Proceedings of the 1990 Winter Simulation Conference
Osman Balci, Randall P. Sadowski, Richard E. Nance (eds.)

THE OAK RIDGE SPREADSHEET BATTLE MODEL

Dean S. Hartley III

Center for Modeling, Simulation, and Gaming
Oak Ridge National Laboratory and other DOE Facilities
P. O. Box 2003, K-1001, MS 7180
Oak Ridge, TN 37831-7180

ABSTRACT

The Oak Ridge Spreadsheet Battle Model is an historically based, spreadsheet implemented battle outcome predicting model. It is a static model that predicts the outcome of one battle at a time. This outcome includes the predicted length of the battle; however, the predictions are not based on a dynamic, time stepped process, but on a straightforward set of (mostly linear) equations relating the predictions to the input data.

The construction of the model is based on statistical analyses of historical data and simulations based on these analyses. The model, in turn, can provide the basis for a dynamic simulation of warfare. The spreadsheet model assumes forces X and Y, supplied by the user, are to fight and yields the estimated results. However, one of the major factors of warfare is the selection and positioning of forces for battle and the timing of the actual engagement. A simulation containing the battle assessment model could address the majority of its processing to the strategic and logistic issues of war.

1. INTRODUCTION

All of the predictions of the Oak Ridge model are based on a database containing over 600 historical battles and the analysis of possible causative factors, both numerical and judgmental [Dupuy, et al.]. Portions of the model are not only historically based, but have been validated against separate sets of data and may be considered very reliable. Portions of the model are based on sparse data and are relatively unreliable. The accuracy of this model is known to be no more than two significant figures for the expected value. When the known random effects are included, the amount of variability in battle results becomes clear.

The simplicity of the battle assessment model is based on its orientation to available data, rather than theoretical constructs. The data on warfare that are collected, analyzed, and available for model construction are limited in quantity and in number of attributes for each battle. This model was built from the largest known database that contains a large number of attributes. The putative causal attributes and the outcome attributes of the database define the inputs and outputs of the model.

The simplicity of the model means that it is both fast running and robust. By not being excessively specific in its predictions, it is more often correct. The philosophy is, "better to be approximately right than exactly wrong." The value of the results is increased by the presentation of confidence intervals for each result.

2. PREPARING THE INPUT

The input to the battle model requires four screens of input and two combination input/output screens. The actual number of input variables is fairly small; however, space is taken to identify each input variable and cue the user for the expected value ranges or type of data. Space is also reserved to display error indicators.

2.1 Battle Identification

Figure 1 shows the format of cues and input fields much as they will appear in the spreadsheet. The time and date are displayed in the upper right corner of the screen. The battle identification is a free format, alphanumeric field which, together with the date and time, allows for scenario definition.

```
                          OAK RIDGE BATTLE PREDICTOR
INPUT DATA                BATTLE IDENTIFICATION            03/09/90 RUN DATE
                          Russia-USA/Scenario 4         04:11:37 PM RUN TIME

ATTACKER IDENTITY                                   DEFENDER IDENTITY
  Arabs                        0       c              Arabs                  0
  Austria                      0       h              Austria                0
  England                      0       o              England                0
  European                     0       i              European               0
  France                       0       c              France                 0
  Germany                      0       e              Germany                0
  Israel                       0                      Israel                 0
  Italy                        0       =              Italy                  0
  Japan                        0                      Japan                  0
  Russia                       1       1              Russia                 0
  USA                          0                      USA                    1
  Other                        0                      Other                  0
                             (ERR)                                         (ERR)
```

Figure 1. First Input Screen

2.2 Combatants' Identities

Figure 1 also includes the input section for identifying the combatants. The selections and the groupings (Arabs, European, Other) were delineated by the historical data used to create this model. Certain countries had sufficient numbers of battles and differed significantly from other countries in battle results to warrant a separate listing. Other countries either had insufficient battle data or were insufficiently different from others for a separate listing and were grouped with similar countries. As examples, all of the Arab countries are grouped together; countries such as Greece and Spain are grouped as (other) European countries; and most non-European countries are grouped as "Other".

The selection illustrated shows Russia as the attacker (choice = "1") and the United States as the defender (choice = "1"). The "0" values for the other choices remove them from the calculations. If there are multiple countries on one or both sides, simply enter the fraction of the total force that each country's forces represent. The only consideration is that the total of the fractions for each of the sides must be "1.0." The "(ERR)" entry in the illustration represents the position of the error test indicator. If the sum = 1.0, then that area is blank; otherwise, the area contains the letters "ERR," indicating an error.

2.3 Combatants' Forces

Figure 2 includes the input variables for the attacker and the defender forces. These are the number of personnel, the number of armored attack vehicles (tanks and armored guns, but not armored personnel carriers), the number of artillery pieces (cannon, artillery mortars, and multiple rocket launchers), and the number of air sorties (single aircraft missions flown in the engagement area against enemy targets by fighters, fighter-bombers, and bombers) to be performed each day.

```
ATTACKER FORCES                                     DEFENDER FORCES
  Personnel                40000                      Personnel                16000
  Armor                     4000                      Armor                     1000
  Artillery                 2500                      Artillery                  500
  Air Sorties/Day            400                      Air Sorties/Day            400
                           (ERR)                                               (ERR)

ATTACKER HUMAN FACTORS                              DEFENDER HUMAN FACTORS
  Technology                   2                      Technology                   2
  Leadership                   0                      Leadership                   1
  Morale                       1                      Morale                       1
  Training                     1  -2 <---> 2          Training                     1
  Initiative                   2                      Initiative                   0
  Intelligence                 2                      Intelligence                 1
  Momentum                     1                      Momentum                     0
  Logistics                    1                      Logistics                    0
  Combat Effectvness           1                      Combat Effectvness           2
                           (ERR)                                               (ERR)
```

Figure 2. Second Input Screen

2.4 Combatants' Human Factor Values

Figure 2 also contains the human factors input data. Each factor may range from "-2" to "+2." Fractional values are permitted, but are probably only justified when a series of battles is being examined, with initial integral values deriving from human judgment and successive values being judged as changes from the initial values depending on the results of the previous battles.

The values should be relative to the time frame of the battle. For instance, the presence of cannon might represent a Technology value of +2 at a time when cannon were relatively unusual, and a value of 0 when cannon were common battle systems. A force with cannon technology on today's battlefield would rate a -2 technology score. Leadership should be judged from the commander of the given battle down. A value of +2 indicates the presence of leadership on the order of Napoleon. A +1 Leadership value indicates competent, well-trained leaders with a tradition of initiative. The value of -1 should be used for normally ill-trained or "green" leaders and -2 reserved for exceptionally incompetent leadership. Morale and Training represent the state of the situation, including what has happened previously in the war. Thus troops with previous battle losses might have lower morale but better training (which includes experience in battle) than the same troops had for the first battle. Initiative refers to proactive actions as opposed to reactive actions. The attacker usually has the initiative; however, in a substantial number of situations, especially at the battle level, the defender may have the initiative, for instance by choosing the time and place of the battle. Intelligence refers to the military capability of gaining and using information concerning the organization, dispositions, intentions, and activities of the forces of the opponent. Momentum is correlated with initiative, but not identical. Momentum is the result of immediate past successes in movement against the enemy. These successes may make taking the initiative easier; however, they do not guarantee that the initiative will be taken. Logistics refers to the ability of the forces to obtain and perform resupply in the appropriate time frame. The logistics ability of a force may be good for the first battle of a series, but decline for successive battles. This change may become significant because one side's logistics ability declines faster than the other side's logistic ability.

Combat effectiveness is a complex judgment factor that is partially redundant when considered together with other factors above; however, it correlates better with some results than does the combination of those factors and is included for that reason. It is an overall judgment about the leadership, training, experience, morale, and logistics of a force and the interaction of these factors. This same effect may be seen in sports teams in which the individual numbers concerning the team don't add up to the overall effectiveness of the team and in teams with exceptionally talented individuals, but without the cohesion of a true team.

2.5 Operational Data

The operational data input variables include the attackers' plan, the defender's scheme of defense, the defender's posture and frontal width, and the presence or absence of air superiority by one or the other side. These variables are found in Figure 3.

Analysis of the historical data yielded the six statistically different attack plans shown in Figure 3. The significance of the difference actually varies depending on the result being modeled. For some results the attack plan makes no discernable difference at all and is not used. For other results it might be the case that only four of the plans are really different in effect. This is modeled by having relatively close weights for the similar plans. For simple battles, choose one plan and place a "1" in its position. Larger battles with combination assaults may be represented with fractional values (with total values adding to 1.0). The defense scheme was analyzed in the same manner as the attack plan and seven different schemes were found to be significant. The desired defense scheme or schemes should be chosen in the same way as described for the attack plan.

```
                    ------ OPERATIONAL DATA ------

ATTACK PLAN                        c      DEFENSE SCHEME
  Single envelopment        0      h        Strict Defense          0
  S. env + Frontal          0      o        Defense + Offense       1
  Double envelopment        0      i        D/O, S. env             0
  D. env + Frontal          0      c        D/O, D. env             0
  Frontal attack            1      e        D/O, Frontal            0
  Riv Crossing              0      =        D/O, Frntl + S env      0
                          (ERR)    1        D/O, Frntl + D env      0
                                                                  (ERR)

DEFENDER'S WIDTH: km      30.0             DEFENSE POSTURE         6.0
                          (ERR)              Fortified = 8        (ERR)
AIR SUPERIORITY             -1               Prepared  = 6
  Defender = -1           (ERR)              Hasty     = 4
  None     =  0                              Delay     = 2
  Attacker =  1                              Withdraw  = 0
```

Figure 3. Third Input Screen

The defender's frontal width represents a measure (in kilometers) of the contact zone between the attacker and the defender. (Other measures based on the width are calculated internally.) The defense posture is represented on a scale from 0 to 8. The value "0" is used when the defender is withdrawing (attempting to disengage from the fighting). The value "2" is used when the defender is delaying (moving to the rear, but attempting to fight small skirmishes to delay the attacker's advance). The value "4" is used when the defender has chosen to stand and fight, but has not had time to prepare a defensive position (e.g., foxhole only). The value "6" is used when the defender has had time to prepare his position (log bunkers, etc.). The value "8" is used when the defender has built concrete bunkers or the equivalent. Intermediate and non-integral values may be used either to represent intermediate values or averages across varying amounts of preparation among the defender's forces. The final input variable in Figure 3 is the air superiority variable. Three values are defined. Either the defender has air superiority over the battlefield (-1); or no one has air superiority (0); or the attacker has air superiority (+1).

2.6 Environmental Data

The environmental data (climate and season, weather, temperature, and terrain visibility) are found on Figure 4. The historical data support differentiation among the five combinations of climate and season shown. Generally a battle will fall within only one of these categories and should have a value of "1" for only one of the categories. However, if after running the model, it is determined that the battle length is sufficient to span two seasons, that span can be indicated with fractional values. The weather variable ranges from 1 to 4 and may have fractional values to indicate an average value. The temperature variable ranges from -1 for cold (freezing temperatures) weather to +1 for hot (tropic temperatures) weather. Intermediate values are allowed. Terrain visibility is a judgmental variable based on the visibility allowed by the type of terrain occupied by the battlefield. Bare terrain is indicated by a "1." Mixed bare and wooded terrain is indicated by a "2." Heavily wooded terrain is indicated by a "3." Urban terrain is indicated by a "4." Intermediate or averages for the entire battlefield may be given by non-integral values.

2.7 Battle Date and Confidence Interval

The battle starting date and the desired confidence interval also are found in Figure 4. This battle model computes expected values for the results of the input values and also a low and high value based on the expected values and a requested confidence interval size, based on standard deviations. The low and high values are roughly the number of standard deviations input below and above the expected values.

2.8 Output Overrides

The output values will be discussed in Section 3; however, there are optional input values contained in the output section. The reason for these optional inputs is that some of the output values are used in calculating other output values and, like all the output values, the results are only guesses (based on historical data, but still guesses) and may be wrong. The optional inputs allow the user to override the output value and view the results on dependent variables. The optional overrides include duration, surprise, advance rate, and casualties.

```
                    ----- ENVIRONMENTAL DATA -----
                                                                    (ERR)
CLIMATE & SEASON          (ERR)           WEATHER                     2.5
  Temperate - Spring          1             Sunny=1, Overcast=2
  Temperate - Summer          0             Light Rain=3, Heavy Rain=4
  Temperate - Fall            0                                     (ERR)
  Temperate - Winter          0           TEMPERATURE                -0.5
  Non Temperate               0             Cold=-1, Mild=0, Hot=1

TERRAIN VISIBILITY          2.5
  Bare=1, Mixed=2,        (ERR)
  Wooded=3, Urban=4

          ----- BATTLE DATE AND CONFIDENCE INTERVAL -----

STARTDATE                          20-Apr-90
  Day                        20
  Month                       4           NUMBER OF STD DEVS          1.0
  Year                       90                                     (ERR)
```

Figure 4. Fourth Input Screen

The first output screen, Figure 5, contains the duration override. A value of "0.0" indicates that no override is intended. Any value greater than 0 is used to replace the duration values in all subsequent calculations. The value is taken to be an exact value and thus replaces the low and high duration values, as well as the expected duration value.

The surprise override is also contained in Figure 5. A value of "0.0" indicates that no override is intended. Because the value of surprise ranges from -3 (complete surprise by the defender) to +3 (complete surprise by the attacker), there is a problem if the user wishes to override the surprise with a 0 surprise value. The special value of "0.001" is chosen to represent this situation. The mathematical effect of this value is the same as a zero surprise value and allows the program to use "0.0" to indicate no override.

```
PREDICTIONS             BATTLE IDENTIFICATION              03/09/90 RUN DATE
                        Russia-USA/Scenario 4           04:11:39 PM RUN TIME

DURATION - LOW                1.8 Days
         - EXP                3.1
         - HI                 5.2
  DURATION override           0.0
                            (ERR)
SURPRISE - LOW                1.4  Att. Complete    =3  Def. Complete    =-3
         - EXP                2.2  Att. Substantial=2  Def. Substantial=-2
         - HI                 3.0  Att. Minor       =1  Def. Minor       =-1
  SURPRISE override         0.001  None             =0.001 (override)
                            (ERR)

ADVANCE RATE - LOW           -0.9 km/day      DISTANCE - LOW            -2.9
             - EXP            2.8                      - EXP             8.5
             - HI             6.5                      - HI             19.9
  ADVNC RAT override        0.000
                            (ERR)
```

Figure 5. First Output Screen

The advance rate override is also contained in Figure 5. A value of "0.0" indicates that no override is intended. Because the value of the advance rate may be virtually any number (negative for advance by the defender and positive for advance by the attacker), there is a problem if the user wishes to override the advance rate with a 0 kilometers/day advance rate value. The special value of "0.001" is chosen to represent this situation. The mathematical effect of this value is the same as a zero advance rate value and allows the program to use "0.0" to indicate no override.

The casualties override values are in the second output screen, Figure 6. The attacker casualties and the defender casualties may be overridden; however, if one is overridden, both must be overridden. (If, for some reason, the user wishes to insist that one side has zero casualties, a small positive value, such as "1," must be used.) A value of "0.0" indicates that no override is intended.

```
                          --- CASUALTIES AND LOSSES ---
ATT CASUALTIES - LOW      843   k/w/mia  DEF CASUALTIES - LOW      177
               - EXP     1972                           - EXP      450
               - HI      4614                           - HI      1140
  ATT CAS override          0              DEF CAS override          0
                        (ERR)    (ERR)                           (ERR)
ATT ARMOR LOST - LOW      713            DEF ARMOR LOST - LOW      174
               - EXP      775                           - EXP      210
               - HI       837                           - HI       246

ATT ARTIL LOST - LOW        6            DEF ARTIL LOST - LOW        0
               - EXP       22                           - EXP        0
               - HI        38                           - HI       127

ATT AIR LOSSES - LOW        0            DEF AIR LOSSES - LOW        0
               - EXP        0                           - EXP        0
               - HI         2                           - HI         1
```

Figure 6. Second Output Screen

3. UNDERSTANDING THE OUTPUT

The output of the battle predictor takes two forms. The first is that of the output screens of the spreadsheet. Two of these screens were shown in Figures 5 and 6 in the discussion of output overrides in the input section. The third output screen is shown below in Figure 7. The second output form is the printed output. The printed output is divided into two pages. The first page combines the four input screens and the second page combines the three output screens. Because the two output forms have the same format and data content, the detailed descriptions in the following paragraphs address the screen output.

The first third of the output is shown in Figure 5. The battle identification and date/time data are repeated to connect the results to the input. The Duration predictions consist of the expected value and the low and high estimates, given in days. In this example, the estimates range from 1.8 days to 5.2 days with an expected value of 3.1 days. The Duration override value of 0.0 indicates that the predicted duration values were not overridden.

The Surprise predictions range from a low of 1.4 (minor attacker surprise) to a high of 3.0 (complete surprise by the attacker) with an expected value of 2.2 (substantial attacker surprise). For this particular scenario, it was decided by the user that the battle being analyzed was part of a larger situation and that attacker surprise was essentially impossible because of major troop movements. Hence, the surprise override value is shown as set to 0.001, forcing a "no surprise" condition on subsequent calculations.

```
                         --- PREDICTING VICTORY ---
ADVANTAGE - LOW          -4.8
          - EXP          -2.5
          - HI           -0.2

          VICTOR - LOW                      >> DEFENDER <<
                 - EXP                      >> DEFENDER <<
                 - HI                       >> DRAW     <<

     The predictions should be used in conjunction with the graph that
places this battle within the context of the historical data used to
produce the predictions.  The three data points representing "This
Battle" show the low-, expected-, and high-predictions.  When the
points fall within the "Core group" ellipse, the results may be given
a high confidence level.  If the points fall outside the "Outer
limit" ellipse, "This Battle" is too far from the underlying data for
good predictions.
```

Figure 7. Third Output Screen

The Advance Rate predictions are based on the input values and the predicted (or override) surprise values. The predictions range from a low of -0.9 kilometers/day (the defender advances 0.9 km/day) to a high of 6.5 km/day (the attacker advances), with an expected value of 2.8 km/day. The advance rate override of 0.000 indicates that the advance rate predictions have not been overridden. The expected Distance prediction is the product of the expected advance rate and the expected duration. The low and high predictions require slightly more complicated calculations.

The middle third of the output is shown in Figure 6 and contains the casualties and losses predictions. The casualty predictions are based on total casualties, including killed, wounded, missing in action, and captured. These predictions are based on the input data, the predicted (or overridden) duration, and the predicted (or overridden) advance rate. The attacker and defender casualty override values are both 0, indicating the casualty predictions have not been overridden.

The predictions for casualties and victory are the most "accurate" predictions made by the model in the sense that the combined expected value predictions and range predictions for these elements have theoretical components that have been validated against multiple datasets. The other predictions are the best fits available for a large set of data; however, they have not been validated against a second set of data. The predictions for attacker and defender losses of armor, artillery, and aircraft follow the same pattern of expected, low, and high predictions. The aircraft loss predictions are the least reliable of all the predictions because the historical data were sparsest in this area.

The last third of the output is shown in Figure 7. Its principal contents are the victory predictions. The Advantage predictions are used to predict the victor. The values for advantage range from -8 for an absolute advantage for the defender to +8 for an absolute attacker advantage. In the illustrated scenario, the predicted advantage ranges from a low of -4.8 to a high of -0.2, with an expected value of -2.5. Advantage is converted to a victory prediction by selecting the defender as the predicted victor when the advantage value is less than -1, the attacker as the victor when advantage > 1, and calling the battle a draw otherwise. In this case, the expected winner is the defender, with a possibility (at the approximate 1.0 standard deviation level) of a draw.

The final part of the output is a discussion of a method for determining whether the input scenario falls reasonably well within the historical data that were used to create the battle predictor model. Because interpolation is a more reliable prediction method than extrapolation, predictions based on scenarios that are within the historical data space are more likely to be nearly correct than those falling well outside the historical data space. This data space, called Helmbold space, is illustrated in Figure 8. The solid ellipse contains the core group of the historical data. The ellipse of boxes contains the majority of the remainder of the historical data. If all of the historical data were plotted in the graph using the same scale symbols, the interior of the solid ellipse would appear completely covered and the space between the solid ellipse and the boxed ellipse would show a decreasing density from the inside out. Only a few data points would appear outside the boxed ellipse.

The range of the predicted results from the current model scenario is shown by the three diamonds (representing the low, expected, and high predictions). In this case the predictions are relatively close together and fall on the edge of the solid ellipse. The results of this test are that the predictions may be given a high level of confidence.

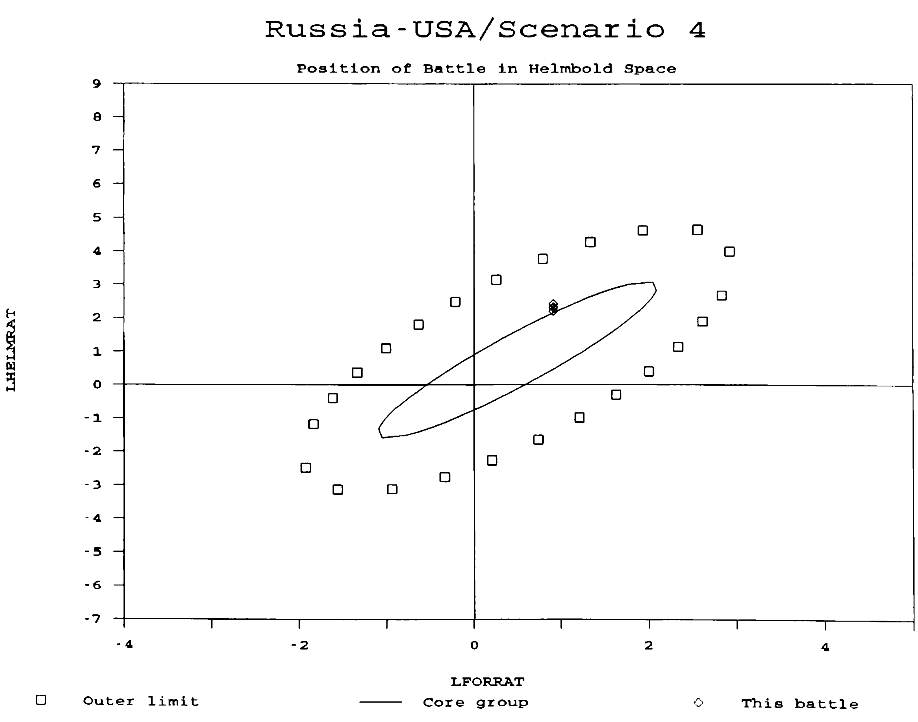

Figure 8. Output Results In Helmbold Space

4. CONCLUSION

The Oak Ridge Spreadsheet Battle Model is a simple and fast model of combat battles. It is a robust model, having a large domain of reliable predictions, predictions based on a large set of historical data. Its output consists of expected values for the duration of the battle, achievement of surprise, the advance rate and total distance advanced, casualties, major combat systems losses, and victory. It also produces estimates of the range of the predictions, based on an input number of standard deviations. The model is useful for simple "what-if" analyses and may also provide a simple, historically based battle model for a larger combat simulation that addresses the more complex issues of deciding what forces will engage in battles and what they do afterward.

A complete description of the derivation process for this model is found in Hartley [1990b]. The research that validated the basic attrition model is contained in a series of papers [Hartley and Kruse 1989; Hartley 1989a, b, c, 1990a]. A more detailed description of the implementation of this model as a spreadsheet, complete with formulas suitable for spreadsheets, is found in Hartley [1990c]. The model may be made available to interested parties.

ACKNOWLEDGEMENT

Research sponsored by the Office of the Secretary of Defense/Net Assessment under Interagency Agreement 1612-1612-A1 between the Department of Energy and The Joint Staff/J-8 and performed under Martin Marietta Energy Systems, Inc., Contract No. DE-AC05-84OR21400 with the U.S. Department of Energy.

REFERENCES

Dupuy, T., et al. (1988 and 1989), *HERO's Land Warfare Data Base (LWDB)*, Data Memory Systems, Inc, Fairfax, VA.

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

Hartley, D. S. III (1989a), "The Constraint Model of Attrition," K/DSRD-114, Martin Marietta Energy Systems, Oak Ridge, TN.

Hartley, D. S. III (1989b), "The Constraint Model of Attrition," In *Proceedings of the 1989 Winter Simulation Conference*, E.A. MacNair, K.J. Musselman, and P. Heidelberger, Eds., Piscataway, NJ, 989-996.

Hartley, D. S. III (1989c), "Historical Validation of an Attrition Model," K/DSRD-115, Martin Marietta Energy Systems, Oak Ridge, TN.

Hartley, D. S. III (1990a), "Confirming the Lanchestrian Linear-Logarithmic Model of Attrition," K/DSRD-263, Martin Marietta Energy Systems, Oak Ridge, TN.

Hartley, D.S. III (1990b), "Predicting Combat Effects," K/DSRD-412, Martin Marietta Energy Systems, Oak Ridge, TN.

Hartley, D.S. III (1990c), "Users' Guide to the Oak Ridge Spreadsheet Battle Model," K/DSRD-413, Martin Marietta Energy Systems, Oak Ridge, TN.