IVAL OVERVIEW

By Dr. H.S. Simrin
July 18, 2023

A set of ranking problems that have a truth model is used to test the accuracy of the weighted sum model (WSM). WSM is found to generate rankings that correlate poorly, and sometimes randomly, with the truth model, no matter the choice of weights.

The Intrinsic Value (IVAL) model, a multi criteria decision analysis (MCDA) model that does not use weighting factors, was thus developed. Rather than using weights, IVAL mimics the human decision-making thought process. Correlation of IVAL with the truth model is exceptional (0.95 – .99+) in all cases examined.

Because there are no weights, there is no need with IVAL for pairwise comparison or other complex processes that generate relative weights. For example, there is no need for a structure like House of Quality, used in Quality Function Deployment (QFD).

Because IVAL ratings have such high fidelity, techniques to improve initial outcomes, such as backscatter, used in deep neural networks, may not be necessary.

The IVAL model easily incorporates criteria dependencies and hierarchical processes. It requires only very simple code that could easily replace just the WSM portion of code used in many current MCDA models. It would be interesting to see the results of doing this.

Acknowledgments
(Specifics included in main body)
Gordon B. Dobbins, Jr.
Dr. Charles B. McLane

Keywords
criterion space
multi criteria decision analysis (MCDA)
multiple criteria hierarchy
ranking process
utility function
weighted sum model (WSM)

Table of Contents

I INTRODUCTION 3

II CASE STUDY – THE WEIGHTED SUM MODEL (WSM) 5

III DISECTING THE WSM MODEL 8

IV THE IVAL METHODOLOGY 9

V CASE STUDY USING THE IVAL METHODOLOGY 12

VI ADDITIONAL COMPARISONS OF WSM AND IVAL 14

VII SUMMARY 16

VIII APPENDICES 17

- A Measuring Goodness of Fit of Rankings 17
- B IVAL Handling of Dependent Criteria 19
- C IVAL Treatment of Hierarchies 23
- D S-Curve Formulas 24
- E Spline Curve Formulas 27
- F Exponential Curve Formulas 28
- G Fundamental Equation of Military Aircraft Combat Effectiveness 30

List of Tables

Table 2–1 WSM Formula Versus "Truth"………………………….. 6
Table 2–2 WSM Formula Does Not Preserve Rank Order………. 7
Table 5–1 IVAL Case Study Results………………………………. 13
Table 6–1 Results of Full Factorial Designs……………………….. 14
Table A–1 Quality of Full Factorial Design Results………………. 17

List of Figures

Figure 4–1 Just 3 Features Are Needed To Mimic Human Decision-Making…… 9
Figure 4–2 Types of Curves…………………………………………………….. 10
Figure 5–1 IVAL Curves Used In Case Study…………………………….………..12
Figure B–1 IVAL Exponen…………………………………………………….... 19
Figure B–2 Dependent Criteria……………………………………………………. 20
Figure B–3 Criterion With One Strong and One Other Dependency……………. 22
Figure E-1 Spline Curves……………………………………………………….... 27

List of Equations

2–1 Weighting Factor Equation 5
2–2 Fundamental Equation of Military Aircraft Combat Effectiveness 5
4–1 IVAL Utility Function 11
A–1 Pearson r Correlation Coefficient 17
B–1 IVAL General Equation 21

I INTRODUCTION

Multiple-criteria decision analysis (MCDA) utilizes a wide range of methods and procedures for rating and ranking alternatives. Alternatives are rated by assigning values for selected criteria. There is generally no truth model, a “natural equation” to combine the criteria values into a single metric, a total value.

Even though there may be no optimum solution to a given ranking problem, it is still possible to measure how well an MCDA model performs by first applying it to a ranking problem for which there is a truth model. Such a ranking problem is presented in Section II, and it is used to evaluate the weighted sum model (WSM).

The truth model is used to measure the goodness of WSM ranking. Several case studies are conducted. It is found that WSM performs very poorly at ranking no matter what choice of weighting factors is selected. In Section III, the reasons for this poor performance are discussed and analyzed.

One of the key issues identified is the very use of weights, themselves. Assigning relative weights feels logical. How, one might ask, could weights go wrong? First, they go wrong because there is no possible choice of weights in the Section II case study that provides even remotely reasonable rankings. A different example, an even simpler study for which weights make no sense, is presented in Section III.

Weights also go wrong because there cannot be more than one dominant weight (i.e., greater than 0.50). Consider the design of an airplane. Suppose that airplane range of 3000 miles and speed of 250 knots are critical. How much weight can be assigned to these criteria? If, say, twenty criteria are being considered, then the average criterion weight is only 0.05, and we cannot realistically assign a weight larger than, say, 0.20, to either of these criteria. This is not enough to eliminate an aircraft alternative that has an unacceptable 2500-mile range but exceptional criteria otherwise. Such an alternative might even receive the highest total value.

If that happens, one expects the decision maker to spot the flawed design and throw it out. The problem is that *there will be* many other poor designs that are highly rated, but which are difficult or impossible to spot. We will see exactly this situation in the Section II case studies.

A weighted sum is of course the perfect tool when it comes to computing statistical means, like the average of student test grades. This does not, however, justify the use of weighting factors for evaluating criteria. Weighting factors have been universally used for MCDA because of the need for a process to perform such evaluations, and weighting factors “feel” like they should be appropriate. But it is important to measure the goodness of an MCDA process before using it, and WSM is sadly lacking.

In Section IV, we introduce the **intrinsic value model (IVAL)**. This model performs MCDA ranking without the use of weights. Rather than weights, IVAL uses criteria curves. The curves are defined in terms of the three properties that decision-makers unconsciously use when evaluating criteria. The properties are (1) regions of diminishing returns, (2) regions of escalating returns, and (3) criticality of each criterion. Each curve represents the total value contribution of a single criterion, which is why we call the curves “intrinsic”.

In Section V, we repeat the case studies, using IVAL instead of WSM, and we illustrate the three intrinsic properties. In Section VI, we perform additional case studies using both WSM and IVAL. We find that IVAL achieves extremely high correlation with the truth model in all the cases. WSM generates unacceptable correlation across the board except that it generates acceptable results in certain extreme circumstances.

Because the IVAL utility function consists of a single equation, just like WSM, it should be possible to easily modify existing WSM computer models by replacing the WSM summation with the IVAL product. It may also be possible to replace the use of weights in some other MCDA models by IVAL curves. It would be interesting to do this and compare the results.

It turns out that the “simple” use of weighting factors in MCDA problems is not simple at all. Pairwise comparison of criterion weights must be performed, and there are many combinations to compare. Large-scale structures like House of Quality have been developed just to handle this unwieldly process. In IVAL, this problem just disappears. This is because IVAL curves are intrinsic, not relative, and so there is nothing to pairwise compare.

The simplicity of IVAL is apparent in Appendix B where we show how independent and dependent criteria are handled. Even very complex dependencies are simple and straightforward to include. We underscore this in Appendix C where treatment of hierarchies, whether simple Analytic Hierarchic Processes (AHP) or more complex Analytic Network Processes (ANP), is reduced to triviality.

Appendix A provides a reminder of the Pearson r Correlation Coefficient, the measure we use to determine goodness of rankings. Additional case studies were performed for Appendix A, and the author was able to quantify the relationship between correlation coefficient and ranking fidelity. It was found that an extremely high correlation coefficient of 0.90 is required just to produce a marginally acceptable ranking study, and 0.95 is required to achieve a fully correct ranking. IVAL achieved 0.95 or higher across the board.

Appendix G provides the derivation of the truth model. The other appendices provide templates for commonly used IVAL equations to facilitate readers who may wish to explore this approach.

II CASE STUDY – THE WEIGHTED SUM MODEL (WSM)

The goal of the case studies in this paper is to evaluate and rank design alternatives for a new military aircraft. In military combat analysis, there are always three criteria to measure: offensive capability, defensive capability, and frequency of combat events[1]. In this study, defense is measured by **Survivability**, an aircraft's ability to fly its mission into enemy territory and return safely. Offense is measured by **Lethality**, the average number of targets an aircraft can destroy per mission. Frequency is measured by sortie rate. A **sortie** is a mission; it is a round-trip flight into enemy territory. **Sortie rate** is the number of sorties an airplane can generate per day.

The weighed sum model begins with the assignment of **weights, w_1 through w_n**, to n criteria. In these case studies, n = 3 because there are three criteria: survivability, lethality, and sortie rate. For a given alternative, let x_i denote **the criterion value** and let v_i be **the normalized value**. In most cases, normalization amounts to dividing each x_i by some maximum x-value to obtain $0 \le v_i \le 1$. This leads to the WSM utility equation:

Weighting Factor Equation

$$V = w_1 v_1 + w_2 v_2 + \ldots + w_n v_n \tag{2–1}$$

where V is the **total value**.

The true relationship between the case study criteria is specified by the

II

$$K = \sqrt{P_S}\,\frac{1-P_S^{SR \cdot 10}}{1-P_S} K_S \tag{2–2}$$

where

P_S = Probability of survival per sortie,

SR = Sortie rate,

10 = Ten days of combat,

K_S = Number of target "kills" per sortie, given that the aircraft reaches the target area, and

K = Number of target kills in 10 days.

P_S, SR, and K_S are the criterion values, and K represents the combat outcome for a single aircraft. It is the **total value** of the truth model. For the interested reader, the Fundamental Equation is derived in Appendix G.

When choosing weighting factors, usually survivability is given top priority. A typical weight is 0.5. Typical weights for SR and K_S are 0.25 each.

[1] Cost, also, is always considered, but is omitted to keep this case study as simple as possible.

To illustrate the problem with the WSM model, we present three case studies in this section. Eight aircraft configuration alternatives are defined. Case Study 1 compares alternatives 1 – 4. Case Study 2 compares alternatives 4 – 8. Case Study 3 compares alternatives 1 – 8.

The first four alternative aircraft configurations are tabulated in Table 2–1. On the left side of the table, the configuration criteria values (*SR, P_S*, and K_S) are shown, and the truth model total value (*K*) from the Fundamental Equation is computed. On the right-hand side, the WSM model is shown. The normalized criteria values v_1 and v_3 are obtained by dividing *SR* and K_s values by 10 (the max value of each) to obtain quantities between zero and one. v_2 is set equal to P_S because probability values are already normalized.

Due to using normalized values, the WSM total value, *V*, falls between zero and one. The Fundamental Equation's total value, *K*, can be any size. To make it easier to visually compare the *K* and *V* answers, *V* is multiplied by 100 in the table. Total values are colored **green**, **yellow**, and **red** for **excellent**, **mediocre,** and **poor** scores. We see that the Fundamental Equation has two groups: excellent and poor. WSM has one group, mediocre.

Table 2–1 WSM Formula Versus "Truth"

Configuration				**Fund Eq**	**Weighting-Factor Equation**						
No.	**SR**	P_S	K_S	K	w_1	v_1	w_2	v_2	w_3	v_3	**100 x** V
1	10.0	0.99	0.1	6.3	0.25	1.00	0.50	0.99	0.25	0.01	74.8
2	10.0	0.30	10.0	7.8	0.25	1.00	0.50	0.30	0.25	1.00	65.0
3	0.1	0.99	10.0	9.9	0.25	0.01	0.50	0.99	0.25	1.00	74.8
4	5.0	0.95	5.0	90.0	0.25	0.50	0.50	0.95	0.25	0.50	72.5

Alternatives 1 - 3 have been contrived to have fundamental design flaws. They are designed with two excellent criteria and one unacceptable one. Alternative 1 has excellent *SR* and P_S, but the small offensive value, K_S, makes this alternative militarily useless. Similarly, alternative 2 has excellent values for *SR* and K_S but its very small defensive value, P_S, makes it also useless as a military aircraft. Finally, alternative 3 has excellent values for P_S and K_S, but an awful sortie rate. The balanced design, configuration 4, is by far the best choice.

The Truth Model correctly identifies alternative 4 as best, and it gives it a total value of 90, an order of magnitude better than the three "flawed" airplanes. However, WSM scores two of the three flawed airplanes higher than alternative 4.

The reader might think that the WSM results can be improved by a better choice of weights. In fact, this is not the case. **No assignment of weights can remove the problem**. No matter what choices one makes, one or more of the flawed configurations is the clear winner. This is because each of the flawed configurations fails in a different criterion. Decreasing any one weighting factor will make the airplane with that flaw the winner because it has the highest value for the other two criteria. Decreasing two weights makes the 3rd weight dominant, and two of the flawed airplanes have the highest value of that criterion.

Table 2-1 contains one "reasonable alternative" (configuration 4) and three outliers (configurations 1 – 3). Thus, the clearly wrong answers are easy to find and discard. To counter this, in Case Study 2, we compare five alternatives, all of which are feasible. This case study is tabulated in Table 2-2. As always, WSM gives answers, but this time the wrong answers will not be easy to spot.

Table 2–2 WSM Formula Does Not Preserve Rank Order

Configuration				**Fund Eq**	**Weighting-Factor Equation**						
No.	***SR***	P_S	K_S	K	w_1	v_1	w_2	v_2	w_3	v_3	**100 x V**
1	10.0	0.99	0.1	6.3	0.25	1.00	0.50	0.99	0.25	0.01	74.8
2	10.0	0.30	10.0	7.8	0.25	1.00	0.50	0.30	0.25	1.00	65.0
3	0.1	0.99	10.0	9.9	0.25	0.01	0.50	0.99	0.25	1.00	74.8
4	5.0	0.95	5.0	90.0	0.25	0.50	0.50	0.95	0.25	0.50	72.5
5	1.0	0.98	5.0	45.3	0.25	0.10	0.50	0.98	0.25	0.50	64.0
6	5.0	0.98	3.0	94.4	0.25	0.50	0.50	0.98	0.25	0.30	69.0
7	7.0	0.95	3.0	56.9	0.25	0.70	0.50	0.95	0.25	0.30	72.5
8	3.0	0.95	7.0	107.2	0.25	0.30	0.50	0.95	0.25	0.70	72.5

In Table 2-2, configurations 4 - 8 are compared. (Configurations 1–3 are dimmed, for reference.) It is difficult to intuitively rank the five alternatives because each of the last four alternatives has one value the same, one value better, and one value worse than configuration 4. So, which are better than alternative 4 and which are worse? The Fundamental Equation uncovers the true total values.

Notice that the Fundamental Equation separates configurations 4 - 8 into two groups. Configurations 5 and 7 (colored yellow) are grouped together, having roughly ½ the value of the other three configurations (colored green). WSM does not even identify that there are two groups. Plus, it generates the same total value for the best and the 2nd worst of the five alternatives. In fact, none of alternatives 4 - 8 outscores the flawed configurations 1 and 3!

The danger with the WSM model is that one never knows which answers are wrong. It is practically impossible to identify mediocre alternatives 5 and 7 using intuition, and so they aren't automatically discarded like alternatives
1 –3. In real case studies, there is no truth equation and, so, alternative 7 would be accepted as achieving the highest total value.

This illustration suggests that we need to find an MCDA model that does not rely on weighting factors. Such a model is developed in Section IV, and we measure it's goodness of fit in Section V.

In Section VI we conduct three additional case studies. We will see that WSM *can* generate acceptable rankings if a design space is sufficiently tight. But, if the design space contains variety, then the WSM approach quickly worsens.

WSM will always give answers, but are they almost random as in Table 2-2 ,or are they acceptable as in a tight design space?

III DISECTING THE WSM MODEL

Why is WSM acceptable for ranking a tight design space but not a wider one? The answer is that complex relationships in mathematics are represented by non-linear equations, i.e., curves. The WSM utility function is linear. If you try to approximate a curve using a line, there is a small region of validity where the line and the curve are not far apart. This constitutes a tight design space where WSM can be acceptable.

A more fundamental flaw of WSM is the very notion that weights are an appropriate tool for decision making. A simple example of this flaw appeared in early psychology studies (prior to 1970) where it was debated whether heredity or home environment is the more important discriminator of a person's behavior. Conflicting studies "proved" that each criterion (heredity, home environment) contributes 70% to the results. That is, some studies showed conclusively that bad heredity led to a poor personality even when the home environment was good. But other studies conclusively showed that a bad home environment resulted in a poor personality even if heredity was fine. However, two 70% factors sum to 140%, a nonsense result.

Psychologists finally "resolved" this "inconsistency" by realizing that the two criteria should be multiplied instead of added. By multiplying the factors, either discriminator could play a dominant role in the final answer. That is, if either criterion had a low value, then it dominated and caused a low total value. In this way, both criteria could contribute more than half.

As a side note, this example shows that WSM can have difficulties even when there are only two criteria being considered.

The author found that a multiplicative rating system is a good start for a fix, but it is not quite enough. In this paper, we expand that approach and abandon the concept of weights. In the next section, the IVAL scheme is introduced whose very equations are an attempt to imitate the human way of thinking when making decisions.

IV THE IVAL METHODOLOGY

Instead of relative weights, the IVAL model uses intrinsic functions, i.e., a curve for each criterion. The label "Intrinsic" expresses that a curve represents a criterion's individual contribution to the total value.

The formula for a curve, $v_i = f(x_i)$, is called an **IVAL equation**. The domain of a curve consists of ***criterion values***, x_i. For the case study, x_i is a value of P_S, K_S, or SR. The variable v_i is called an ***IVAL factor*** (because we will later multiply by it).

It has been discovered that just three features are needed to define a curve to make it mimic the human decision-making thought process: steep regions, flat regions, and floors, illustrated in the Figure 4–1, below.

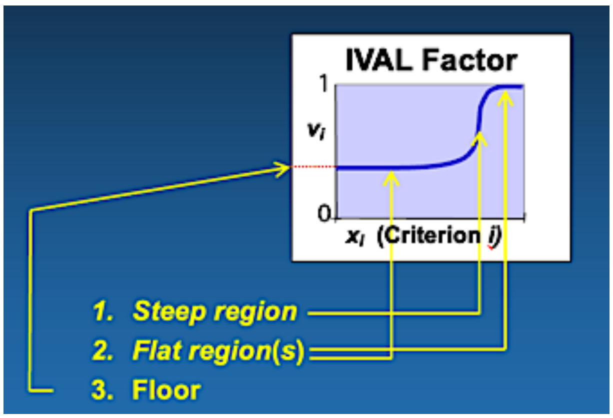


Figure 4–1 Just 3 Features Are Needed To Mimic Human Decision-Making

To motivate steep and flat regions, consider an airplane that carries a missile with a 50-mile range. How much range must the airplane's radar have to enable the missile to be launched at the maximum 50-mile range? After an enemy aircraft is detected by the radar, the pilot needs time to react, identify, maneuver, aim, and then launch the missile. Suppose the bare minimum time to complete this sequence of tasks requires a radar range of 60 miles, and the optimal time requires 80 miles. The lower flat region of the radar-range intrinsic curve in Figure 4–1 would be from 0 to 60 miles on the *x*-axis. Any improvement to radar range from 0 to 60 miles would be of little value. Above 80 miles would be the upper flat region in the figure. Even if the pilot has excess time to launch the missile, there is little benefit because the pilot must wait until the two aircraft close to within 50 miles of each other. The steep region of rapidly increasing value in Figure 4–1 would be from 60 to 80 miles on the *x*-axis.

In the figure, the flat regions are regions of *diminishing returns*. **The upper flat region** is where "enough is enough". Once we achieve enough of this capability, additional capability does little to improve total value.

The lower flat region is where the criterion value is too small to provide a meaningful contribution to the desired total value. It is where a little short is the same as being a lot short.

The steep region, between two inflection points, is the region of *escalating returns*, where small improvements in criterion value result in escalated total value.

The floor is the measure of *a criterion's importance.* A critically important criterion has a floor near zero. This causes a poor score for this criterion to generate a near-zero IVAL factor, and multiplying by this factor makes the total value near zero. A criterion that is "nice-to-have" but not essential will have a floor closer to one. For example, a 0.9 floor means that the total value is only reduced by 10%, even if there is no capability at all in this criterion.

Curves can have various shapes (Figure 4-2). Commonly, an IVAL curve will be an "S" curve with two flat regions and two inflection points as in Figure 4-1. But some criteria will be described by an exponential curve, a curve with only one inflection point and only one flat region. Or, an IVAL curve can be a straight line, a curve having no flat region (i.e., no region with slope zero) and no inflection point. Curves can be increasing, where more is better, or decreasing, where less is better. Curves can have a finite range of criterion values (x-axis) or an infinite range of criterion values. Equations for all these types of curves are provided in the appendices. If some other curve shape is appropriate, the analyst is free to use his own curve.

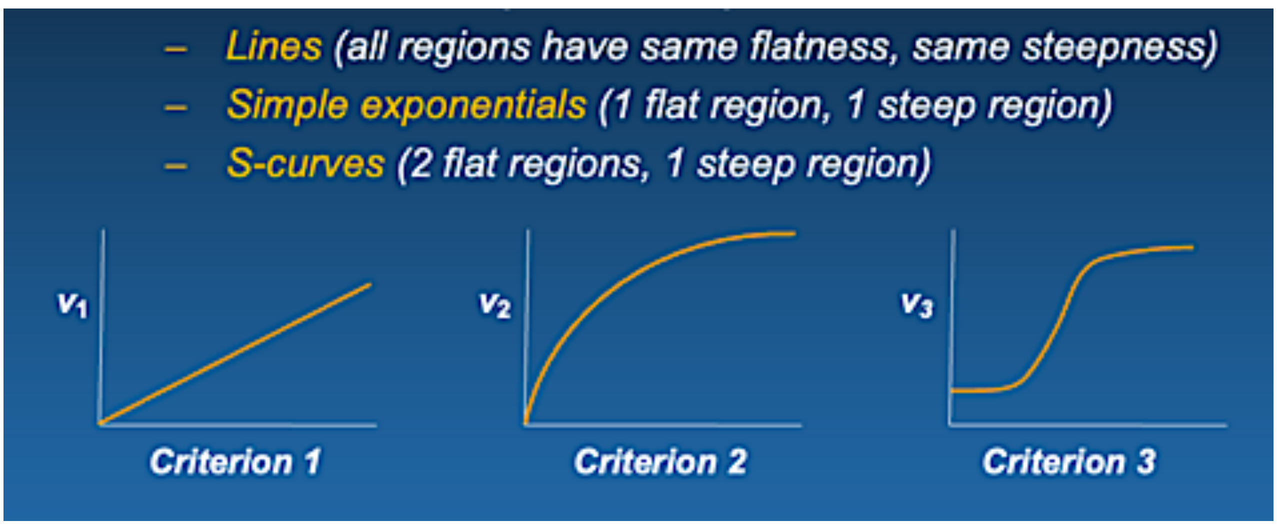


Figure 4–2 Types of Curves

When developing an IVAL curve, like the radar curve just discussed, the analyst should assume that all the other criteria are excellent and that the total value is driven by the IVAL factor of just this criterion. The analyst's task is to develop a curve for the criterion that reflects the amount of degradation from a perfect score that will occur over the range of values of the criterion. This curve then represents the intrinsic value of the criterion.

Development of the IVAL curves is the first step in an IVAL ranking study. The second step is to plug the criteria values of each alternative into the IVAL curves to get the IVAL factors, v_l, for each alternative.

The third step in the IVAL process is simply to multiply the IVAL factors for each alternative to get the total values. We call this product the

IVAL Utility Function

$$V = v_1\ v_2\ v_3\ \ldots\ v_n \qquad (4\text{–}1)$$

At first glance, it may appear that the utility function is nothing more than a simple multiplicative scheme. Keep in mind, however, that each v_i is the result of a curve. That is,

$$V = f(x_1)\ f(x_2)\ f(x_3)\ \ldots\ f(x_n)\ ,$$

Thus, the IVAL utility function is non-linear both because multiplication is non-linear and because the IVAL curves can be non-linear. Recall that linearity was one of the shortcomings of WSM, restricting its use to a set of tightly similar alternatives. The steep regions and the flat regions of IVAL change the criterion values in an appropriate, non-linear fashion that enables faithful ratings even for sets of highly varied alternatives.

In this way, IVAL factors represent human decision-making by capturing just three criteria features: diminishing and escalating returns and criteria importance. Having just a single value, a weight, cannot begin to capture the behavior of the regions of diminishing and escalating returns.

In a sense, the (intrinsic) floors replace the (relative) weights, but they do a much better job of modeling “importance”. There can be at most one weight that is large enough to eliminate an alternative that lacks it, but there is no limit to the number of critical floors, i.e., curves having floors near zero that can eliminate alternatives. Moreover, in a large study with 10, 20, or more criteria, each WSM weight is so small as to be rendered meaningless, whereas a single deficient IVAL factor will eliminate an alternative.

V CASE STUDY USING THE IVAL METHODOLOGY

In this section, IVAL is applied to Case Study 3, all eight alternatives (Figure 5-1). Expert judgment was used to select the types of curves and to identify the region or regions of diminishing returns.

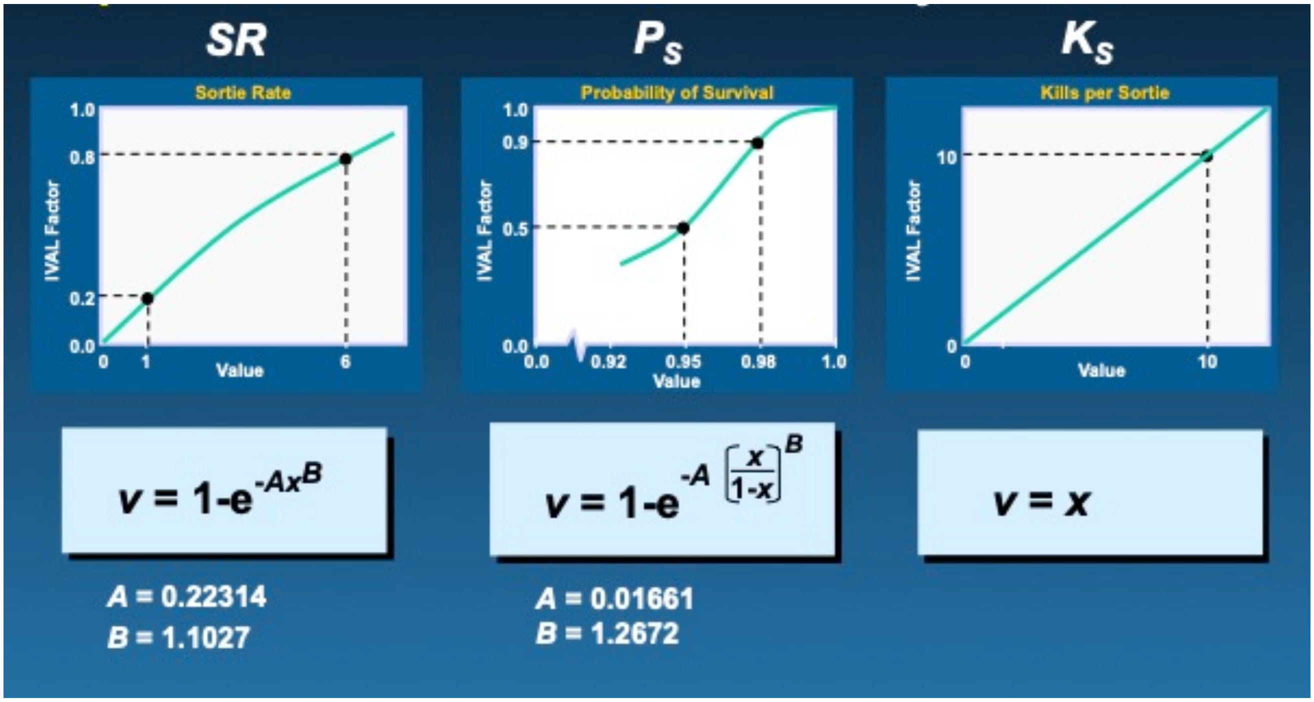


Figure 5–1 IVAL Curves Used In Case Study

Only Probability of Survival (P_S) was modeled as an "S" curve. 98% survivability was judged to represent a 10% capability loss over the period of 10 days of combat, hence the upper inflection point[2] at (0.98, 0.9). A 95% survival rate, which is a whopping 5% loss rate per sortie, was judged to have reduced capability by 50%, hence the (0.95, 0.5) lower inflection point. (There was not room to show the entire *x*-axis. Notice the broken *x*-axis near 0.)

Sortie rate was judged to be almost a linear improvement until about 6 sorties per day, with diminishing returns after that. It was modeled as an exponential curve with an inflection point at (6, 0.8) and a region (not clearly shown) of diminishing returns above that.

As for lethality (K_S), it was decided that more target kills per sortie means proportionately more target kills in ten days, without bound. For example, if x_i triples, then v_i triples. Thus, a really simple $v = x$ linear curve. The implicit assumption here is that there is an unlimited target set; the offense never runs out of targets. In a less target-rich environment, an exponential curve would have been more appropriate. The version of the Fundamental Equation of Combat Aircraft Survivability used in this study was developed for a target-rich environment, so we used that for our lethality curve guestimate.

---

[2] The term "inflection point" is used loosely. An S-curve can be defined by selecting any two points.

Lastly, consider the floors. All three criteria are critical criteria. Weakness in offense, defense, or frequency of flight leads to a terrible design. Hence, all three criteria were given floors of zero. In all three equations, an *x*-axis value of zero yields a *y*-axis value of zero.

Table 5–1 IVAL Case Study Results

Configuration Values				IVAL Factors			Total Values		Fund Eq
No.	*SR*	P_S	K_S	*SR*	P_S	K_S	*Y*	50 *Y*	*K*
1	10	0.99	0.1	0.941	0.996	0.1	0.094	5	6
2	10	0.30	10	0.941	0.006	10	0.053	3	8
3	0.1	0.99	10	0.017	0.996	10	0.174	9	10
4	5	0.95	5	0.732	0.500	5	1.830	91	90
5	1	0.98	5	0.200	0.900	5	0.900	45	45
6	5	0.98	3	0.732	0.900	3	1.976	99	94
7	7	0.95	3	0.852	0.500	3	1.277	64	57
8	3	0.95	7	0.527	0.500	7	1.846	92	107

Table 5-1 shows that the simple IVAL utility function matches the truth model very closely. IVAL picked out the reds, yellows, and greens. The Pearson r correlation coefficient for this case is 0.99.

Statisticians consider a correlation such as 0.90 to be outstanding, but this is not the case for a ranking study. As can be seen, even this 0.99 correlation has some errors in ranking. The best design configuration according to the truth model is #8, followed by #6. IVAL got these reversed. The reader can just imagine how many rank reversals occur with a .95 correlation, and with a .90 correlation there will be numerous rank reversals and also many badly ranked alternatives.

In Appendix A, it is shown that 0.95 is the smallest correlation to ensure correctly identifying all the top candidates. Anything below .90 is problematic. The Pearson r correlation coefficient for this case study using WSM was .038, almost random, which is not surprising since WSM could not even identify the correct colors.

Just a quick note. The author did not “tweak” the IVAL curves to achieve this result. I simply gave my expert opinion best guess at the curves, once, and this is the result. I might have “tweaked” a different approach for a lifetime and not achieved this level of correlation.

I was astounded that the results were this good. It was unexpected. Gordon Dobbins[3] and I had been trying for 15 years to develop a useful utility function for ranking problems, and we never had any success at all.

[3] Gordon B Dobbins, Jr (deceased), worked at General Dynamics, Fort Worth Division which later became Lockheed Martin Tactical Aircraft Systems. I developed a forerunner to this approach while working with

VI ADDITIONAL COMPARISONS OF WSM AND IVAL

Next, we examine alternatives in three different design spaces. In the tight design space, all the alternatives have similar design criteria. In the medium design space, we allow greater variability, and more so in the wide design space. To cover the range of designs in each space, we conduct a full-factorial study with the following criterion ranges:

	SR	*PS*	K_S
Tight	1–3	.97–.99	1–3
Medium	1–5	.94–.99	1–5
Wide	1–9	.91–.99	1–9

In each design space we examine 27 aircraft designs: (3 sortie rates) x (3 *PS*'s) x (3 K_S's). The correlation coefficients for these spaces are listed in Table 6–1.

Table 6–1 Results of Full Factorial Designs

	Pearson r Correlation Coefficient				
	Original 8 Designs	Only the 5 Feasible Designs	Tight design space	Medium design space	Wide design space
WSM	.038	.584	.979	.897	.736
IVAL	.987	.949	.996	.986	.985

We observe that WSM generates almost random answers for the original 8 designs. It fairs considerably better when the infeasible designs 1 – 3 are thrown out. The .584 result it achieves shows clear correlation but is insufficient for ranking. As explained in appendix A, a 0.95 correlation coefficient is the minimum desired for a ranking study. So, WSM is very acceptable for ranking for the tight design space where the alternatives are quite similar. The .897 coefficient obtained by WSM for the medium design space can identify only 2/3rd of the top candidates, and that is not ideal. For the wide design space, WSM can identify less than 1/3rd of the top candidates.

By contrast, IVAL has consistent correlation coefficients across the design spaces, from the tight to the wide, and can identify 100% of the top candidates in all examined cases.

Gordon, and I fully developed IVAL after retirement. Charles B McLane has developed advanced equation templates for IVAL, and he has provided many keen observations.

We summarize why IVAL does so well.

1. **IVAL factors are intrinsic**. *Intrinsic* means that an IVAL curve is developed for each criterion without regard for the other criteria. The intrinsic values are defined in terms of how much or how little the criterion contributes to the total value.

2. **The IVAL factors are degradation factors.** This means that if an alternative has perfect v_i scores, say 1.0, for all the criteria except one, and the last criterion has, say, v = 0.20, then the total value is degraded from a perfect 1.0 down to 0.20. As a result, there can be as many critical criteria as needed, and “nice-to-have” criteria can be assigned appropriate less-critical degradations. Criticality is defined by the criterion’s floor.

3. **The IVAL curves identify the regions of slowly and rapidly changing degradation.** This enables the IVAL utility function to give proper credit (or demerit) to criteria values. This is what enables IVAL to succeed even when there is large design variability.

We also summarize how similar the WSM and IVAL model steps are.

	Step	Description
WSM	1	Develop criterion relative weights w_i
	2	Generate criterion normalized values v_i (for each alternative)
	3	Add, to get total value (for each alternative)
		$V = w_1v_1 + w_2v_2 + w_3v_3 + \ldots + w_nv_n$

	Step	Description
IVAL	1	Develop IVAL curves
	2	Use curves to generate IVAL factors v_i (for each alternative)
	3	Multiply, to get total value (for each alternative)
		$V = v_1 \bullet v_2 \bullet v_3 \bullet \ldots \bullet v_n$

Consequently, WSM is easily replaced by IVAL, even in complex computer code. Replace the inputting of weights with the inputting of the IVAL curves. Replace the normalized WSM values by IVAL factors from the curves. Replace the WSM summation in Step 3 by the IVAL multiplication. Except for input, and maybe some output displays, no other code change might be required.[4]

[4] Code that might be included to compute the relative weights, or to provide “corrected” weights, can be discarded.

VII Summary

IVAL reflects the considerations that people make when evaluating criteria, and it was shown to generate excellent ratings and rankings across a broad range of cases.

WSM was shown to generate good rankings only for tight design spaces and, even then, IVAL performed better.

The weighting sum model, and weighting factors in general, thus constitutes a weak foundation for MCDA models. Because it is so easy to replace weights by IVAL curves and to replace the WSM summation by the IVAL product, it may be possible to strengthen the foundation of existing MCDA models by replacing WSM by IVAL.

MCDA studies using IVAL are simpler than studies that use WSM because it is much easier to generate IVAL curves than it is to generate weighting factors. IVAL curves are quickly drawn because they graphically represent the way we already think. Weighting factors require pairwise comparison or other time-consuming processes to generate.

Finally, it should be possible to eliminate the need for techniques such as backscatter, which was developed for deep neural networks to “correct” the weighting factors. In neural science it is widely known the WSM does not represent the way that neurons work when decisions are made, and so backscatter is one technique that was developed to improve the weights.

VIII APPENDICES

A Measuring Goodness of Fit of Rankings

Pearson's r correlation coefficient is the standard way of measuring the goodness of fit of two random distributions. Let *X* be the random variable representing the truth model distribution and *Y* the random variable representing either the WSM or IVAL distribution. The Pearson r correlation coefficient is defined as

VIII

$$\rho = \frac{\mathrm{cov}(X,Y)}{\sigma_X \sigma_Y} = \frac{E\left(\left(X-\mu_X\right)\left(Y-\mu_Y\right)\right)}{\sigma_X \sigma_Y} \qquad \text{(A–1)}$$

where cov(*X*, *Y*) is the covariance of *X* and *Y*, E(…) is the expected value, μ_X is the mean of the *X* values, μ_Y is the mean of the *Y* values, and σ_X and σ_Y are the standard deviations. ρ ranges from -1 to +1. Zero means the two distributions are uncorrelated, that the results are random. +1 represents perfect correlation, and -1 represents perfect anti-correlation (not relevant here).

Table A-1 shows correlations in green (excellent), yellow (mediocre), and red (poor). It shows that the correlation of WSM varies widely. Without a truth model to tell us, we have no way of knowing whether a WSM ranking is good or bad. IVAL correlations, on the other hand, are consistently excellent across studies.

Table A–1 Quality of Full Factorial Design Results

	Pearson r Correlation Coefficient				
	Original 8 Designs	Only the 5 Feasible Designs	Tight design space	Medium design space	Wide design space
WSM	.038	.584	.979	.897	.736
IVAL	.987	.949	.996	.986	.985

In large studies it is important to be able to correctly identify the top 10% of alternatives so they can be carried forward for more detailed study and analysis. We examined the spreadsheet used to generate this data, and additional cases as well, to determine how many of the alternatives in the Truth Model's top 10% are identified. We found:

1. Only 1/3rd of the top alternatives are identified when $\rho = 0.75$.
2. 2/3rd of the top alternatives are identified when $\rho = 0.90$.
3. All the top alternatives are identified when $\rho = 0.95$.
    a. However, there are many rank reversals when $\rho = 0.95$.
    b. There are still a few rank reversals even when $\rho = 0.99$.
    c. To eliminate rank reversals requires $\rho = .995$, which IVAL achieved for the tight design space.

We conclude that $\rho = 0.95$ is the minimum desirable Pearson r coefficient, and a coefficient less than $\rho = 0.90$ is not acceptable. A correlation coefficient of 0.90 will mis-identify 1/3 of the top alternatives, and lower coefficients will lead decision makers to make even poorer decisions.

B IVAL Handling of Dependent Criteria

Criteria can range from independent to dependent. Assessment of criteria dependence is often included in alternative ranking studies.

Suppose that

$$V = v_1\ v_2\ v_3\ v_4.$$

This represents IVAL treatment of four mutually **independent** criteria. But suppose that criteria 3 and 4 are (fully) dependent, that all values of v_3 are the same as v_4. The equation above would then yield

$$V = v_1\ v_2\ v_3\ v_4 = v_1\ v_2\ v_3\ v_3$$
$$= v_1\ v_2\ v_3^2\ ,$$

and we see that the equation double accounts v_3. Because v_1, v_2, and v_3 are independent, they should each be represented just once. That is, we should have

$$V = v_1\ v_2\ v_3\ .$$

But this is the same as

$$V = v_1\ v_2\ v_3^{1/2}\ v_4^{1/2}\ .$$

This is the recommended IVAL utility function equation for when two criteria are (fully) **dependent.**

Figure B–1 plots the exponent of the IVAL factors. Figure B-1a is a straight-line graph of the exponent of v_3 (or v_4), from the dependent point in the lower right to the independent point in the upper left.

In Figure B–1b, if v_3 and v_4 are only **moderately dependent** (that is, not equal, but rather tending in the same direction), we could choose an exponent of .75, halfway between 0.5 and 1:

$$V = v_1\ v_2\ v_3^{3/4}\ v_4^{3/4}\ .$$

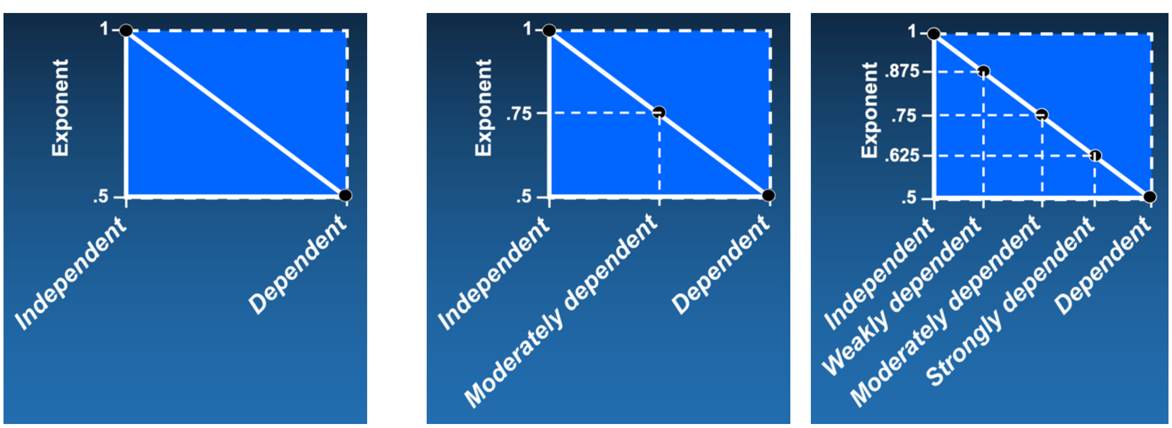


Figure B–1 IVAL Exponent

(a) (b) (c)

¾ + ¾ adds to less than 2 but more than 1. This is appropriate because, together, they represent a mix of independence (two criteria) and dependence (one criterion).

If the criteria are only slightly dependent rather than moderately, we could split the difference again, between 0.75 and 1, as shown in Figure B-1c, to get an exponent of 0.875:

$$V = v_1\ v_2\ v_3^{7/8}\ v_4^{7/8}\ .$$

In the other direction, if the criteria are more than moderately dependent but not fully dependent, we could split the difference between 0.5 and 0.75 to get 0.625:

$$V = v_1\ v_2\ v_3^{5/8}\ v_4^{5/8}\ .$$

We named the latter two cases "**weakly" and "strongly" dependent** in Figure B-1c.

The cases can all be written as $V = v_1\ v_2\ v_3^{q3}\ v_4^{q4}$ where $q_3 = q_4$ is chosen from the graph in Figure B-1c according to the degree of dependency.

Similarly, if three variables are (fully) dependent and $v_2 = v_3 = v_4$, then including all three terms would be triply redundant:

$$V = v_1\ v_2\ v_3\ v_4 = v_1\ v_2^3.$$

If they are mutually dependent, we should write the total value as

$$V = v_1\ v_2 = v_1\ v_2^{1/3}\ v_3^{1/3}\ v_4^{1/3}\ .$$

This is the recommended IVAL treatment for when three criteria are (fully) dependent, and is shown in Figure B–2.

If the criteria are only moderately dependent, then instead of 1/3 we would split the difference between 1/3 and 1 as shown in Figure B-2a, to get an exponent of 2/3:

$$V = v_1\ v_2^{2/3}\ v_3^{2/3}\ v_4^{2/3}\ .$$

Figure B-2a also shows that the exponents for three strongly or weakly dependent criteria are found by again splitting the differences

Dependent	0.333
Strongly Dependent	0.500
Moderately Dependent	0.667
Weakly Dependent	0.8335
Independent	1.000

Dependent	0.250
Strongly Dependent	0.4375
Moderately Dependent	0.625
Weakly Dependent	0.8125
Independent	1.000

Figure B–2 Dependent Criteria
(a) Three Criteria (b) Four Criteria

Figure B-2b extends this result to four correlated criteria. For example, if criteria 1 and 2 are independent of the other criteria but criteria 3 – 6 are (fully) dependent, then total value would be

$V = v_1\ v_2\ v_3^{1/4}\ v_4^{1/4}\ v_5^{1/4}\ v_6^{1/4}$.

Were criteria 3 – 6 moderately dependent, then total value would be

$V = v_1\ v_2\ v_3^{.625}\ v_4^{.625}\ v_5^{.625}\ v_6^{.625}$.

The general formula for total value can be written

VIII

$$V = v_1^{q_1}\ v_2^{q_2}\ v_3^{q_3} \ldots v_n^{q_n} \qquad \text{(B–1)}$$

where v_i is the ith IVAL factor and q_i is the ith correlation exponent. V is total value. This generalizes the IVAL utility function, equation 4–1.

Lastly, consider the more complex case where the subject matter expert judges that criteria 1 and 2 are strongly dependent, criteria 2 and 3 are weakly dependent, yet criteria 1 and 3 are independent. The question is how to determine the exponents, especially of criterion 2, when the criterion has two (or more) dependencies, and the dependencies are possibly not the same.

For example, let's say we have five independent criteria. Then total value is

$V = v_1\ v_2\ v_3\ v_4\ v_5$.

Were (only) criteria 1 and 2 strongly dependent, the total value would be

$V = v_1^{.625}\ v_2^{.625}\ v_3\ v_4\ v_5$.

Were (only) criteria 2 and 3 weakly dependent, the total value would be

$V = v_1\ v_2^{.875}\ v_3^{.875}\ v_4\ v_5$.

With both dependencies, the exponent of v_1 is clearly .625 and the exponent of v_3 is clearly .875. Moreover, clearly the exponent of v_2 must be some combination .625 and .875.

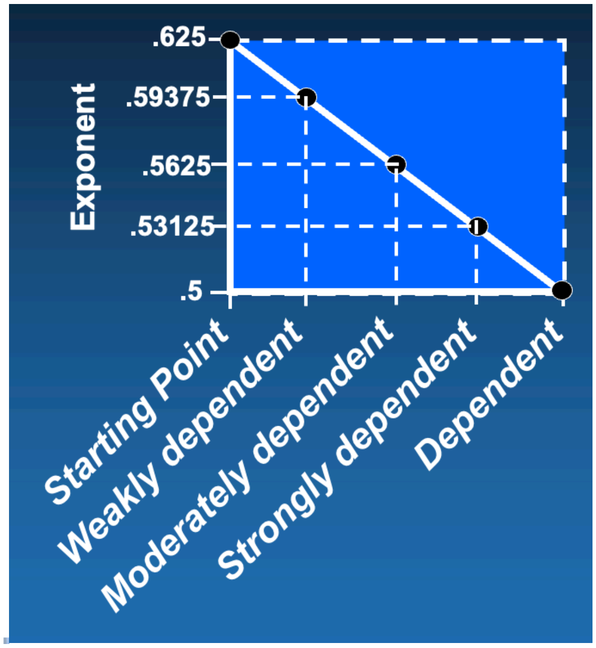


Figure B–3 Criterion With One Strong and One Other Dependency

To find the exponent, we simply put .625 at the top of the *y*-axis, as shown in Figure B-3. We place .5 at the bottom, as usual. We then split the differences to find the exponents for various other dependencies. In this case, our answer for v_2 having one strong dependency and one weak dependency is

$$V = v_1^{.625}\ v_2^{.59375}\ v_3^{.875}\ v_4\ v_5\ .$$

We could just as well have started by putting .875 at the top and then chosen the "strongly dependent" exponent from the resulting graph. The answer would be the same.

This last example is mostly overkill. It is given to provide guidance to a subject matter expert who might desire this level of correlation analysis. In most studies, nothing as complicated as this is considered. In many studies correlations are ignored altogether, with minimal consequence.

If a subject matter expert gives a criterion an even more complicated correlation, say correlation with not two (as assumed above), but with three different criteria, start with Figure B-1c to find the first correlation exponent. Use that exponent at the top of the y-axis to work out another exponent, and then use that exponent at the top of the y-axis to generate the final exponent.

C IVAL Treatment of Hierarchies

A lot of work can be required to handle hierarchies in WSM studies. Treatment of hierarchies in IVAL is trivial, as shown here.

Suppose Total Value is computed from three criteria, A – C. Further, suppose criterion A is composed of sub-criteria A_1 – A_3; B is composed of B_1 – B_3; C is composed of C_1 – C_3 .

We compute total value starting at the bottom of the pyramid:

$$V_A = v_{A_1}^{\,q_{A_1}}\, v_{A_2}^{\,q_{A_2}}\, v_{A_3}^{\,q_{A_3}},$$

$$V_B = v_{B_1}^{\,q_{B_1}}\, v_{B_2}^{\,q_{B_2}}\, v_{B_3}^{\,q_{B_3}},$$

$$V_C = v_{C_1}^{\,q_{C_1}}\, v_{C_2}^{\,q_{C_2}}\, v_{C_3}^{\,q_{C_3}},$$

and

$$V = V_A^{\,q_A}\, V_B^{\,q_B}\, V_C^{\,q_C},$$

where the “*q*” exponents, if any, are the correlation exponents developed in Appendix B.

D S-Curve Formulas

In this appendix, formulas for four generic S-curves are developed. There are four curves because we include curves that are either increasing or decreasing, and which are defined either on a closed interval [*a*, *b*] or a half-open interval [*a*, ∞).

The equations are developed in a way that gives the analyst control over the knees of the curve and the floor. The actual knees of a curve are determined through differentiation, but it is sufficient to approximate the knees by using the 20% and 80% points (on the *y*-axis). The 30% and 70% points, or any two other points, can be used instead. Formulas are provided in this appendix for the 20/80 case. The 30/70 and other cases are easy to generate following the patterns established.

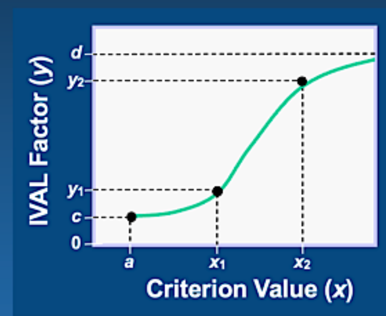
Note: For $y = c + (d - c)\left[1 - e^{-A(x-a)^B}\right]$ to represent an S-curve, we must have $A > 0$ and $B > 1$. It is okay to use other values but just be aware that the curve will then not be "S" shaped. The values for *A* and *B* presented in this appendix satisfy these requirements.

S-Curve Formulas

Case 1: Infinite Interval [a, ∞) Increasing Function

Givens

(x_1, y_1) = Point of Lower Diminishing Returns
(x_2, y_2) = Point of Upper Diminishing Returns
a = Lower Bound for Criterion Value
c = Lower Bound for IVAL Factor
d = Upper Bound for IVAL Factor



Solution

$$y = c + (d - c)\left[1 - e^{-A(x-a)^B}\right], \quad x \geq a$$

where

$$B = \frac{\ln[\ln(d-c) - \ln(d-y_2)] - \ln[\ln(d-c) - \ln(d-y_1)]}{\ln(x_2 - a) - \ln(x_1 - a)}$$

$$A = \frac{\ln(d-c) - \ln(d-y_1)}{(x_1 - a)^B}$$

20%-80% Inflection points:

$y_1 = c + .2\,(d - c)$
$y_2 = c + .8\,(d - c)$

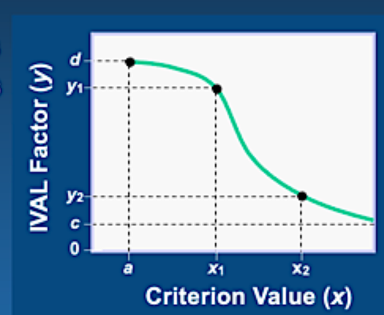
50

S-Curve Formulas

Case 2: Infinite Interval [a, ∞) Decreasing Function

Givens

(x_1, y_1) = Point of Upper Diminishing Returns
(x_2, y_2) = Point of Lower Diminishing Returns
a = Lower Bound for Criterion Value
c = Lower Bound for IVAL Factor
d = Upper Bound for IVAL Factor



Solution

$$y = c + (d - c)\, e^{-A(x-a)^B}, \quad x \geq a$$

where

$$B = \frac{\ln[\ln(d-c) - \ln(y_2 - c)] - \ln[\ln(d-c) - \ln(y_1 - c)]}{\ln - \ln(x_1 - a)(x_2 - a)}$$

$$A = \frac{\ln(d-c) - \ln(y_1 - c)}{(x_1 - a)^B}$$

20%-80% Inflection points:

$y_2 = c + .2\,(d - c)$
$y_1 = c + .8\,(d - c)$

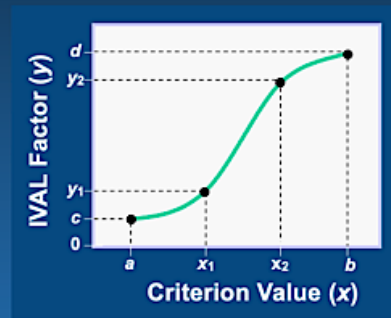
54

S-Curve Formulas

Case 3: Finite Interval [a, b] Increasing Function

Givens

- (x_1, y_1) = Point of Lower Diminishing Returns
- (x_2, y_2) = Point of Upper Diminishing Returns
- a = Lower Bound for Criterion Value
- b = Upper Bound for Criterion Value
- c = Lower Bound for IVAL Factor
- d = Upper Bound for IVAL Factor



Solution

$$y = c + (d - c)\left[1 - e^{-A\left[\frac{x-a}{b-x}\right]^B}\right], \quad a \le x \le b$$

where

$$B = \frac{\ln[\ln(d-c) - \ln(d-y_2)] - \ln[\ln(d-c) - \ln(d-y_1)]}{[\ln(b-x_1) - \ln(x_1-a)] - [\ln(b-x_2) - \ln(x_2-a)]}$$

$$A = \frac{\ln(d-c) - \ln(d-y_1)}{\left[\frac{x_1-a}{b-x_1}\right]^B}$$

20%-80% Inflection points:

$y_1 = c + .2\,(d - c)$

$y_2 = c + .8\,(d - c)$

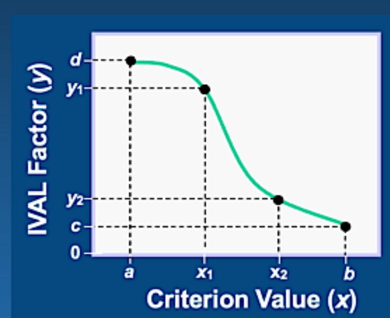
57

S-Curve Formulas

Case 4: Finite Interval [a, b] Decreasing Function

Givens

- (x_1, y_1) = Point of Upper Diminishing Returns
- (x_2, y_2) = Point of Lower Diminishing Returns
- *a* = Lower Bound for Criterion Value
- *b* = Upper Bound for Criterion Value
- *c* = Lower Bound for IVAL Factor
- *d* = Upper Bound for IVAL Factor



Solution

$$y = c + (d - c)\; e^{-A\left[\frac{x-a}{b-x}\right]^B}, \quad a \le x \le b$$

where

$$B = \frac{\ln[\ln(d-c) - \ln(y_2-c)] - \ln[\ln(d-c) - \ln(y_1-c)]}{[\ln(b-x_1) - \ln(x_1-a)] - [\ln(b-x_2) - \ln(x_2-a)]}$$

$$A = \frac{\ln(d-c) - \ln(y_1-c)}{\left[\frac{x_1-a}{b-x_1}\right]^B}$$

20%-80% Inflection points:

$y_2 = c + .2\,(d - c)$

$y_1 = c + .8\,(d - c)$

59

E Spline Curve Formulas

There is nothing sacred about using the formulas for S-curves developed in Appendix D. For example, a cubic spline is a very good alternative to an S-curve[5] when the *x*-interval is finite and the S-curve is symmetrical about a horizontal line. For curves from [0, 0] to [1, 1], cubic splines have the very nice properties of $f(0) = 0$, $f(1) = 1$, $f'(0) = 0$, and $f'(1) = 0$. However, these benefits come at the expense of relinquishing control of the inflection points. The equations are

$y = -2x^3 + 3x^2$ (increasing) and $y = 2x^3 - 3x^2 + 1$ (decreasing).

More generally, cubic splines defined from $[a,b] \rightarrow [c,d]$ as shown in Figure E-1 have the formulas

$$y = -2(d-c)\left(\frac{x-a}{b-a}\right)^3 + 3(d-c)\left(\frac{x-a}{b-a}\right)^2 + c$$

(increasing)

and

$$y = 2(d-c)\left(\frac{x-a}{b-a}\right)^3 - 3(d-c)\left(\frac{x-a}{b-a}\right)^2 + d$$

(decreasing).

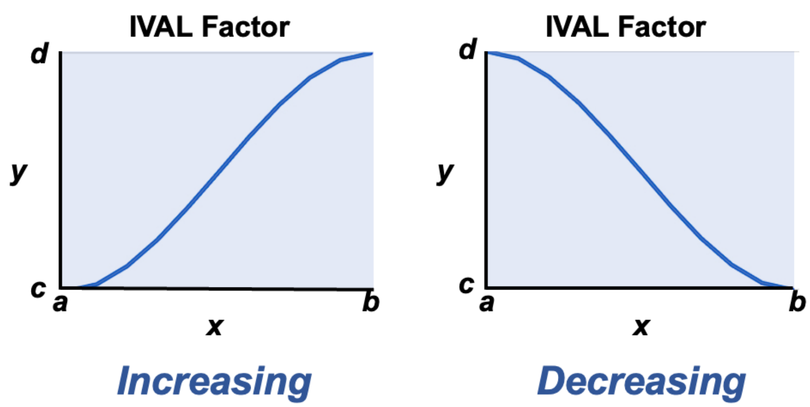


Figure E-1 Spline Curves

[5] Thanks to Dr. Charles B. McLane, retired Lockheed Martin Aerospace, for this observation.

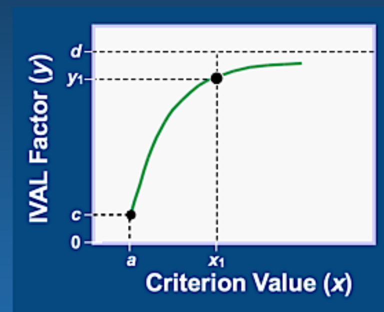
F Exponential Curve Formulas

Exponential Curve Formulas

Case 1: Infinite Interval [a, ∞)
Increasing Function

Givens

(x_1, y_1) = Point Diminishing Returns
a = Lower Bound for Criterion Value
c = Lower Bound for IVAL Factor
d = Upper Bound for IVAL Factor



Solution

$$y = d - (d - c)\, e^{-A\,(x - a)}, \quad x \geq a$$

where

$$A = \frac{\ln(d - c) - \ln(d - y_1)}{x_1 - a}$$

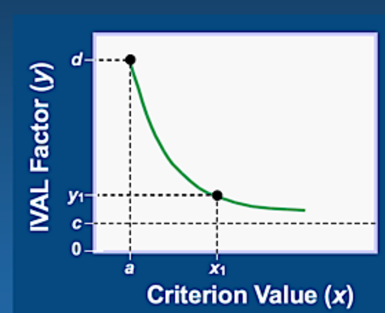
64

Exponential Curve Formulas

Case 2: Infinite Interval [a, ∞)
Decreasing Function

Givens

(x_1, y_1) = Point of Diminishing Returns
a = Lower Bound for Criterion Value
c = Lower Bound for IVAL Factor
d = Upper Bound for IVAL Factor



Solution

$$y = (d - c)\, e^{-A\,(x - a)}, \quad x \geq a$$

where

$$A = \frac{\ln(d - c) - \ln y_1}{x_1 - a}$$

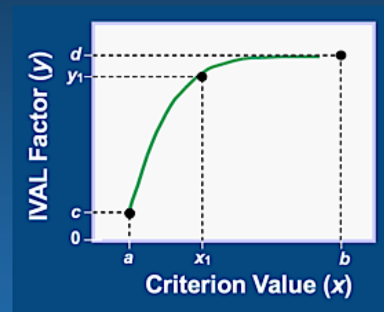
67

Exponential Curve Formulas

Case 3: Finite Interval [a, b] Increasing Function

Givens

(x_1, y_1) = Point of Diminishing Returns
a = Lower Bound for Criterion Value
b = Upper Bound for Criterion Value
c = Lower Bound for IVAL Factor
d = Upper Bound for IVAL Factor



Solution

$$y = d - (d - c)\left[\frac{b - x}{b - a}\right]^{-A}, \quad a \le x \le b$$

where

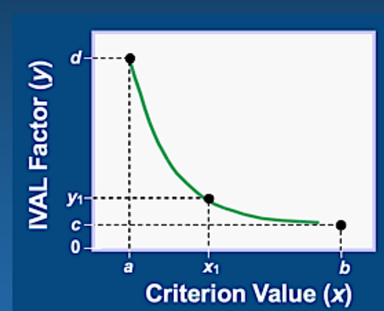
$$A = \frac{\ln(d - c) - \ln(d - y_1)}{\ln(b - x_1) - \ln(b - a)}$$

Exponential Curve Formulas

Case 4: Finite Interval [a, b] Decreasing Function

Givens

(x_1, y_1) = Point of Diminishing Returns
a = Lower Bound for Criterion Value
b = Upper Bound for Criterion Value
c = Lower Bound for IVAL Factor
d = Upper Bound for IVAL Factor



Solution

$$y = (d - c)\left[\frac{b - x}{b - a}\right]^{-A} + c, \quad a \le x \le b$$

where

$$A = \frac{\ln(d - c) - \ln(y_1 - c)}{\ln(b - x_1) - \ln(b - a)}$$

A04-12368120 73

G Fundamental Equation of Military Aircraft Combat Effectiveness[6]

The equation to be derived has four givens and one unknown. The givens are:

P_S = Probability of survival per sortie
SR = Sortie rate (number of sorties per day),
N = Number of days of combat, and
K_S = Number of target "kills" per sortie, given that the aircraft reaches the target area.

The unknown is:
K = Number of target kills in N days.

Let TO = Expected number of takeoffs by a single airplane in N days

A single aircraft takes off on its first sortie with probability 1. It has a probability P_S of surviving the sortie, so the probability of a 2nd takeoff is P_S. Its probability of surviving two sorties is P_S^2, which is the probability of a 3rd takeoff, etc.

The probability of takeoff on sortie SR, the last sortie of day 1, is P_S^{SR-1}. So, the probability of the 1st takeoff on day 2 is P_S^{SR}. Continuing this to the last sortie on day N, we see that TO is represented by a geometric series:

$$
\begin{array}{llllllll}
TO = 1 & + P_S & + P_S^{2} & + \ldots & + P_S^{SR-1} & \text{(Day 1)} \\
\quad + P_S^{SR} & + P_S^{SR+1} & + P_S^{SR+2} & + \ldots & + P_S^{2\,SR-1} & \text{(Day 2)} \\
\quad + P_S^{2\,SR} & + P_S^{2\,SR+1} & + P_S^{2\,SR+2} & + \ldots & + P_S^{3\,SR-1} & \text{(Day 3)} \\
\quad \vdots & & \vdots & & \vdots & \\
\quad + P_S^{(N-1)\,SR} & + P_S^{(N-1)\,SR+1} & + P_S^{(N-1)\,SR+2} & + \ldots & + P_S^{N\,SR-1} & \text{(Day } N\text{)}
\end{array}
$$

whose sum is $$TO = \frac{1 - P_S^{N \cdot SR}}{1 - P_S}$$

The target kills occur in the target area of the enemy, so we are interested in the number of trips to the target area. Presumably, the target area is the halfway point of the sortie, after which the airplane flies home. In the absence of information about the placement of enemy defenses, aircraft flight plan, or enemy firing doctrine, we assume that aircraft have 50% of their losses to threats during ingress to the target and 50% during egress. That is, we assume

$$P_{IN} = P_{EG}$$

where

[6] Developed in the 1970's by Gordon B Dobbins, Jr. of General Dynamics, Fort Worth Division.

P_{IN} = Probability of survival during INGRESS to the target area and

P_{EG} = Probability of survival during EGRESS from the target area.

Since probability of survival is multiplicative,

$P_S = P_{IN} P_{EG} = P_{IN}^2$,

and the ingress (half-way) probability of survival is

$$P_{IN} = \sqrt{P_S} \ .$$

The expected number of trips, T, to the target area in N days is simply the expected number of takeoffs times the probability of survival during ingress to the target area:

$$T = P_{IN}\, TO = \sqrt{P_S}\ \frac{1 - P_S^{\,N \cdot SR}}{1 - P_S} .$$

Finally, the Fundamental Equation of Military Aircraft Combat Effectives, the expected number of targets killed in N days, by a single aircraft is simply the expected number of trips to the target area times the expected number of targets killed per sortie:

$$K = T\, K_S = \sqrt{P_S}\ \frac{1 - P_S^{\,N \cdot SR}}{1 - P_S} K_S .$$

Note. This version of the Fundamental Equation rests on the assumption that the expected number of targets killed, K_S, remains the same, sortie after sortie, day after day. This is tantamount to assuming there is an infinite target set. In what is termed a “low density” target environment, this assumption would be changed.