Journal of the
Military Operations Research
Society of Korea, Vol. 5, No. 2
December, 1979

Aggregation of Measures of Effectiveness with Constant Sum Scaling Method and Multiple Regression

Kim, Hyung Bae *

ABSTRACT

This method explores a method of aggregating the measures of effectiveness of a weapon system from its characteristics. With this method, the constant sum method and multiple regression are used to develop a functional relationship between system effectiveness and system characteristics. As an example, a study of a tank weapon system was conducted with data from the U.S. Army Armor School. It was concluded that the aggregation method is feasible, and that for the tank system studied, the reciprocals of system characteristics give a good estimating equation for measuring tank system effectiveness.

1. Introdution

We often find measurement problems in O. R. that are difficult in that widely used concepts have not been made operational. How to measure the effectiveness of a weapon system is one of the most important tasks in military affairs. What is needed is some method to give answers to such questions as " How much better is a M60A1 than a T-62 ?"
We want to know much better one weapon system is than another among similar systems.

In this paper we will propose and demonstrate a way of structuring such a relationship using the effectiveness of weapon systems values which originally came from military experts' judgements. Once we have found such a function, we would not necessarily require experts' judgements again since one can use the function to calculate the effectiveness of a proposed weapon system from its characteristics.

* 国防部 PPBS室

2. Concept of the problem

The problem we are interested in is to calculate the overall effectiveness of a weapon system from its characteristics. The approach here will be to demonstrate how to estimate the effectiveness of a weapon system using a scaling method [1].

Every weapon system has its own characteristics and if we have a value for the overall effectiveness of that system, we would be able to obtain or fit a relationship expressing :

Overall effectiveness of a weapon system $= f(x_1, x_2, \cdots, x_m)$ ………………… (1)

where

$x_1, x_2, \cdots\cdots, x_m$ are system characteristics.

The purpose of this study is to show a procedure for obtaining values for overall effectiveness, and a way of determining the function f. Since we will be fitting functions to data, the more instances of the system we use, the better the functional relationship we can find.

There are n instances of the system and thus n effectiveness values which have to be obtained. The system has m characteristics which we presume to relate to effectiveness, and assume $m < n$. If we express this figure in a mathmatical equation we would write ;

$$Y_i = f(X_{i1}, X_{i2}, \cdots\cdots, X_{im}) \quad \cdots\cdots (2)$$

describing overall effectiveness of a system as a function of its characteristics, and we propose to show how to obtain values for the Yi, and how to find a good fitting function f.

3. Study procedure

The genersal study procedure is composed of 3 steps.

a) Design of the study

b) Computation of the effectiveness of each instance of the system

c) Determination of the functional relationship

A. Design of the study

a) Selection of the competing systems

What kind of systems should we choose ? It depends on what we are going to do with those systems, and since we want to decide which tank system is good for battle, we should choose Main Battle Tanks (MBTs) [3] For this study six Main Battle Tanks were chosen :

M 48 A 5 (Korea)

	M 60 A1	(U. S. A.),
	AMX - 30	(France),
	LEOPARD 1	(W. Germany),
	T - 62	(U. S. S. R.),
and	CENTURION (MK. 13)	(Great Britain.).

As mentioned before, the more tanks used in the study, the better is the relationship that may be obtained. The reason for using only six MBTs here relates to the workload that is placed on the judges.

In the constant sum method which we are going to use, each judge is asked to consider each possible pair of instances and split 100 points between the two instances in each pair. Thus for n instances to be scaled, $n(n-1)/2$ pairs must be considered by each judge. Therefore we have to consider the number of instances to be scaled because the greater the number of instances, the greater the workload asked of each judge and the less careful he might be in his rating and the smaller proportion of questionnaires we would expect to be returned.

Since "the more judges the better" is particularly an axiom in scaling, tradeoffs may have to be made between the effort that will be required of a judge and the amount of confidence one wishes to have in the resulting scale. Our selection of six tanks (MBTs.) requires $\frac{6(6-1)}{2} = 15$ pairs of tanks to be considered, and this mumber is thought reasonable for a judge to handle in a short period of time.

b) Selection of the Major Characteristics

There are of course many characteristics which affect the effectiveness of a system. Some of the characteristics have very similar valuses among various MBTs, and these characteristics should not make any substantial difference in the comparison of effectiveness among the competing system.

We have chosen 4 tank characteristics as follows :

1. Speed
2. Silhouette
3. Hp/ ton
4. Armor

These characteristics are not necessarily the most important ones. For example, fire power is a very important consideration, but obtaining useful numerical data on fire power is very difficult due to a lack of a standard measurement criterion.

Therfore we have chosen the above characteristics as generally accepted important factors, which should serve well in our demonstration of a method for assessing system effectiveness.

c) Collection of Data

After selecting the system instances and the major characteristics for the study, we have to collect data for each characteristic.
Table I shows the basic data of the six MBTs, which we will use.

Table I. CHARACTERISTICS OF SIX MAIN BATTLE TANKS

CHARACTER-ISTICS / SYSTEMS	ARMOR (mm on nose)	SILHOUETTE (height in m)	SPEED (Km/ hr)	Hp/ ton
LEOPARD 1	70	2.64	65	20.7
M 60 A 1	110	3.26	48.3	15.3
T- 62	100	2.4	50	19.2
M 48 A 5	110	3.09	48.2	15.2
CENTURION (Mk. 13)	118	3.01	34.6	12.5
AMX- 30	48	2.85	65	19.4

d) Selection of the Method

There are many scaling methods we could use for our study, among those scaling methods the constant sum method will give a ratio scale which is easy to use. We can convert the information from judges about system effectiveness into a ratio scale. Therefore the constant sum method was chosen for the illustrative study of tank effectiveness.

e) Preparation of the questionnaire

The questionnair should be prepared very carefully with a clear explanation of how to fill it out, together with information about the systems which the judges can use to assist them in their ratings. Since we are going to use the constant sum method to compute the effectiveness the judges will be asked to make ratio scale judgements by splitting 100 points between the two instances represented by each possible pair of n tanks. For our study we have six instances, and therefore there are 15 pairs to be presented to the judges.

f) Selection of the Judges

There appears to be no rule or standard for designating individual as " experts ". It depends on common sense or military judgements.

We believe that the armor officers in the U. S. Army Armor school may be considered experts about tanks.

B. Computation of system effectiveness

The constant sum scaling method of computing the overall effectiveness values for each system instance, using the information received from the judges, will be explained.

a) Calculation of the effectiveness of each weapon system using the constant sum scaling method.

The constant sum scaling method is designed to scale a property with a natural origin upon which the judges agree.

Let a_{ij} be the number of points out of 100 which a judge awards to instance j when it is compared to instance i. If we arrange a judge's responses in an array a_{ij}, there will be one array for each judge in which values on the diagonal would be set at 50 because comparing something with itself should be 50 : 50.

We could average these arrays over the N judges to produce an array $\overline{a_{ij}}$ where

$$\overline{a_{ij}} = \frac{\sum_{\text{over all judges}} a_{ij}}{N} \quad \cdots\cdots (3)$$

and the values on the diagonal remain set at 50.

From Table Ⅱ one may see, for example, that the average of the judges' 100-point split in overall tank effectiveness between the T-62 and the AMX-30 was 55.9 for the T-62 and 44.1 for the AMX-30.

Table Ⅱ. a_{ij} Array Computed from Judges' Responses

	M48 A5	M60 A1	AMX-30	LEOPARD1	T-62	CENTURION (MK. 13)
M 48 A 5	50	50.94	45.1	64.22	54.7	45.4
M 60 AI	43.06	50	39.14	54.44	45.04	39.36
AMX-30	54.9	60.86	50	63.86	55.9	49.7
LEOPARD 1	35.78	45.56	36.14	50	41.46	37.54
T-62	45.3	54.96	44.1	58.54	50	44
CENTURION (MK. 13)	54.6	60.64	50.3	62.46	56	50

The next step is to construct a new W_{ij} array where the entries are the ratios of the instance values across the diagonal, or

$$W_{ij} = \frac{\overline{a_{ij}}}{\overline{a_{ji}}} \quad \cdots\cdots (4)$$

In this array, it is immediately apparant that the entry in the ith row and jth column is the reciprocal of that jth row, and ith column, i.e.,

$$W_{ij} = \frac{1}{W_{ji}} \quad \cdots\cdots (5)$$

Table Ⅲ. W_{ij} Array

	M48 A5	M60 A1	AMK-30	LEOPARD 1	T-62	CENTURION
M 48 A 5	1	1.322	0.821	1.785	1.207	0.831
M 69 A 1	.756	1	.643	1.195	.819	.649
AMX - 30	1.217	1.555	1	1.767	1.267	.988
LEOPARD 1	.557	.837	.566	1	.708	.601
T - 62	.838	1.22	.789	1.412	1	.786
CENTURION	1.203	1.542	1.012	1.664	1.273	1

Since W_{ij} is the ratio of the average points awarded to j (when compared to i) to the average points awarded to i (when compared to j), then in general, if S_i and S_j are the scale values we seek, W_{ij} is an estimate of the ratio S_j/S_i. Thus in terms of the Table Ⅲ data, for example, judges have indicated that they feel that the M60 A 1 tank is 1.322 times more effective than the M48A5 tank. The solution is overdetermined, however, since there are far more W_{ij} ratios)(15) than there are scale values to be estimated (6).

We propose to handle this multiple estimate problem by a least squares approach over the estimate.

$$W_{ij} = \frac{S_j}{S_i} \qquad \cdots\cdots (6)$$

by taking the log to both sides of (6), we have

$$\log W_{ij} - (\log S_j - \log S_i) = 0 \qquad \cdots\cdots (7)$$

For the least squares approach we wish to obtain values of S_j and S_i which will make

$$\log W_{ij} - (\log S_j - \log S_i)$$

close to zero over all instance pairs i, j. Thus we want to find values S_1, S_2 ...., Sn such that

$$Q = \sum_{i=1}^{n} \sum_{j=1}^{n} \left[\log W_{ij} - (\log S_j - \log S_i) \right]^2 \qquad \cdots\cdots (8)$$ is minimized.

Algebraically expanding (8), and taking the n^{th} partial derivative of Q with respect to S_j and set them equal to zero, we have

$$\log S_j = \frac{\sum_{i=1}^{n} \log W_{ij}}{n} + \frac{\sum_{i=1}^{n} \log S_i}{n}, \quad j = 1, 2, \cdots, n \qquad \cdots\cdots (9)$$

since the choice of a unit for scale is arbitrary, we will chose one such that the average of the logs of the scale value is zero, or

$$\frac{\sum_{i=1}^{n} \log S_i}{n} = 0.$$

Thus equation (9) becomes

$$S_j = \prod_{i=1}^{n} (W_{ij})^{1/n} , j = 1 , 2 , \cdots , n \quad \cdots\cdots\cdots\cdots (11)$$

Therefore the scale value of instance j., S_j, derived from the least squares approach can be interpretated simply as a Geometric mean of the jth column of the W_{ij} array. Using Table Ⅲ (W_{ij}) array, we obtained S_j values as follow:

$S_1 = 0.89$ $\quad S_4 = 1.44$
$S_2 = 1.22$ $\quad S_5 = 1.02$
$S_3 = 0.78$ $\quad S_6 = 0.79$

b) System Effectiveness Values

At last we have ratio scale values for the relative effectiveness of each tank as shown in Table Ⅳ.

Table Ⅳ. EFFECTIVENESS OF SIX MAIN BATTLE TANKS

TANK	EFFECTIVENESS
M 48 A 5	0.89
M 60 A 1	1.22
AMX - 30	0.78
LEOPARD 1	1.44
T- 62	1.02
CENTURION(MK. 13)	0.79

The effectiveness values calculated above are relative effectiveness values, and have no physical values.

An advantage of using a scaling method which requires that the judges provide ratio scale information is that the output is also a ratio scale. Thus we can say the Leopard 1 is more effective (or better) than the AMX- 30 by

$$\frac{1.44 - 0.78}{1.44} \times 100 = 46\%$$

The effectiveness value by itself is meaningless, but because this is a ratio scale, we can compare the two systems directly by ratios, and can say how much better one is than the other.

This scaling approach provides an effective way of computing the overall effect-iveness of weapon systems.

However, this use of a scaling method alone requires that we have to send que-stionnaires everytime we want to calculate the effectiveness of, say, a new or different Main Battle Tank. This is because the effectiveness we computed is based on the information given by the judges, but not directly on the system characteristics.

We are going to find a function which relates the characteristics to the effectiveness values. With the data shown in Table I (characteristics), tank characteristics can be thought of as explanatory variables and the effectiveness can be thought as a dependent variable for the multiple regression [2] analysis which will now be discussed.

C. Determination of the functional relationship

a) Searching for a functional relationship between effectiveness and system characteristics using multiple regression.

Since there are six dependent variables (6 MBTs) and four explanatory variables (4 characteristics), computational work [4] is simplified with a computer program for linear and non- linear multiple regression. There exists a very useful APL computer program for multiple regression named " REGRESS ", which does a multiple regression analysis relating the dependent variable Y to a set of explanatory variables X.

Here Y is a vector of size n and the right hand argument X is an $n \times m$ matrix consisting of n observations on each of m variables, corresponding to the tank characteristics data in Table I .

Output consists of an ANOVA table, coefficient of determination R^2, standard error, regression coefficients b_j, and a vector of predicted values of Y and residuals.

We used the computer program " REGRESS" on APL by taking

(1) a linear combination of the characteristics, or

$$Y_i = a + \sum_{j=1}^{m} b_j x_{ij} ,$$

(2) a linear combination of logs of the characteristics, or

$$Y_i = a + \sum_{j=1}^{m} b_j \log x_{ij} ,$$

(3) a linear combination of logs of both the characteristics and effectiveness, or

$$Y_i = \exp \left(a + \sum_{j=1}^{m} b_j \log x_{ij} \right)$$

(4) a linear combination of reciprocal of the characteristics, or

$$Y_i = a + \sum_{j=1}^{m} b_j \left(\frac{1}{x_{ij}} \right)$$

(5) a linear combination of reciprocal to both the characteristics and the effectiveness, or

$$Y_i = \left(a + \sum_{j=1}^{m} b_j \left(\frac{1}{x_{ij}} \right) \right)^{-1}$$

(6) a linear combination of square root of the characteristics, or

$$Y_i = a + \sum_{j=1}^{m} b_j \sqrt{x_{ij}} ,$$

(7) a linear combination of reciprocal of square root of the characteristics, or

$$Y_i = a + \sum_{j=1}^{m} b_j \left(\frac{1}{\sqrt{x_{ij}}} \right)$$

and finally

(8) a linear combination square root of both the characteristics and effectiveness, or

$$Y_i = \left(a + \sum_{j=1}^{m} b_j \sqrt{x_{ij}} \right)^2$$

The resulte were shown in Table Ⅴ. Here R^2, the coefficient of determination, shows the proportion of total variance accounted for by the estimating equation as a measure of dispersion, and thus a bigger R^2 is better. The third column shows the standard error which is definded as the square root of the unexplained variance of the dependent variables y. Therefore the smaller the standard error, the better the estimating equation.

The F-statistics is used to test whether the incremental improvement assosiated with the addition of a variable is significant.

Thus the larger the F-ratio the better. The last column shows the coefficient of the variation which relates the standard error (SE) to the mean of the dependent variables y's, or

$$CV = \frac{SE}{\bar{y}} \quad \cdots\cdots (13)$$

This value is used in comparing one standard error with another; a lower CV value is better.

Table Ⅴ. SUMMARY OF RESULTS

REGRESSION	R^2	SE	F	CV
$Y_i = a + \sum_{j=1}^{m} b_j X_{ij}$	0.783	0.267	0.955	0.261
$Y_i = a + \sum_{j=1}^{m} b_j \log X_{ij}$	0.888	0.196	1.978	0.192
$Y_i = \exp[a + \sum_{j=1}^{m} b_j \log X_{ij}]$	0.877	0.194	0.1775	0.189
$Y_i = a + \sum_{j=1}^{m} b_j \left(\frac{1}{X_{ij}} \right)$	0.980	0.083	12.24	0.081
$Y_i = [a + \sum_{j=1}^{m} b_j \left(\frac{1}{X_{ij}} \right)]^{-1}$	0.958	0.110	5.704	0.107
$Y_i = a + \sum_{j=1}^{m} b_j \sqrt{X_{ij}}$	0.838	0.236	1.29	0.231
$Y_i = a + \sum_{j=1}^{m} b_j \frac{1}{\sqrt{X_{ij}}}$	0.937	0.147	3.947	0.143
$Y_i = [a + \sum_{j=1}^{m} b_j \sqrt{X_{ij}}]^2$	0.833	0.116	1.249	0.113

where Y_i : Effectiveness
a : Constant term
b_i : Coefficients
X_{ij} : Characteristic Data (Table I)

b) Selection of the best equation

Looking at Table V, the largest measure of dispersion (R^2) is 0.98, the smallest standard error is 0.083, the hightest F-ratio is 12.24 and the smallest CV is 0.081. Fortunately, all of these values for measures of fit occur when we linearly combine the reciprocal of the data. Here the CV is 8.1 %, which tells us the estimating equation is fitted very well in this case.

Therefore the best estimating relationship developed among those forms investigated is

$$\text{EFFECTIVENESS} = -11.62 - \frac{74.15}{\text{ARMOR}} + \frac{27.03}{\text{SILHOUETTE}} - \frac{622.54}{\text{SPEED}} + \frac{271.94}{\text{Hp/hr}} \quad \cdots\cdots (14)$$

The closeness of the fit of this function to the total effectiveness values furnished by the judges is shown, for individual points, in Table VI.

Table VI. JUDGE EFFECTIVENESS VS FUNCTION EFFECTIVENESS

TANK	JUDGE EFFECTIVENESS	FUNCTION EFFECTIVENESS	DEVIATION	PERCENT DEVIATION
LEOPARD 1	1.44	1.46665	-0.26652	1.85 %
M 60 A 1	1.22	1.16532	0.05468	4.48 %
T- 62	1.02	0.99533	0.024668	2.42 %
M 48 A 5	0.89	0.939137	-0.049137	5.52 %
CENTURION (MK. 13)	0.78	0.800448	-0.010448	1.32 %
AMX - 30	0.78	0.773111	0.006888	0.88 %

We can see from Table VI, that The percent deviations of effectiveness for the six weapon systems are less than 5.52 %.
This suggest that the estimating function of reciprocals of the data fits quite well the data upon which it was developed.

A common practice in attempting to evaluate the effectiveness of a weapon system is to use a simple linear combination of the characteristic values, with coeficients determined by any of several rather arbitray ways. One appreach which is rarely undertaken is to do as was done in this paper, using a least square fit with effectiveness values obtained from expert's judgement.

One possible reason why an estimating function using reciprocals of the characteristice is better fitted than the common procedure of evaluating the effectiveness of a weapon system using a simple linear combination of the characteristics, is that a property like system effectiveness may possess diminishing marginal returns with respect to increasing characteristic values. In the fitting equation (reciprocals) the partial derivative of effectiveness with respect to a characteristic value (with a negative coefficient) was then the reciprocal of the square of the characteristic value, hence diminishing marginal returns.

4. Conclusions

The principal purpose of this study was to determine whether we could compute the overall effectiveness of a weapon system from its characteristics, to establish the existence of functions which could be used to relate system characteristics and effectiveness, and to identify the best estimating relationship. To do this we proposed a procedure, sent appropriate questionnaires, and computed the overall effectiveness values for tank weapon systems by using the constant sum scaling method. Then, using multiple regression we found functional relationships as in table V and evaluating these results we finally found that the best estimating equation occurred when we took the reciprocal, i.e.,

$$Y_i \text{ (Effectiveness) } = a + \sum_{j=1}^{m} b_j \left(\frac{1}{X_{ij}} \right)$$

where

$a = -11.6119$ is an intercept

$b_j = (-74.149, 29.926, -622.536, 291.944)$, and

X_{ij} are characteristics (Table I)

This approach is felt to have merit as of finding an overall MOE because it is based on the opinion of many experts. A conspicuous limitation in the tank example used here is that this study did not include tank characteristics relating to fire power, simply because of the difficulty of data collection.

The scaling for valuing system effectiveness is, of cause, independent of the number of characteristeristics or presence of data on those characteristics. This provides, however, effectiveness values only for the instances listed in the questionnaires. Developement of the functional relationships between characteristics and effectiveness fortunately, requires more data points (instances on the questionnaires) than characteristics if the function-fitting approach used here is employed.

References

[1] Torgerson, Warren, S., Theory and Methods of Scaling, Wiley, 1976.

[2] Batchelder, C. A., An Introduction to Equipment Cost Estimating, The RAND Corporation.

[3] Foss, Christopher F., Jane's World Armoured Fighting Vehicles, St. Martin's Press, 1977.

[4] McNeil, Donald R., Interactive Data Analysis, Wiley, 1976.