JOURNAL OF SPACECRAFT AND ROCKETS
Vol. 48, No. 1, January–February 2011

Estimating Technology Readiness Level Coefficients

Edmund H. Conrow*
Redondo Beach, California 90278

DOI: 10.2514/1.46753

NASA and U. S. Department of Defense technology readiness level scales are in widespread use and provide inputs for a variety of systems engineering and project management functions. The common nine-level technology readiness level scales have ordinal coefficients that both limit their usefulness and introduce errors if mathematical operations are performed on the technology readiness level scale values (e.g., averaging). The analytic hierarchy process was used to estimate cardinal technology readiness level scale values. The average of the deviations between the ordinal coefficients (one through nine) and the analytic-hierarchy-process-estimated cardinal coefficients was 166%. A high-quality curve fit of the analytic-hierarchy-process-estimated coefficients was also developed (statistical coefficient of determination equal to 0.997), which permits the generation of noninteger technology readiness level values for use in mathematical operations.

Nomenclature

R^2 = statistical coefficient of determination
t = t value (tests the hypothesis that the specified regression coefficient is zero)

Introduction

THE purpose of this paper is to 1) present the technology readiness level (TRL) scale commonly used to evaluate hardware maturity on a variety of aerospace and nonaerospace applications, 2) discuss the fact that the TRL scale is ordinal and that common mathematical operations cannot be performed on the scale values, 3) provide a methodology to calibrate the scale coefficients to yield cardinal estimates for each TRL value, 4) provide the estimated TRL cardinal coefficients, 5) present the differences between the ordinal and resulting coefficients, 6) provide a curve fit of the cardinal coefficients to permit TRL estimates for fractional values ($1 \leq \text{TRL} \leq 9$), and 7) discuss limitations of the cardinal estimation methodology and the resulting coefficients. Of these seven items, all but the first represent new work that advances the state of the art and can be applied to a wide variety of both aerospace and nonaerospace programs.

"Technology readiness levels are a systematic metric/measurement system that supports assessments of the maturity of a particular technology and the consistent comparison of maturity between different types of technology."[†] TRLs have a wide variety of uses in aerospace systems engineering and project management, including permitting technology maturation tracking; helping to balance cost, performance, and schedule; and risk management (via the use of a condensed, reversed scale to represent the technology maturity component of probability of occurrence).

Scales involving TRLs or similar proxies for hardware maturity were developed and used by the U. S. Department of Defense (DOD) and NASA in the 1980s but more widely used following dissemination of John C. Mankins's 1995 paper.[†] Mankins's TRL definitions are given in Table 1.

TRL definitions have been applied and tailored to a variety of industries beyond NASA and uses other than hardware technology [1–3]. However, the basic idea associated with these other applications and uses remains the same as in Mankins's TRL scale.[†] For example, the DOD TRL scale definitions [3] are essentially the same as Mankins's NASA definitions,[†] with identical definitions for levels 1 through 6 and only minor definition wording changes for levels 7 through 9. The DOD TRL scale definitions are given in Table 2.

A variety of different scale types exist, including nominal, interval, ordinal, calibrated ordinal, estimative probability, and ratio [4–7]. Ordinal scales, as the name suggests, have levels that are monotonic and rank ordered and coefficient values that are only placeholders for the true cardinal values. Consider a three-level maturity scale with coefficients 1, 2, and 3. Here, the maturity of level2 > level1, the maturity of level3 > level1, and the maturity of level3 > level2. In this case, the scale is entirely nondecreasing. However, because coefficients 1, 2, and 3 are only placeholders, we can just as easily substitute coefficients A, B, and C, where C > B and B > A. (In fact, this representation is often preferable to using numbers to preclude people from performing mathematics on the resulting values.) Ordinal scales that have been calibrated using an additive utility function (as done with the TRL scale in this paper) are termed calibrated ordinal scales. Such scales are ratio scales of relative magnitudes [8]. However, the zero point associated with the resulting scale may not be meaningful if it 1) falls outside of the range of derived scale coefficients and 2) does not have physical significance. Mathematical operations between calibrated ordinal scale coefficients are generally possible, so long as it is recognized that the individual coefficients only have relative magnitudes, and the zero point may lack physical meaning (e.g., in contrast to the Kelvin and Rankine ratio temperature scales, where the zero point does have physical meaning).

The TRL scale (Tables 1 and 2) is a nine-level ordinal maturity scale related to hardware technology. The scale level values are inverted versus probability of occurrence ordinal scales commonly used in performing a risk analysis. For example, with a five-level ordinal probability maturity scale, the least mature level = 5, and the most mature level = 1. This is the opposite ordering of a TRL scale.

As previously mentioned, the ordinal TRL coefficients are only placeholders for the true cardinal coefficients, the scale levels are only rank ordered, and the interval values between scale levels are both potentially different and unknown. It is interesting to note that Mankins[†] never claimed that the TRL scale level values were cardinal nor that mathematical operations could be performed on the results. For example, an item rated as TRL = 8 [actual system completed and flight qualified through test and demonstration (ground or space)] is likely much more than twice as mature as another item rated as TRL = 4 (component and/or breadboard validation in laboratory environment). In addition, mathematical operations should never be

Received 17 August 2009; revision received 22 June 2010; accepted for publication 3 September 2010.

*Principal, Post Office Box 1125. Associate Fellow and Life Member AIAA.

[†]John C. Mankins, "Technology Readiness Levels," NASA Office of Space Access and Technology White Paper, April 1995; data available at www.hq.nasa.gov/office/codeq/trl/trl.pdf [retrieved 2009].

Table 1 NASA TRL values and corresponding definitions

TRL value	Definition
1	Basic principles observed and reported
2	Technology concept and/or application formulated
3	Analytical and experimental critical function and/or characteristic proof of concept
4	Component and/or breadboard validation in laboratory environment
5	Component and/or breadboard validation in relevant environment
6	System/subsystem model or prototype demonstration in a relevant environment (ground or space)
7	System prototype demonstration in a space environment
8	Actual system completed and flight qualified through test and demonstration (ground or space)
9	Actual system flight proven through successful mission operations

Table 2 DOD TRL values and corresponding definitions [3]

TRL Value	Definition
1	Basic principles observed and reported
2	Technology concept and/or application formulated
3	Analytical and experimental critical function and/or characteristic proof of concept
4	Component and/or breadboard validation in laboratory environment
5	Component and/or breadboard validation in relevant environment
6	System/subsystem model or prototype demonstration in a relevant environment
7	System prototype demonstration in an operational environment
8	Actual system completed and qualified through test and demonstration
9	Actual system proven through successful mission operations

performed on results obtained from ordinal scales, because very large errors can result [7]. (Irrefutable examples with errors of 600% or more between ordinal and cardinal scale coefficients can readily be generated; see [7], p. 262.) While the practice of performing mathematical operations on results from ordinal scales was widespread in the DOD and its contractors in the 1980s and 1990s, this erroneous practice was greatly curtailed by the late 1990s as a result of language included in the Risk Management Guide for DOD Acquisition [9] and other sources [6,10,11].

Attempts have been made to relate ordinal TRL values to potential risk, specifically cost growth and/or schedule slippage, in aerospace programs [12–18]. However, most of these discussions are subjective and provide only anecdotal information. Note that values derived from a TRL or similar scale are only related to a portion of the probability of the occurrence risk term and have nothing whatsoever to do with the consequence of occurrence risk term. Hence, TRL estimates will only be weakly correlated with risk. A simple illustration of this point follows.

The U. S. Government Accountability Office (GAO) attempted to relate TRL values to cost growth and schedule slippage. In the case of a commercial satellite solar array used on the HS 702 (later, the BS 702, when a portion of Hughes Space and Communications Company was sold to The Boeing Company), the GAO stated in July 1999 that the solar array was at TRL = 6 when launched, and that zero product development cost growth and schedule slippage existed [19]. However, following satellite launches in late 1999, 2000, and 2001, six satellites that included this solar array design rather rapidly and unexpectedly experienced problems associated with reduced onorbit power: so much so that insurance claims were filed for the six spacecraft totaling $1.04 billion.[‡] An inherent solar array design flaw led to a higher than expected outgassing of the solar cells, fogging of the solar array concentrator mirrors, reduced available power, and diminished satellite operational capability. While ground tests were performed on the arrays before launch, the tests did not identify the potential outgassing problem. The resulting cost and reputation impact of this design flaw was considerable; clearly, the technology was not as mature as had originally been estimated. (Even if the TRL value was correctly estimated, the risk level was effectively much higher than anticipated.) Incredibly, in a May 2007 presentation [20], the GAO still touted that the same commercial satellite solar array with TRL = 6 had zero product development cost growth and schedule slippage five or more years after onorbit problems had been observed, more than a billion dollars in insurance claims had been filed, and redesigns had been performed. In a communication from Frost and Sullivan in August 2004, it was stated,[†] "The disastrous introduction of the BS 702 serves as an example that there is considerable risk and no guarantees in being among the first to use a new satellite."

A quantitative analysis of TRL value versus schedule change (SC) was performed using a data set developed by Lee and Thomas [21]. Lee and Thomas estimated a cost-weighted TRL (WTRL) by taking the TRL of each available component and multiplying it by the component percent cost (component cost versus total program cost). The resulting sample contained WTRL, cost, and schedule-related information for 28 NASA space programs [21].

The SC was estimated from the initial schedule duration estimate (IDE, years) and the final total schedule duration (FTD, years) data as

$$SC = (FTD/IDE) \tag{1}$$

The author determined that a weak relationship existed when the WTRL was regressed against the SC with $R^2 \sim 0.26$ for several different equation forms (e.g., linear, modified exponential, and decay) that were curve fit to the full 28-program sample. Thus, only about one quarter of the variance in the dependent variable (SC) can be explained by the regression equation; the other three quarters is unexplained. (This falls to only about one fifth of the variance being explained in the dependent variable when the regression results are adjusted for degrees of freedom.) (Note that if the percent change in final versus initial schedule durations is regressed against the WRTL, the resulting R^2 values are effectively the same for each equation form as those estimated using the SC.) These results show that schedule slippage is primarily related (74 + % of the total variance) to factors other than the hardware development TRL value, and they are likely related to a number of other items that are not accounted for. Such factors may include, at a minimum, the following: 1) that the WTRL data are weighted by a cost fraction: this is unrelated to the technical and schedule dimensions and introduces another level of uncertainty into the data subsequently used in regressions between the WTRL and schedule slippage; 2) other maturity-related components outside of hardware technology (e.g., design, integration, manufacturing, and software systems); 3) trades made between cost, performance, schedule, and risk that affect buyer and seller

[‡]Data available at http://www.lr.tudelft.nl/live/pagina.jsp?id=ed766dad-dc20-44b8-9153-c89014139d0b&lang=en [retrieved 2010].

utility (e.g., preferences) but are not solely accounted for by the hardware TRL; and 4) the TRL scale values are ordinal, not cardinal, which may introduce an error when the WTRL is estimated and subsequently regressed.

In addition, the WTRL is estimated at the authority to proceed, which is when the project development contract was let.[§] However, the SC is a measure of the SC between the final, realized development schedule and the initial estimated schedule (which corresponds to IDE). Hence, SC contains two points in time, while WTRL is measured only at development initiation. The TRL for an item will change during the course of the program, both due to changes in maturity as well as refinements in the estimate of the TRL value. While the former cause will tend to increase the TRL value, the latter cause may increase or decrease the TRL value, depending upon the information and the state of the world. For example, in the commercial satellite solar array case, the actual TRL value (not the one reported by the GAO) was effectively lower (poorer) than the reported value. Here, the state of the world did not change, but the onorbit degradations revealed weaknesses in how the technology was implemented into the design that were not visible during ground-based testing. [Attempting to relate TRL values to schedule slippage does not take into account that programs are often started with an insufficient budget and/or schedule in order to meet the required level of performance or the potential changes in cost, performance, and schedule (and the reason for these changes) that occur during the course of the development program.]

The remainder of this paper provides five unique results. First, an approach that calibrates the TRL ordinal scale coefficients given in Tables 1 and 2, using the analytic hierarchy process (AHP), which uses an additive utility function [8] is given. The calibrated TRL scale coefficients are cardinal, and limited mathematical operations can be performed on them. Second, the resulting cardinal calibration coefficients for TRL values one through nine are provided: the results obtained are identical for both the DOD and NASA TRL scales. Third, the percent deviations between the ordinal and cardinal coefficients associated with TRL scale values one through nine are provided. These deviations are substantial. For example, using the calibrated coefficients, TRL = 8 is approximately six times more mature than TRL = 4 rather than a factor of two from simply and incorrectly taking the ratio of the ordinal (uncalibrated) scale coefficients ($8/4 = 2$). Fourth, a high-quality curve fit is presented for the nine calibrated TRL scale coefficients. This curve fit permits estimates to be made for noninteger TRL values, so long as the estimates are performed in the obvious, valid range ($1 \le \text{TRL} \le 9$). Fifth, the limitations of the resulting methodology used to calibrate the TRL coefficients and the coefficients themselves are discussed. The previously mentioned methodology used to generate results will be discussed, and it can be applied to TRL-like maturity scales other than the hardware technology scales evaluated in Tables 1 and 2.

Analytic Hierarchy Process Methodology and Implementation

TRL scales are ordinal with increasing maturity relative to the level number. For example, TRL = 9 is more mature than TRL = 8, TRL = 4 is more mature than TRL = 3, TRL = 2 is more mature than TRL = 1, and so on. The AHP was used to estimate the cardinal coefficient value for each ordinal TRL scale value. While this approach has been used to calibrate risk analysis probability of occurrence and consequence of occurrence scales for the past 16 years on a variety of programs, the application and results presented in this paper are the first known use of AHP (or related methodologies) to estimate calibrated TRL scale coefficients.

The relative magnitude scale values are estimated by expert opinion (e.g., knowledge and experience) making paired comparisons [8] in terms of how much more mature (ready) TRL scale level n is than TRL scale level $n-1$, and so on. For example, if TRL = 9 is judged to be twice as mature (or ready) as TRL = 8, then the pairwise comparison value for these two scale increments is equal to 2.0. For TRL = 9, there are eight pairwise comparisons made (TRL = 9 versus TRL = 8 through TRL = 9 versus TRL = 1). Similarly, seven pairwise comparisons are made for TRL = 8 (TRL = 8 versus TRL = 7 through TRL = 8 versus TRL = 1), and one pairwise comparison is made for TRL = 2 (TRL = 2 versus TRL = 1). Hence, the total number of pairwise comparisons that have to be made for the TRL scale are the sum of

$$8 + 7 + 6 + 5 + 4 + 3 + 2 + 1 = 36$$

The pairwise comparisons are used to populate a square matrix containing an equal number of rows and columns. Thus, in the nine-level TRL scale case, the matrix has 81 entries [36 entries in the upper triangle, 9 diagonal entries that are self comparisons (for which the values are 1.0, e.g., TRL = 9 versus TRL = 9), and 36 entries in the lower triangle]. The numbers in this matrix represent the maturity dominance of the criterion in the column heading (individual scale levels, such as TRL = 9) over the criterion in the row heading (again, individual scale levels; e.g., TRL = 5); see [8], p. 13. "Because a relative ratio scale is used, the matrix is reciprocal which means that the numbers, which are symmetric with respect to the diagonal, are inverses of one another, $a_{ij} = 1/a_{ji}$." (For example, if TRL = 9 is deemed to be twice as mature as TRL = 8, then the TRL = 8 is half as mature as TRL = 9.)

In general, $n(n-1)/2$ pairwise comparisons (or n items taken two at a time, $nC2$) are needed if n is the number of elements being compared above the diagonal populated with values of 1.0; see [8], p. 14. (Thus, with nine TRL scale levels, $9 \times 8/2 = 36$ pairwise comparisons are needed.)

"The pairwise comparisons are converted to a relative ratio scale by estimating the eigenvector of the square matrix," as described previously; see [8], p. 16. With the Expert Choice® software package ([8], p. 79)[¶]:

> Eigenvector computation is based upon the normalized row sums of the limiting power of a primitive matrix (and hence also of a positive matrix). To obtain this vector the matrix is raised to powers. Fast convergence is obtained by successively squaring the matrix. The row sums are calculated and normalized. The computation is stopped when the difference between these sums in two consecutive calculations of the power is smaller than a prescribed value.

For the nine-level TRL case, the eigenvector consolidates the 81 relative maturity (readiness) ratios of the matrix into nine measures of readiness. "This new scale is formally called the derived scale. It is an important property of this scale that the sum of the numbers is always 1.00" ([8], p. 15).

While the previous process may appear somewhat complicated, the TRL evaluation in this paper is a simplistic hierarchy with a single criteria (and, by extension, a single subcriteria). A more complex (but still simple) evaluation, for example, would be the calibration of cost, performance, and schedule consequence of occurrence ordinal scales, with a single criteria (consequence of occurrence), and three subcriteria (one each corresponding to cost, performance, and schedule consequences). Here, each subcriteria includes an n-level ordinal scale. An even more complex risk analysis evaluation would be the calibration of various hardware, software, and integration ordinal probability of occurrence scales, where each of these items (e.g., hardware) corresponds to an individual criteria, and each scale within the criteria (e.g., hardware: design/engineering, manufacturing, technology, threat) is a separate subcriteria. As in the last example, each subcriteria includes one or more n-level ordinal scales. David Graham, of the U.S. Air Force, likely first developed the approach of applying AHP to ordinal probability of occurrence scales in 1994 [7]. The first AHP application to a series of probability

[§]Dale Thomas, NASA Marshall Space Flight Center, personal communication, July 2009.

[¶]Numerous AHP estimation software exists and includes both commercial (e.g., Expert Choice, Decision Lens Suite™, and SuperDecisions™) and noncommercial packages.

Table 3 Estimated calibrated TRL coefficients (raw and adjusted)

Ordinal TRL values	AHP-estimated TRL values	AHP-estimated TRL values adjusted to 9.0 (TRL 9)
1	0.01	0.26
2	0.02	0.53
3	0.03	0.71
4	0.04	1.14
5	0.07	1.97
6	0.10	2.74
7	0.16	4.26
8	0.25	6.81
9	0.33	9.00

Table 4 Percent deviation between ordinal and AHP-estimated TRL coefficients

Ordinal TRL values	AHP-estimated adjusted TRL values	Percent deviation ordinal vs AHP estimated
1	0.26	284.6
2	0.53	277.4
3	0.71	322.5
4	1.14	250.9
5	1.97	153.8
6	2.74	119.0
7	4.26	64.3
8	6.81	17.5
9	9.00	0.0

scales was likely performed by Graham, Jason Dechoretz (MCR), and the author in 1994.

Results

The author has estimated the 36 pairwise comparisons for the TRL scales given in Tables 1 and 2. He then derived the resulting eigenvector (calibrated coefficients) using the Expert Choice software package and the methodology described previously. The calibrated ordinal coefficients for the TRL scales (Tables 1 and 2) are given in Table 3. The resulting eigenvector was identical to the eigenvector estimated using an online AHP calculator module, which is available to internet users free of charge.** The module will estimate the eigenvector for a single criteria and subcriteria having one to nine levels; hence, it is suitable for estimating†† coefficients for the nine-level TRL scale.

Both coefficients calculated directly from Expert Choice (raw) and those adjusted to 9.0 for the TRL 9 coefficient are provided. (In the former case, the coefficients resulting from the eigenvector result in a sum to 1.01, which is effectively 1.00, given rounding for the nine coefficients. In the latter case, each calibrated coefficient was multiplied by $9/0.33 = 27.27$ to obtain coefficients adjusted to 9.0 for TRL 9.)

The calibrated TRL coefficients presented in Table 3 have a small random noise component (as measured from Expert Choice AHP calculations and the online AHP calculator). Random noise is akin to inconsistency, ranging from 0 to 1.0, and is a "number closely related to the principal eigenvalue of the (square) matrix" ([8], p. 13), previously discussed. Inconsistency is "an adjustment needed to improve the consistency of the comparisons" ([8], p. 84). On a scale that ranges from zero to one, the overall inconsistency should be about 10% (0.1) (see [8], p. 84) and ideally smaller than this. While a consistency index value close to zero is desirable, it should not result from gaming the pairwise comparisons to drive the inconsistency down. The consistency index estimated by both Expert Choice and the online AHP calculator for the pairwise comparisons associated with the results in Table 3 was 0.01, which corresponds to a very small random noise term.‡‡ This very low consistency index resulted, in part, from Mankins's† clear and carefully defined wording used for each TRL scale level: ambiguous or subjective wording would have contributed to a larger consistency index.

Table 5 Difference and percent difference in adjacent TRL (*X* observed) values

TRL values *X* observed	AHP-estimated TRL values	Difference in AHP-estimated TRL values	Percent difference in AHP-estimated TRL values
1	0.26	N/A	N/A
2	0.53	0.27	N/A
3	0.71	0.18	−33.3
4	1.14	0.43	138.9
5	1.97	0.83	93.0
6	2.74	0.77	−7.2
7	4.26	1.52	97.4
8	6.81	2.55	67.8
9	9.00	2.19	−14.1

Percent deviations between the ordinal and cardinal coefficients associated with TRL scale levels 1 through 9 are given in Table 4. The average percent deviation was 166, and in each case, the AHP-estimated and adjusted cardinal coefficient was less than or equal to the ordinal TRL value. In fact, in five of the nine cases (TRL = 1 to 5), the deviation between the estimated and adjusted cardinal and ordinal coefficients was greater than 150%.

As shown in Table 4, the ratio of the AHP-estimated adjusted TRL values was substantially different than that for the ordinal TRL values. For example, using the AHP-estimated adjusted coefficients, TRL = 8 is approximately six times (5.97) more mature than TRL = 4 rather than a factor of two from simply and incorrectly taking the ratio of the ordinal (uncalibrated) scale coefficients ($8/4 = 2$). (The ratio of several other TRL values also showed substantial variations from equivalent ratios of the AHP-estimated adjusted TRL values versus ordinal values. But again, the raw TRL values are only ordinal placeholders, and performing mathematical operations on them will lead to erroneous results.)

The difference and percent difference in adjacent AHP TRL (*X* observed) values are given in Table 5. The largest differences in adjacent AHP-estimated TRL values are between TRL = 6 and 7, TRL = 7 and 8, and TRL = 8 and 9. However, the largest percent differences in adjacent AHP-estimated TRL values are between TRL = 3 and 4, TRL = 4 and 5, and TRL = 6 and 7, and they range from 93 to 139%. While it may appear that the highest payoff in improving maturity may occur for the TRL = 3 and 4, TRL = 4 and 5, and TRL = 6 and 7 cases, this is perhaps an overly simplistic perspective given the financial investments needed and the resulting timeline, and additional performance-related criteria will vary on a case-to-case basis.

A curve fit was then performed on the AHP-estimated adjusted TRL values, given in Tables 3 and 4. The purpose of this curve fit was to provide an approach to estimate noninteger TRL values (e.g., TRL = 5.6), so long as the estimates were performed in the obvious, valid range of $1 \leq \text{TRL} \leq 9$. (Having a method to estimate noninteger TRL values would provide, for example, the ability to more accurately estimate an average or WTRL value or TRL value standard deviation for a series of components, etc.)

Several criteria for accepting the curve fit were established before the process began. While these criteria are not inviolate, each is desirable to insure that the resulting curve fit is of high quality. First, the selected equation with derived coefficients had to be smooth, continuous, and without changes in the sign of the first and second

**This is available at http://www.isc.senshu-u.ac.jp/~thc0456/EAHP/EAHP_manu.html and http://www.isc.senshu-u.ac.jp/~thc0456/EAHP/AHPweb.html [retrieved 2010].

††There was no difference in the pairwise comparison scores between the NASA and DOD TRL scales; hence, the resulting AHP-estimated TRL values are the same for these two scales.

‡‡A consistency index of zero corresponds to no measurable inconsistency in the pairwise comparisons.

Table 6 Selected TRL curve fit equation form

Equation form and statistics	Value
Equation form	$Y = \alpha + \beta^* X^3$
R^2	0.997
Degrees of freedom adjusted R^2	0.996
Fit standard error	0.180
F value	2330.0

Table 7 Coefficient curve fit statistics for selected equation form

Parameter Statistics	Coefficient α	Coefficient β
Value	0.346	0.012
Standard error	0.082	0.0002
t value	4.22	48.27
95% lower bound	0.152	0.011
95% upper bound	0.540	0.013
Probability > $\|t\|$	0.004	0

derivatives over the range of $1 \leq \text{TRL} \leq 9$. Second, the resulting equation had to have as few coefficients as possible to retain relatively high degrees of freedom for the nine values used in the curve fit. Third, the t statistic of each curve fit coefficient had to be much greater than zero, and the corresponding probability level associated with the t statistic for each coefficient had to be much less than 0.05. Fourth, the 95% confidence interval lower and upper bounds for each estimated coefficient could not cross the zero value. Fifth, the resulting equation coefficient of determination (R^2) had to be fairly high (e.g., greater than 0.90 if possible), as should the R^2 adjusted for degrees of freedom.

More than 2000 equation forms were evaluated against the nine AHP-estimated adjusted coefficient values given in Tables 3 and 4. The selected equation form is

$$\text{TRL predicted} = \alpha + \beta^*(\text{TRL observed})^3 \tag{2}$$

The selected equation (in terms of X and Y) and summary statistics are given in Table 6.

Curve fit statistics for the individual coefficients of this equation are given in Table 7.

Predicted values of the AHP-estimated adjusted TRL values (Y predicted), along with the difference between the Y predicted and the AHP-estimated adjusted TRL values (Y residual) and the corresponding residual percentage (Y residual percentages) are given in Table 8. {Note that Y predicted is identical to TRL predicted [Eq. (2)], and X observed is the same as TRL observed [Eq. (2)]}. The average Y residual percentage from Table 8 is −1.7%: a relatively small value.

A graphical representation of the ordinal and estimated TRL coefficients is given in Fig. 1. The dashed line with diamonds in Fig. 1 represents the ordinal TRL coefficients. The slope of the line as drawn is 45 deg, but the y-axis values are an ordinal representation

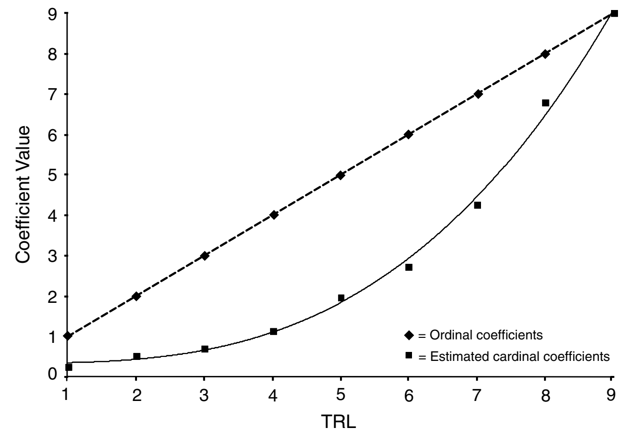

Fig. 1 Graphical representation of ordinal and estimated TRL coefficients.

and simply placeholders for true cardinal coefficient values. The upward sloping curve (with positive first and second derivatives) in Fig. 1 is the fitted curve of the form:

$$Y \text{ predicted} = 0.346 + 0.012 \times X^3$$

[Eq. (2) with coefficients from Table 7]. The square points about this curve are the AHP-estimated TRL adjusted values given in Tables 3, 4, and 8. The only time the selected equation form (Table 6) or corresponding curve (Fig. 1) should be considered is when noninteger estimates of TRL are needed. In such cases, note the residual magnitude and percent error given in Table 8 for each estimated TRL value. For example, for TRL = 5.6, the Y predicted value using Eq. (2) is 2.45, and a linear interpolation of residual values in Table 8 yields an estimated Y residual of −.072.

From Tables 4 and 8, the average of the nine ordinal coefficients is 5.0. Note, this is effectively an invalid result, since the average of ordinal numbers is generally meaningless. However, this result is presented to contrast it with the average of the nine AHP-estimated TRL values, which is 3.1, or about a 60% difference. Hence, using ordinal TRL coefficients may overestimate the technical maturity of an item when mathematical operations are performed on values corresponding to TRL levels 1 through 8. This is graphically illustrated in Fig. 1, where the AHP-estimated TRL coefficient values are less than the corresponding ordinal values for all but TRL = 9 (which is the normalization point to fix the AHP-estimated values to the ordinal values).

Use the following procedure to perform mathematical operations on TRL values. First, convert integer TRL scores (e.g., 5) to TRL-estimated values adjusted to 9.0, with the results given in Tables 3 and 4. For TRL scores involving fractional values (e.g., 5.6), use Eq. (2) to estimate TRL-predicted values. Second, perform the desired mathematical operations on the estimated and/or predicted values. (Only valid mathematical operations should be considered.) Third, convert the resulting value(s) back to observed TRL using Eq. (3), which is derived from Eq. (2):

Table 8 Selected equation predicted TRL values and residuals

TRL values, X observed	AHP-estimated adjusted TRL values	TRL values, Y predicted	Y Predicted: AHP-estimated Y residual	Y Predicted: AHP-estimated Y residual, %
1	0.26	0.36	−0.10	−37.8
2	0.53	0.44	0.09	16.5
3	0.71	0.67	0.04	5.6
4	1.14	1.11	0.03	2.2
5	1.97	1.85	0.12	6.3
6	2.74	2.94	−0.20	−7.3
7	4.26	4.46	−0.20	−4.8
8	6.81	6.49	0.32	4.7
9	9.00	9.10	−0.10	−1.1

$$\text{TRL observed} = [(\text{TRL predicted} - \alpha)/\beta]^{(1/3)} \quad (3)$$

Any item that is a medium or high risk should always be carried independently of results generated by mathematical operations on a series of risks or roll ups, for example, to higher work breakdown structure levels. Otherwise, risks can be overlooked that can reappear later in the project as problems that are very difficult to handle ([7], p. 295). Similarly, when using ordinal or AHP-estimated TRLs, unadjusted TRL values should be carried for each item, and values below a particular predetermined threshold (e.g., TRL = 6) should be reported for such items, no matter what mathematical operations or roll ups are performed. (For example, if the average TRL is estimated for n different hardware components in a given subsystem using the AHP-estimated TRL values, the lowest component TRL level value present in the subsystem should be separately reported.) In addition, the lowest TRL value should be carried forward without modification for each successive level of integration. For example, if a system is composed of components and subsystems, TRL values should first be estimated for appropriate components (e.g., those that are developmental items and not proven off-the-shelf items). The subsequent subsystem TRL is thus the lowest TRL value of any component contained in the subsystem plus the equivalent of TRL for integrating the components into the subsystem and other appropriate considerations (e.g., design/engineering, manufacturing, support, and threat). Similarly, the system-level TRL is the lowest of the subsystem-level TRL scores coupled with the equivalent of the TRL for integrating the subsystems into the system, plus other appropriate considerations. All components, subsystems, and systems that have TRL values below the minimum threshold required by the program should be identified and documented, and appropriate action should be taken to evaluate and alleviate the shortfalls [22].

The same methodology used to generate results given in Tables 3–8 and Fig. 1 can be applied to TRL-like maturity scales, other than the hardware scales evaluated in Tables 1 and 2. In fact, the methodology application is direct and easy for simulation, software, biomedical, manufacturing, and related maturity-based applications that currently exist [1–3]. In these particular cases, the nine-level scales will also have 36 pairwise comparisons that must be evaluated. (The number of comparisons is $nC2$, where n is the number of scale levels.) It is strongly recommended that an independent analysis to estimate scale coefficients be performed for each TRL-like scale, rather than simply using the coefficient and curve fit results for the hardware maturity TRL scale presented here.

Conclusions

Several enhancements have been developed and described in this paper for using widely distributed TRL values and similar maturity-based readiness level scales. These enhancements include 1) application of AHP methodology to provide estimates of cardinal TRL coefficients derived from the original ordinal TRL scale, 2) the resulting cardinal TRL values, and 3) a high-quality curve fit to permit estimation of noninteger TRL values.

The TRL scale developed by Mankins† and similar scales developed for other applications (e.g., software and manufacturing) can be very helpful in estimating the level of maturity (or readiness) for selected items. However, TRL and similar readiness (maturity based) scales are only weakly correlated with risk, because they are unrelated to the consequence of occurrence and only partially represent the probability of the occurrence term. TRL scale levels are also ordinal, not cardinal, and the user has no cardinal indication of how much more mature (ready) one TRL value is versus another. (For example, TRL = 8 is not simply two times more mature than TRL = 4.) Similarly, mathematical operations cannot be performed on results obtained from ordinal scales, such as the TRL scale, because the ordinal scale (coefficient) values are only placeholders for actual but unknown cardinal coefficients, and the resulting error magnitudes can be quite large. Given limitations of the ordinal TRL scale, a procedure was developed to estimate corresponding cardinal coefficients for each TRL value. Mathematical operations, such as computing the average or weighted average, can then be performed on the cardinal coefficients without introducing errors resulting from the use of ordinal scale coefficients.

The AHP was used to determine cardinal coefficients for the nine TRL scale levels from 36 estimated pairwise comparisons. (The 36 pairwise comparisons correspond to nine items taken two at a time, or 9C2.) When the estimated coefficient for TRL = 9 is constrained to 9.00, the other eight coefficients are all less than their corresponding ordinal level values. Mathematical operations can be performed on the AHP-estimated coefficients, because they are cardinal (albeit relative), not ordinal. Thus, the ratio of one TRL value to another (e.g., TRL = 8 to TRL = 4, which is 5.96) provides an estimate of the relative level of maturity (readiness) between the estimates, whereas any such ratios between the original ordinal scale values are practically meaningless and can lead to substantial errors. Similarly, the differences between adjacent scale level coefficients provide an estimate of how much more mature one TRL level is versus another. The resulting ratios and differences can provide the user with information to better allocate resources in order to change (typically increase) an item's TRL value versus other items with typically higher TRL values.

The AHP-estimated cardinal coefficients also permit mathematical operations (e.g., simple or weighted averages and standard deviation) to be performed on TRL values. This provides the user with a method of comparing TRL values between two different units (e.g., subsystems or systems), thus contrasting the maturity of the units themselves. (From an overall maturity and potential contractual perspective, the minimum TRL value should also estimated and reported for each unit.)

A very high-quality curve fit (involving a simple third-order relationship) of the AHP-estimated cardinal coefficients was developed ($R^2 = 0.997$) that allows estimates to be made and subsequent computations to be performed with fractional TRL values. The curve fit permits the direct estimation of noninteger TRLs (e.g., 5.6) for which the resulting values can be used in mathematical operations without introducing errors associated with ordinal TRL coefficients. The mathematical results can, in some cases (e.g., average or weighted average), be converted back to observed TRL estimates using an inverse relationship derived from the curve fit.

While the AHP procedure was specifically applied to the TRL scale in this paper, it can also be directly applied to other ordinal scales, such as manufacturing and software readiness levels to estimate corresponding cardinal coefficients for these scale values. As mentioned, this can be done for both readiness level scales as well as a variety of other scales used in risk analysis and other applications.

Acknowledgments

I gratefully acknowledge the support of Rozann Saaty of RWS publications for granting permission to include excerpts from *The Fundamentals of Decision Making and Priority Theory with the AHP* in this paper. I also gratefully acknowledge the support of Dale Thomas of NASA Marshall Space Flight Center, who provided the NASA schedule and technology readiness level (TRL)-related data used for regressions of weighted TRL versus schedule change.

References

[1] Clay, R. L., Marburger, S. J., Shneider, M. S., and Trucano, T. G., "Modeling and Simulation Technology Readiness Levels," Sandia National Labs. Rept. SAND2007-0570, Albuquerque, NM, Jan. 2007. Unlimited release. Available to Dept. of Energy and Dept. of Energy contractors through the U. S. Dept. of Energy. Available to the public through the U. S. Dept. of Commerce, National Technical Information Service, Springfield, VA.

[2] Hantos, P., "Software Technology Readiness Assessments-Managing Technology Risks in Space System Acquisitions," The Aerospace Corporation, Rept. TOR-2008(8550)-8033, 16 June 2008. Unlimited release. Available to Dept. of Defense and Dept. of Defense contractors through the Defense Technical Information Center (ADA484614). Available to the public through the U. S. Dept. of Commerce, National

Technical Information Service, Springfield, VA.
[3] "Technology Readiness Assessment (TRA) Deskbook," U. S. Dept. of Defense, May 2005. Approved for public release, distribution unlimited. This document also contains TRL scales for biomedical, manufacturing, and software applications. Available to Dept. of Defense and Dept. of Defense contractors through the Defense Technical Information Center (ADA438933). Available to the public through the U. S. Dept. of Commerce, National Technical Information Service, Springfield, VA.
[4] Stevens, S. S., "On the Theory of Scales of Measurement," *Science*, Vol. 103, No. 2684, June 1946, pp. 677–680. doi:10.1126/science.103.2684.677
[5] Velleman, P. F., and Wilkinson, L., "Nominal, Ordinal, Interval, and Ratio Typologies Are Misleading," *The American Statistician*, Vol. 47, No. 1, Feb. 1993, pp. 65–72. doi:10.2307/2684788
[6] Conrow, E. H., "The Use of Ordinal Risk Scales in Defense Systems Engineering," *Proceedings of the 1995 Acquisition Research Symposium*, Defense Systems Management College, Ft. Belvoir, VA, June 1995, pp. 455–463.
[7] Conrow, E. H., *Effective Risk Management: Some Keys to Success*, 2nd ed., AIAA, Reston, VA, 2003, Chap. 6, Appendix H.
[8] Saaty, T. L., "*The Fundamentals of Decision Making and Priority Theory with the Analytic Hierarchy Process*," Vol. 6, revised ed., AHP Series, RWS Publ., Pittsburgh, 2000, Chaps. 1, 3.
[9] "*Risk Management Guide for DoD Acquisition*," 2nd ed., Defense Systems Management College, Ft. Belvoir, VA, March 1998, pp. 16–17.
[10] Pariseau, R., and Oswalt, I., "Using Data Types and Scales for Analysis and Decision Making," *Acquisition Review Quarterly*, Vol. 1, No. 2, Spring 1994, pp. 146–152.
[11] Kerzner, H., *Project Management: A Systems Approach to Planning, Scheduling, and Controlling*, 7th ed., Wiley, New York, 2001, pp. 924–927.
[12] "Defense Acquisitions: Improvements Needed in Space Systems Acquisition Management Policy," U.S. Government Accountability Office, Report GAO-03-1073, Sept. 2003.
[13] "Defense Acquisitions: Assessments of Selected Weapon Programs," U.S. Government Accountability Office, Rept. GAO-07-406SP, March 2007.
[14] "NASA: Implementing a Knowledge-Based Acquisition Framework Could Lead to Better Investment Decisions and Projects Outcomes," U.S. Government Accountability Office, Rept. GAO-06-218, Dec. 2005.
[15] "Defense Acquisitions: Assessments of Selected Major Weapon Programs," U.S. Government Accountability Office Repts. GAO-05-301, March 2005.
[16] "Defense Acquisitions: Assessments of Selected Major Weapon Programs," U.S. Government Accountability Office Rept. GAO-06-391, March 2006.
[17] "Defense Acquisitions: Assessments of Selected Weapon Programs," U.S. Government Accountability Office, Rept. GAO-08-467SP, March 2008.
[18] "Defense Acquisitions: Assessments of Selected Weapon Programs," U.S. Government Accountability Office, Report GAO-09-326SP, March 2009.
[19] "Best Practices: Better Management of Technology Development Can Improve Weapon Systems Outcome," U.S. Government Accountability Office, Rept. GAO/NSIAD-99-162, July 1999, p. 27.
[20] Sullivan, M. J., "GAO Review of Technology Transition Practices," *Fourth Annual Acquisition Research Symposium*, 16 May 2007, p. 22.
[21] Lee, T. S., and Thomas, L. D., "Cost Growth Models for NASA's Programs," *Journal of Probability and Statistical Science*, Vol. 1, No. 2, Aug. 2003, pp. 265–279.
[22] "NASA Systems Engineering Handbook," NASA SP-6105, Rev. 1, Dec. 2007, p. 295.

O. de Weck
Associate Editor