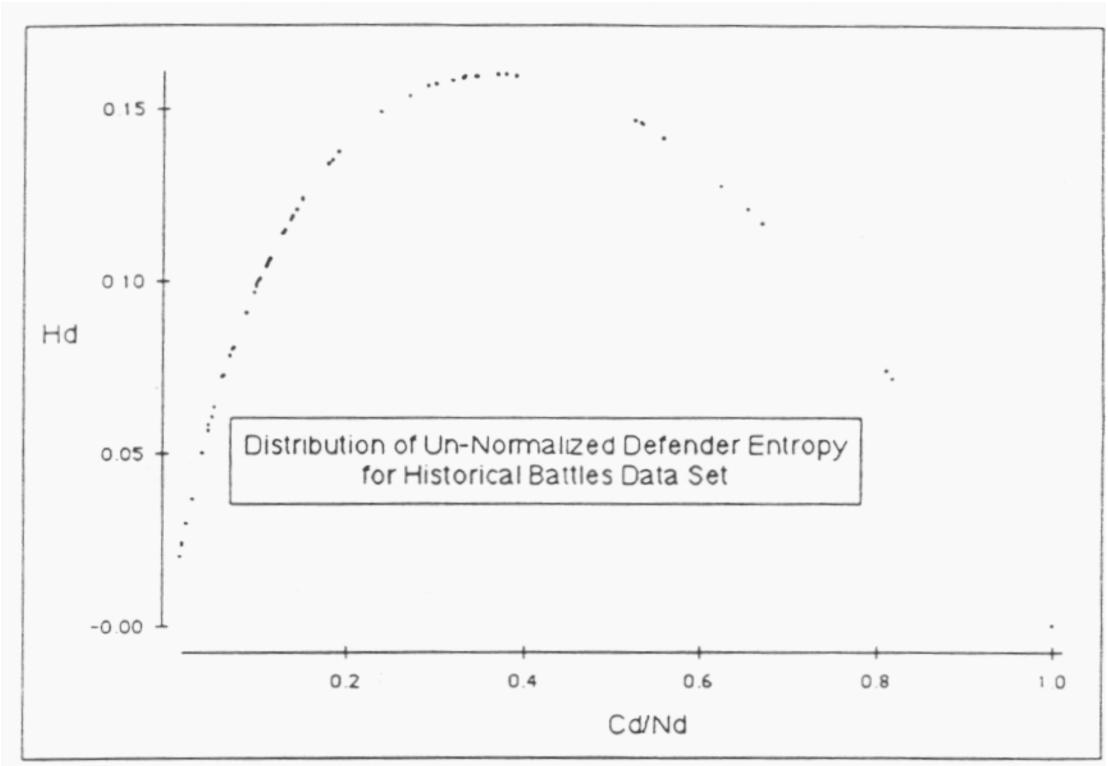

Figure 2: Plot showing the distribution of defenders' entropy for 59 historical battles
_ along the curve of p.ln(1/p)

Case 2

Data was also **collected from Operation** shown below. Entropy is computed for West Wall, a campaign near Aachen in daily time periods based on initial and final World War 11 in 1945; and from the Inchon count of the combatants during the time Operation during the Korean Conflict. period. Entropic space denotes a plot of Hd Two figures from the Inchon Operation are versus Ha-

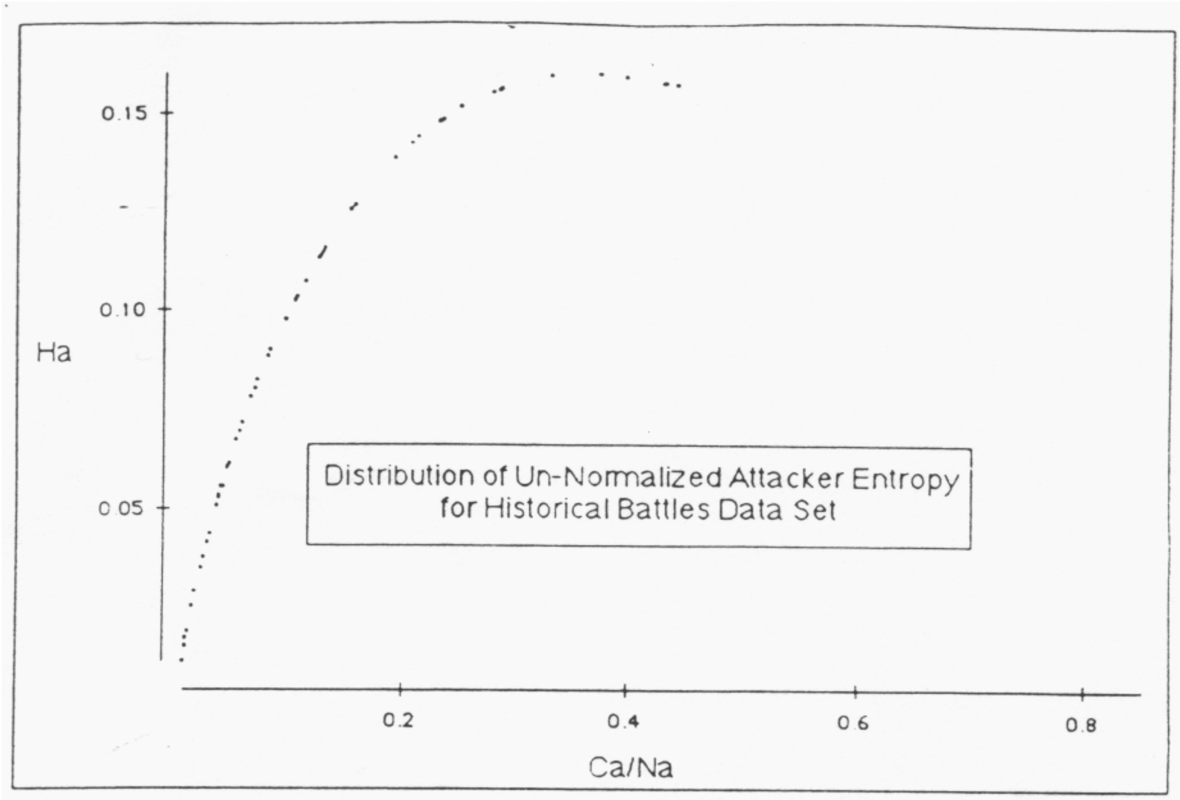

Figure 3: Plot showing the distribution of attackers' entropy for 59 historical battles along the curve of p.ln(1/p).

In Figure 4 one sees a consistently higher entropy for the defender. We also get the first glimmerings of something that showed up more clearly in the field exercise data. -That is that the attacker leads the defender in a buildup of entropy, which then must fail if the attack is to be successful. Figures 5 give a somewhat different view. In both there are are discernable clusters of points which we can correlate with three (or four) military phases of the operation. Note that these figures include reinforcements as for example near day 9 for the UN force. A link up of force occurred on day 13.

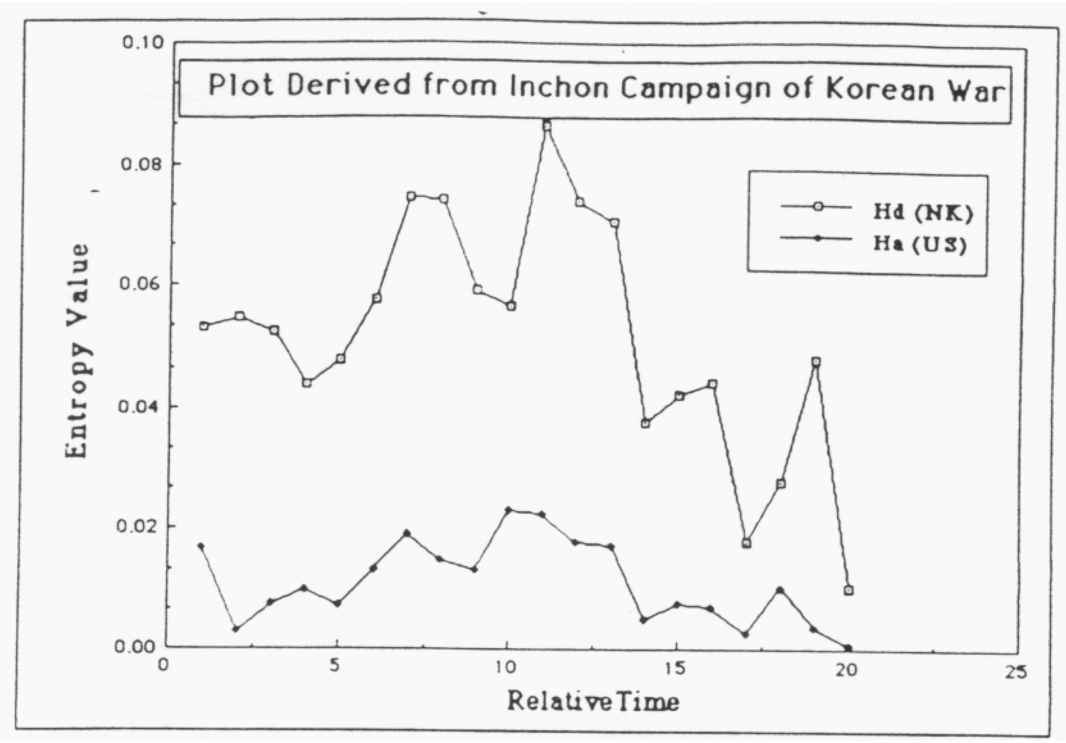

Figure 4: Time plot of the entropy changes from data taken from the Inchon Operation of the Korean War (UN forces are attacking).

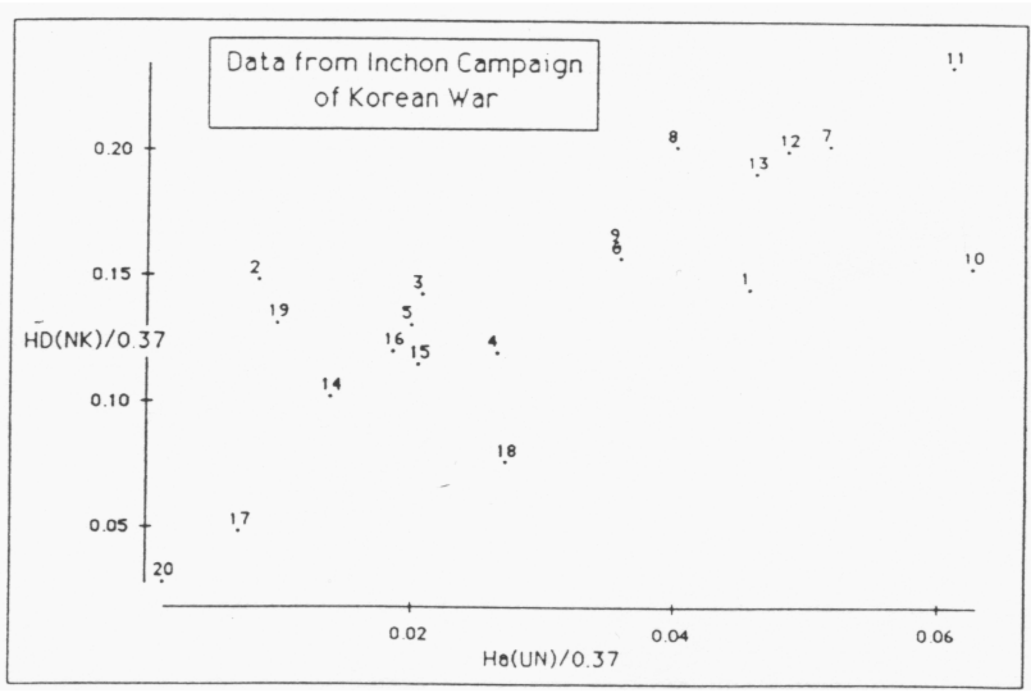

Figure 5a:Plot of data from the Inchon Operation in "entropic" space showing numbered, time ordered points.

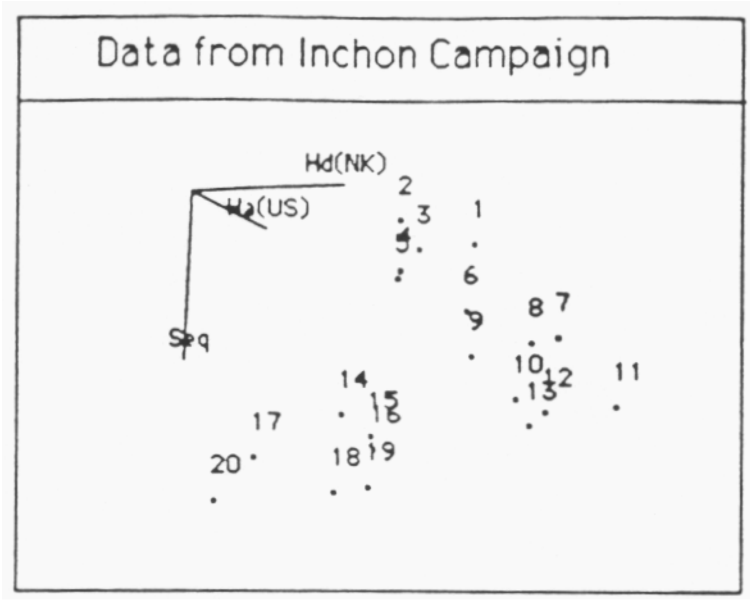

Figure 5b: Three dimensional depiction of the relative entropies versus time for data from the Inchon Operation of the Korean War. Points are numbered in sequence. (Specially scaled for graphical purposes.)

Summarv

We concluded from this portion of our taken from field exercises at the National work that there was merit in continuing an Training Center (NTC) for which the examination of entropy as a combat JANUS simulation had been used as predictor. The next source of data was sensitivity analysis tool.

Comparison with Combat Simulations of Field Exercises.

We now turn to a consideration of the analyzed in two ways. First a normalized time series data which is expected to be cumulative entropy calculation was done in more sensitive to the hypothesis that which the original number of attackers and entropy is a combat outcome predictor. It is defenders were used as the numerator in also the data from which we could hope to equation 1 for all incremental time periods a command and control predictor. tj. Second, the entropy was computed for The time series information from the NTC each successive time period using for Ni the and JANUS simulations thereof was remaining force strength at tj, which was divided into the casualties generated between tj and tj+l. Two examples are shown in Figures 6 and 7 in which the cumulative entropy is plotted for actual data and a corresponding JAN-US simulations. The NTC data is by time interval while the simulation is cumulative

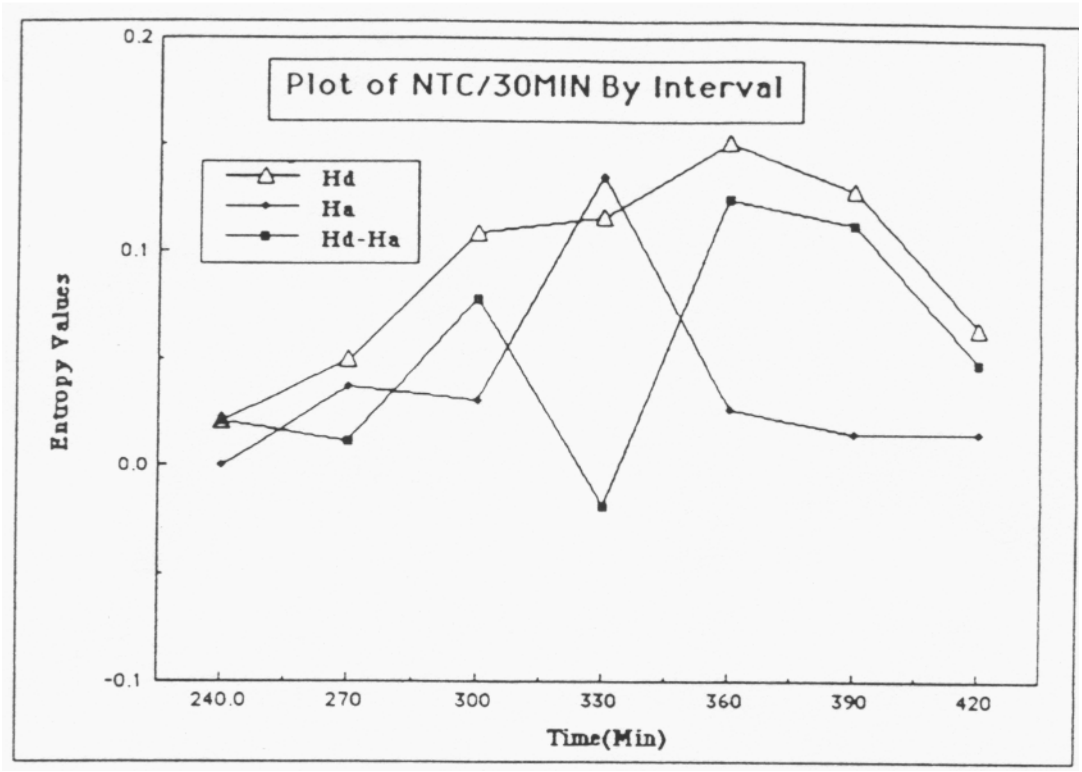

Figure 6: Plot of NTC 30 Minute time sensitive interval data for defender and attacker entropy, and their respective difference all versus time into the battle.

In Figure 6 the record shows that the attacker wins. His first contact is signaled by a steep rise in entropy followed by a decline. The defender on the other hand remains high. The spike in entropy is almost a "signature" of the attack phase. Whether it succeeds or not depends on whether the defenders' entropy remains high while that of the attacker returns to lower levels. The story is different in Figure 7. Here the attacker's contact is signaled by a rise in entropy, but there is no corresponding rise in defender entropy. The defense wins this one.

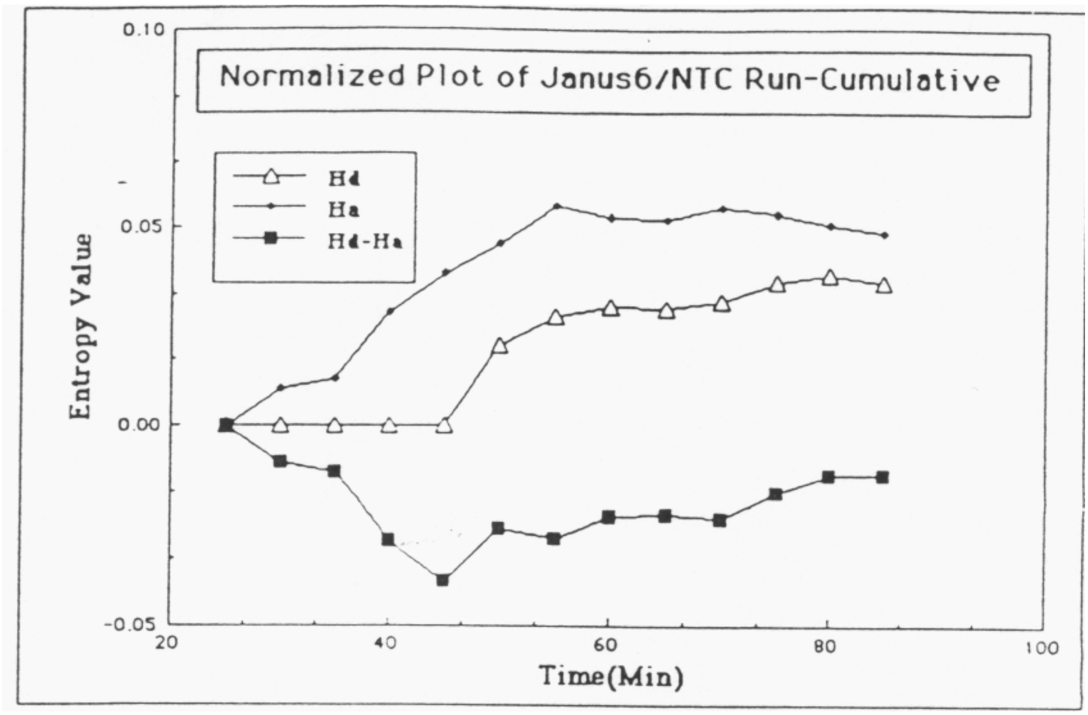

Figure 7: Plot of cumulative data from a JANUS simulation (arbitrarily numbered 6) of NTC data for attacker and defender.

A power spectral analysis was performed on the NTC and JANUS simulation to further investigate the idea of "signature" for the attack. Shown in Figures 8 are NTC data, in which the entropy computation is based on 10 minute intervals. Figure 8a is the profile of a winner with one low and dominant frequency. The defender's profile shows instead three peaks. Although the density of points is marginal for this kind of analysis, Figures 8 do show a main peak which we interpret to mean that the Red attack basically came in a single wave. This is in accord with field observations. The Blue defense appears to have one (or two) higher frequency cycles indicating more than one mode of response to the attack. The foregoing appears to verify that the entropy predictions done with this data are uncomplicated by multiple thrusts so that the engagement can be regarded as a single attack.

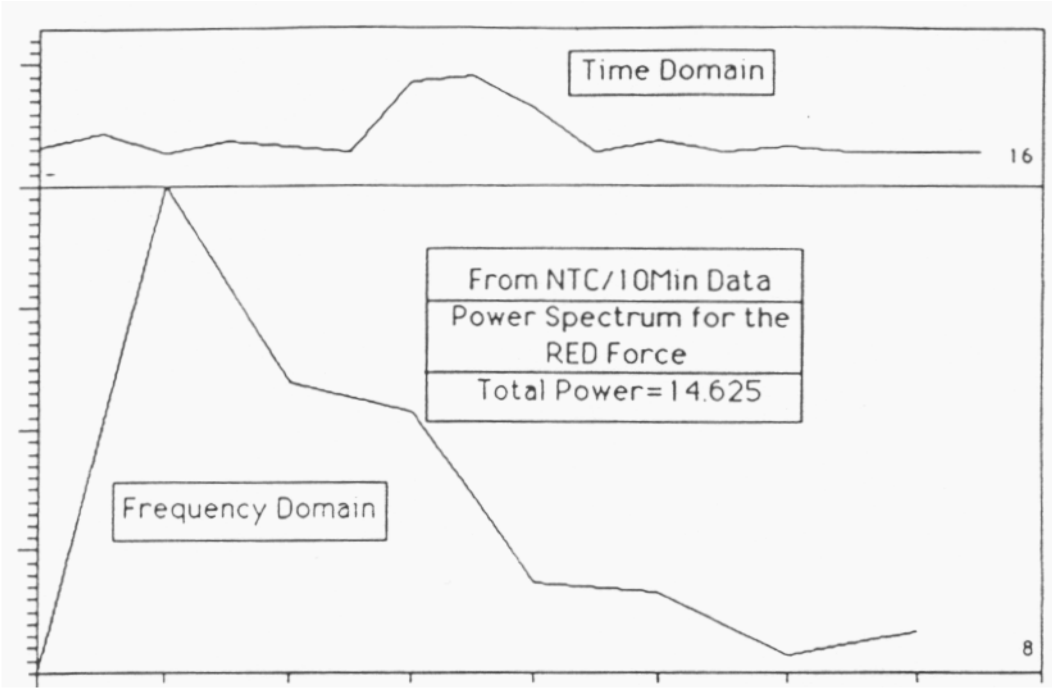

Figure 8a: Power spectral analysis of 10 minute interval data for Red (attacker) from NTC. Shown above is the time domain plot, and below it the corresponding frequency domain plot normalized to the frequency with the highest power contribution.

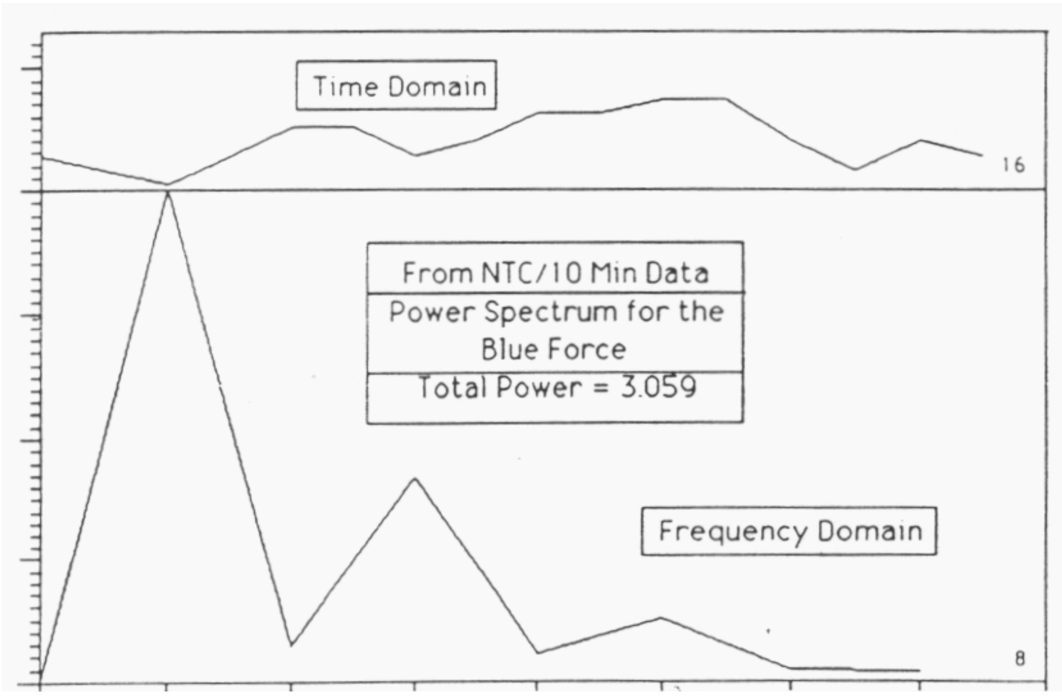

Figure 8b: As in 8a except for Blue (defender) side.

We turn finally to a test of our assertions about casualty based entropy which involves "pure" attrition based combat derived from a Lanchester equation series of simulations programmed in STELLA. The latter is icon driven software tool based on system dynamics and is totally object oriented.

Comparisons with Data from a Mathematical Model

Several scenarios were constructed of dividing through be the peak of the curve increasing complexity. For each we (0.37). The model used was as follows. computed the plot of casualty versus time as well as the plot of Hd versus Ha. (Note that the latter are sometimes normalized by

$$dx/dt = -a1y - b1xy + c1 \quad (2a)$$

$$dy/dt = -a2x - b2xy + c2 \quad (2b)$$

Scenario A

We began with a highly stylized scenario, which used only aimed fire with equal initial combatants (1000) on either side. Respective attrition rates were al = 0.25 and a2 = 0.10 per time step with b =
c= 0 for both sides. Casualty production is depicted in Figure 9. The trajectory of points in entropy space shown in figure 10 was also obtained. Comparing with the results from NTC and actual combat we see that this simulation, while generating a classical curve peaking at 0.37, is also too intense. The latter is a long standing criticism of Lanchester equations; and it appears well founded. It would also explain why deterministic Lanchester equations almost never fit real data.

Scenario C

In a scenario, arbitrarily called scenario C ,we added area fire to produce Figures 11-13. The initial force strengths were unequal being x=1500 and y =1000. For side x the value of al was constant at 0.15 while for y the values of a2 are 0.1 for t < 2.5 rising to 0.3 for 0.5 time units and then declining to 0.05 after t = 3.0. The x force is subject to area fire between t = 2.0 and t = 2.5 time units with a value of b1 = 0.001 with b2 = c = 0.

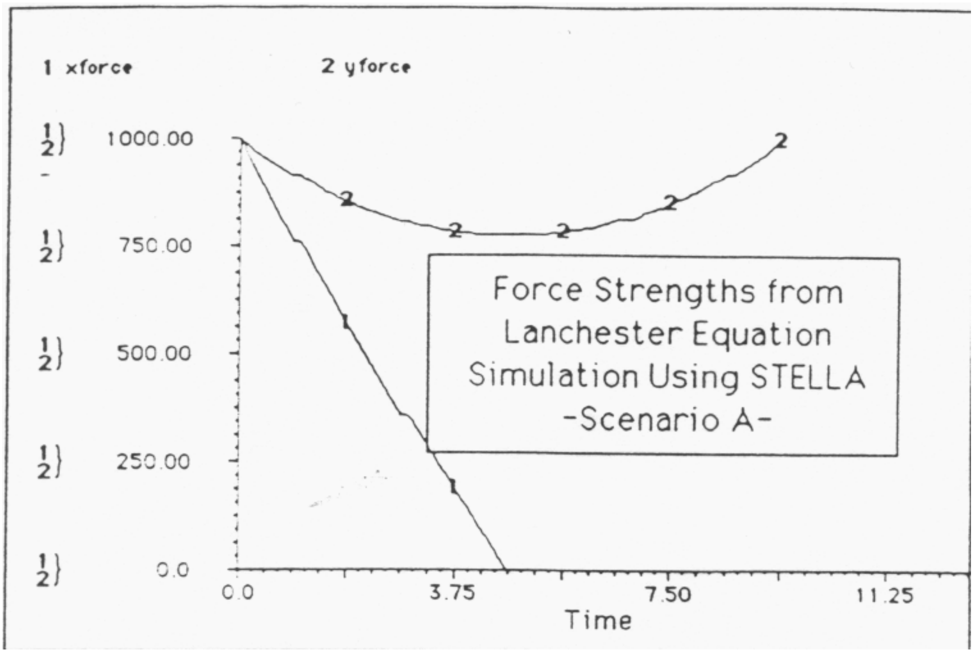

Figure 9: Casualty production with time for STELLA based simulation entitled scenario .A. Equal initial combatants but greatly unequal attrition coefficients (respectively al and a2 are 0.25 and 0.10).

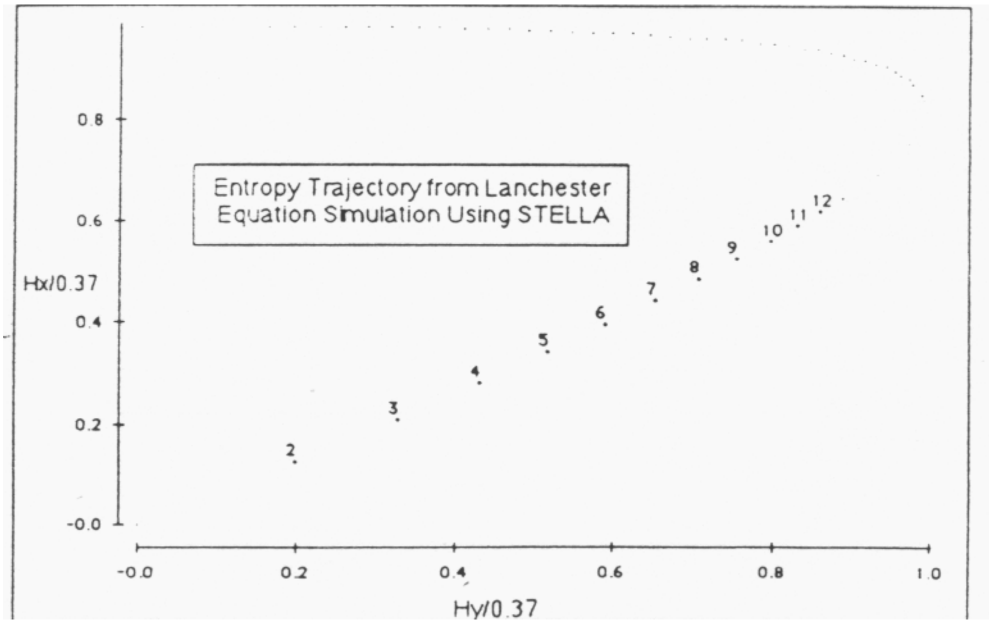

Figure 10: Entropy from a STELLA based simulation entitled scenario A shown normalized. Some initial points are numbered to show the arrow of time.

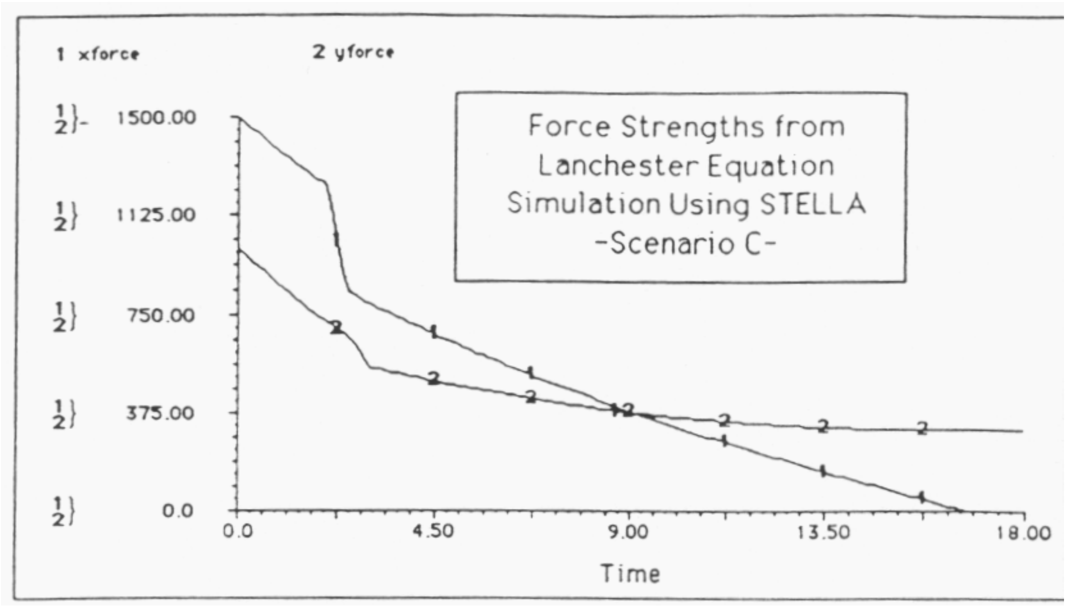

Figure 11: Casualty production with time for a STELLA based scenario entitled C. Initial force strengths favors x by 3:2. Side x inflicts constant attrition but is subject to area fire from y, and also to a time dependent attrition which increases threefold (Initial values of al and a2 are 0.15 and 0.1).

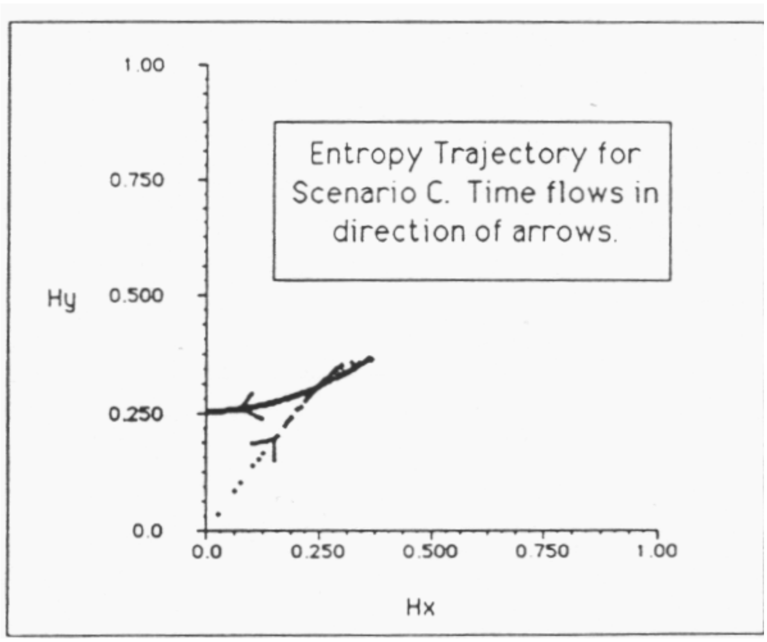

Figure 12: Plot of entropies for scenario C. Time proceeds along the direction of the arrows.

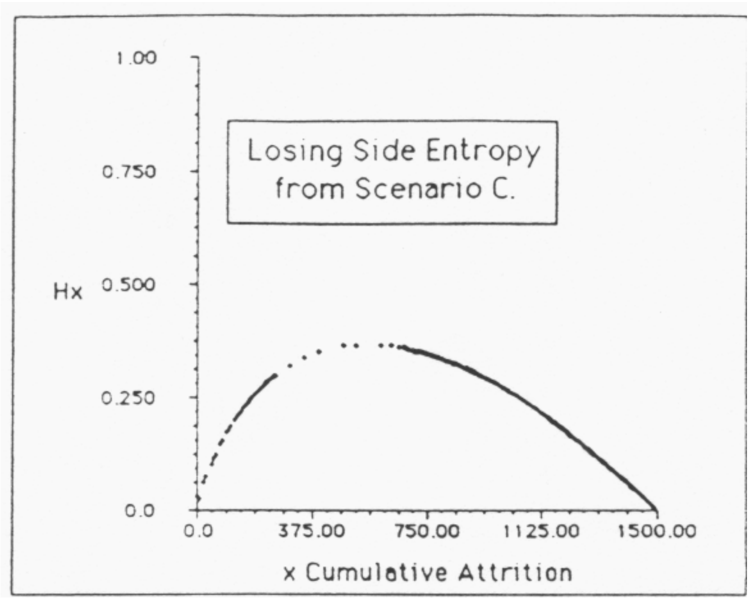

Figure 13a: Plot of the entropy versus cumulative attrition for the losing side x in scenario C.

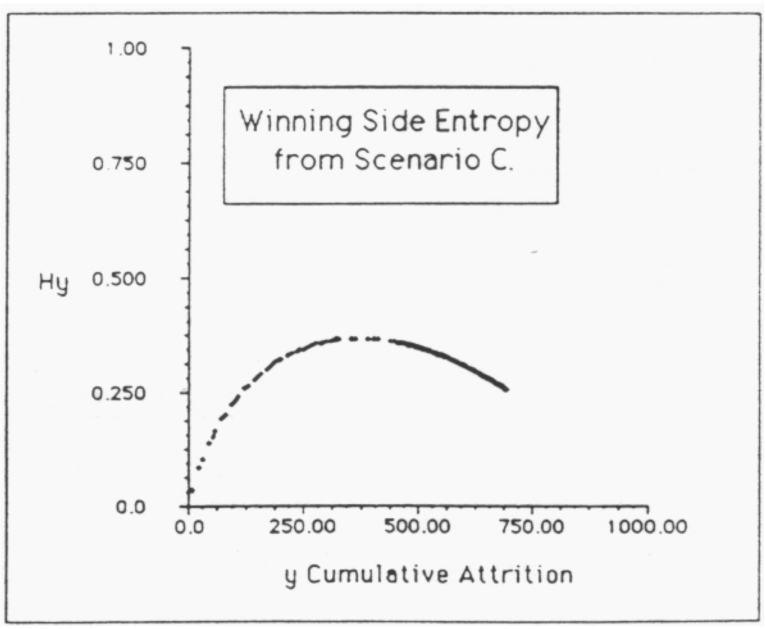

Figure 13b: Plot of the entropy versus cumulative attrition for the winning side y in scenario C.

The effect of the area fire is to reverse the tide of battle in favor of side y as seen in Figure 12. Cumulative attrition plots in Figures 13 show that even for the victor, the Lanchester Equation battle goes well beyond the peak of the p.ln(1/p) curve.

The results of the three STELLA based simulations using Lanchester equations are clearly in accord with our contention that entropy is a predictor of victory where the reversals due to additional influences is clearly seen.

Implications for C2

Because casualties and casualty rates can be measured from observables, the production of entropy based on these figures can be instrumented. Since destruction of infra-structure and materiel also contribute to the breakdown of the inherent cohesive structure required of a fighting force, it should also be possible to measure that destruction in entropic terms. Since casualty data can be notoriously uncertain, the hypotheses about progress of the battle should be cast in entropic terms as well.

The rate of entropy production is also singled out as another C2 indicator. In fact the **coupled casualty and reinforcement** rate are really what are at issue. Measurement and display of both seem close to parameterizing something which is variously described as the "tempo" of battle. The tempo is then seen to characterize, not the physical rate of advance (the usual connection), but rather the rate of structural breakdown of the fighting force.

It also seems plausible to articulate a class of entropy-based planning factors to govern the development and time line of either attack or defense. It may be that the "classical" notions of requirements for numerical superiority of the attacker at numbers like two and three to one may be related to entropy production. We would venture that these ratios are intended to permit the attacker to sustain high transient entropy production rates in order to gain, and hold, the initiative while still relying on a dynamic structure. The defender on the other hand, usually can count on more fixed infrastructure, and so would have different entropy based planning factors.

Summary

We believe we have demonstrated from the limited data available to us that the hypothesis concerning casualty production as a predictor of combat is verified from both time series analysis; and also the historical record based on final casualty counts. Moreover, casualty based entropy emerges as a major indicator for the design of future C2 systems. Entropy is shown to have deep theoretical significance for future analysis of combat as a dissipative system and for the eventual identification of attractors in combat through the use of such techniques as time series and chaotic dynamical systems analysis.

References

1. 1. Lawson, Lectures and Informal Presentations (1965-1975).

2. F. Carvalho-Rodrigues, "A Proposed Entropy Measure for Assessing Combat Degradation," J. Opl. *Res. Soc. (UK), 40* (8), pp 789-93 (1989).

3. T. Dupuy, (A singular combat historian to whose efforts, public and private, must ultimately be attributed the data on which this analysis rests]. HERO (Historical Evaluation and Research Organization), "Combat History Study Analysis Effort." (January 1986).

4. R. Helmbold, "Combat History Analysis Study Effort," CAA-TP-86-2, US Armv Concepts Analysis Agency, 8120 Woodmont Ave, Bethesda, Maryland 20814 (August 1986).

5. R. Helmbold, "Do Battles and Wars Have a Common Relationship Between Casualties and Victory?" CAA-TP--87-16, US Army Concepts Analysis Agency, 8120 Woodmont Ave, Bethesda, Maryland 20814 (November 1987).

6. R. HeImbold, private communication (1989).