NPS-AM-12-C9P21R02-086

EXCERPT FROM THE

PROCEEDINGS

OF THE

NINTH ANNUAL ACQUISITION RESEARCH SYMPOSIUM THURSDAY SESSIONS VOLUME II

An Analysis of TRL-Based Cost and Schedule Models

C. Robert Kenley and Bernard El-Khoury
Massachusetts Institute of Technology

Published April 30, 2012

An Analysis of TRL-Based Cost and Schedule Models

C. Robert Kenley—Kenley holds a PhD and an MS in engineering-economic systems from Stanford, an MS in statistics from Purdue, and an SB in management from MIT. He has over 30 years experience in systems engineering in advanced socio-technical systems research and development in the aerospace, nuclear, and healthcare domains. He has provided clients with insight and understanding of systems problems as an independent consultant, and he currently is a research associate at MIT. He is a published author of several papers and journal articles in the fields of systems engineering, decision analysis, Bayesian probability networks, and applied meteorology.

Bernard El-Khoury—El-Khoury is a graduate student at the Massachusetts Institute of Technology where he is pursuing a master's degree in the technology and policy program. He is simultaneously pursuing a master's degree in industrial engineering at Ecole Centrale Paris, where he also received a BS in engineering. His research interests are in technology cost and schedule forecasting, and power systems modeling.

Abstract

The GAO's, NASA's, and the DoD's adoption of the technology readiness level (TRL) scale to improve technology management has led to the emergence of many TRL-based models that are used to monitor technology maturation, mitigate technology program risk, characterize TRL transition times, or model schedule and cost risk for individual technologies, as well as technology systems and portfolios. In the first part of this paper, we develop a theoretical framework to classify those models based on the (often implicit) assumptions they make; we then propose modifications and alternative models to make full use of the assumptions. In the second part, we depart from those assumptions and present a new decision-based framework for cost and schedule joint modeling.

Introduction

The technology readiness level (TRL) is a discrete scale used by U.S. acquisition agencies to assess the maturity of evolving technologies prior to incorporating those technologies into a system or subsystem. A low TRL (1–2) indicates a technology that is still at a basic research level, while a high TRL (8–9) indicates a technology at the final system level being already incorporated into an operational system. Table 1 presents the formal definitions of the NASA TRL levels defined by Mankins (1995).

Table 1. NASA TRL Scale Definitions

TRL	NASA TRL 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

Government agencies face major challenges when it comes to developing new technologies. For instance, the Department of Defense (DoD) (1) develops a very large portfolio of technologies, (2) develops high complexity system technologies, (3) manages a budget of several hundred billion dollars (GAO, 2009) in a monopsonistic contracting environment with limited market competition, (4) suffers from frequent design changes due to changes in requirements, and (5) is under constant pressure to accelerate the development of technologies required for pressing national security issues.

In evaluating the DoD's performance, the Government Accountability Office (GAO) concluded, "Maturing new technology before it is included on a product is perhaps the most important determinant of the success of the eventual product—or weapon system" (GAO, 1999, p. 12). The GAO (1999) also encouraged the use of "a disciplined and knowledge-based approach of assessing technology maturity, such as TRLs, DoD-wide" (p. 7).

The TRL scale gained prominence in technology management and the defense acquisition community, and a literature soon developed on the use of TRL to monitor technology maturation, to mitigate technology program risk, to characterize TRL transition times, and to model schedule and cost risk for individual technologies, as well as technology systems and portfolios.

Those approaches do not depart from the same assumptions. Although some try to find a theoretical foundation for their models, others implicitly make the assumptions and found their models on the usefulness and robustness of the results. Developing a common framework for TRL-based models can help us better understand the underlying assumptions of those models and better critique them or use them to their full potential.

These models try to study the relationship between TRL and key program management variables, such as cost, schedule, performance, and program risk. We refer to these variables more generally as *maturity variables* because we try to model their evolution (along with the related uncertainties) as the project matures (i.e., as TRL increases). We use maturity variables in an attempt to generalize the results and maximize the scope of the models whenever possible. *Transition maturity variables* are maturity variables defined for one TRL transition (e.g., the TRL 1-to-2 transition time, or the cost of transition from TRL 5 to TRL 7, are such variables).

We provide a theoretical framework of the assumptions generally made to make TRL-based cost and schedule modeling, classify existing TRL-based models within this framework, and propose modifications and alternative models to make full use of the assumptions. We look at one of those proposed models in more detail as a method that integrates both cost and schedule through the use of a decision-modeling framework. In the Part I: Theoretical Framework and Currently Available Models section, we explain the theoretical framework, and for each assumption level of the framework, we list the currently available models and discuss alternatives or suggestions whenever possible. In the Part II: A New Framework for Cost and Schedule Joint Modeling section, we depart from those assumptions and present a new framework for cost and schedule joint modeling.

Part I: Theoretical Framework and Currently Available Models

Figure 1 shows the four levels of increasingly strong assumptions made in TRL-based models of maturity variables.

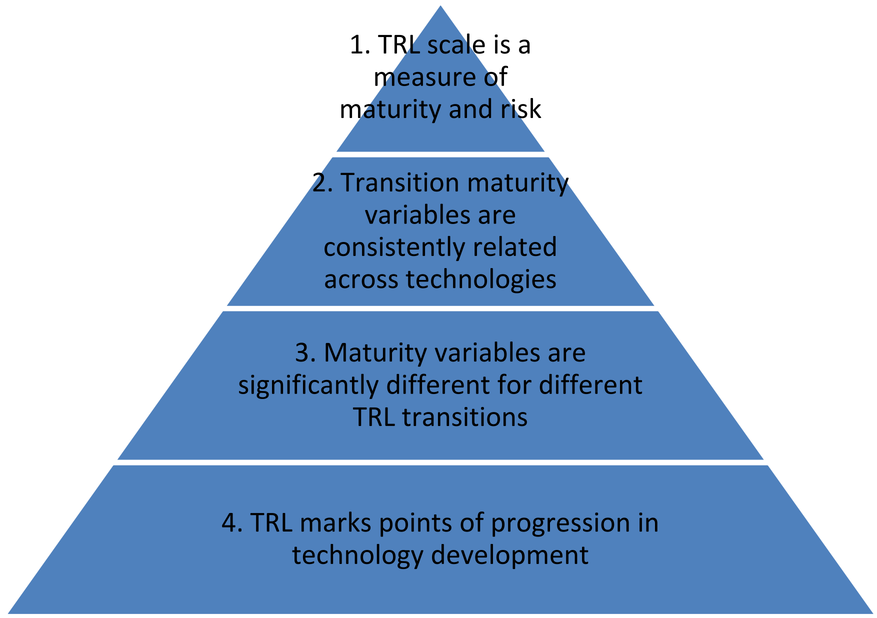

Figure 1. Assumption Levels for the Framework

Level 1 Assumption

The Level 1 assumption is trivial and is hard to contest; it makes use of the ordinality property of the TRL scale. Its direct consequence is that the higher the TRLs, the smaller the remaining overall uncertainty in maturity variables. A project at TRL 3 is subject to risks (cost, schedule, technology) on transitions from TRL 3 to TRL 9, while a project at TRL 7 is subject only to risks on transitions from TRL 7 to TRL 9. Figure 2 depicts the reduction in uncertainty and particularly identifies the TRL 6-to-7 transition as the most important step in reducing the risk of achieving a product launch. Although the GAO identified this for product launch or programmatic risk, this reduction in uncertainty is also true for other maturity variables. The range of uncertainty for cost, schedule, and performance is also reduced as the TRL progresses.

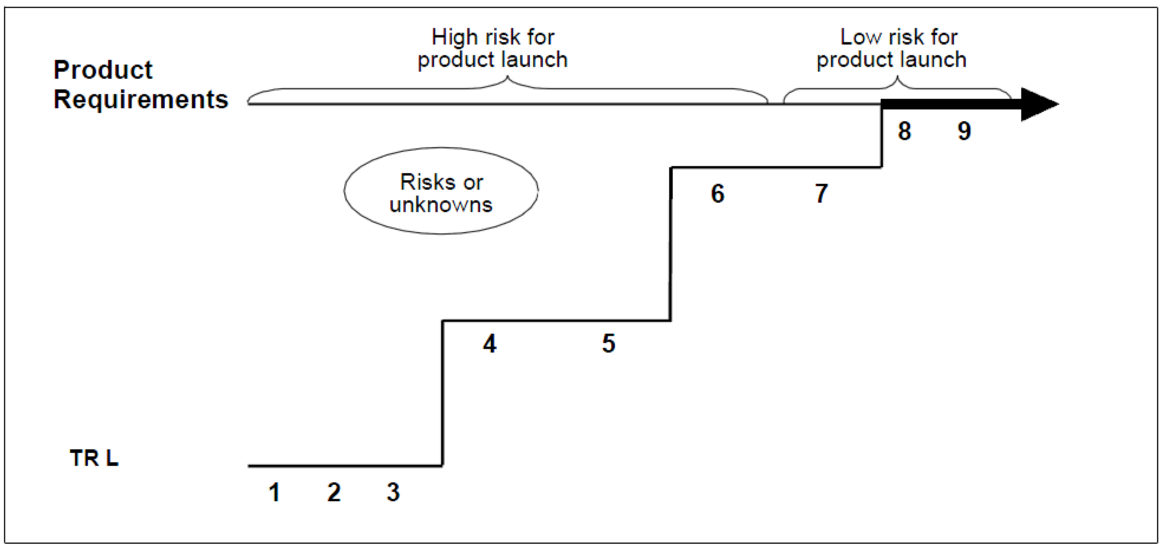

Figure 2. Programmatic Risk as a Function of TRL
(GAO, 1999, p. 24)

Level 2 Assumption

The Level 2 assumption is also a weak assumption. It stipulates that when we look at one technology transition, the maturity variables have a probability distribution different from other TRL transitions. Theoretically, this assumption is supported by the very design of the TRL scale. Because each transition in the scale corresponds to a well-defined action common to any technology development, we expect each of those transitions to share common properties across technologies and thus be significantly differentiable from other transitions. For example, the TRL 1–2 transition that happens when “an application of the basic principles is found” is different from the TRL 2–3 transition that corresponds to “going from paper to lab experiments,” which itself is different from the TRL 6–7 transition that happens when “the prototype is tested in the real environment.” The descriptions of those transition processes are clear enough to expect their properties to be different from each other while being coherent among similar transitions.

We performed analysis of variance (ANOVA) in Table 2, which shows for example that TRL transitions 1–2 and 2–3 are both different from transition 1–3 at a 95% confidence level. Hence, we statistically lose information when we reduce the TRL scale to fewer than the nine defined stages. The ANOVA uses data from a case study done by the Systems Analysis Branch at NASA’s Langley research center, looking at typical times aeronautical technologies take to mature (Peisen & Schulz, 1999). The data was collected through interviews with NASA personnel. The equality of means hypothesis was rejected here, although the dataset contained transition times for only 18 technologies; the power of the test can be significantly improved if more data were available.

Table 2. ANOVA Analysis on NASA SAIC Data Comparing Transitions 1–2 and 2–3 to Transition 1–3

ANOVA Summary					
Total Sample Size	38				
Grand Mean	0.3221				
Pooled Std Dev	0.8973				
Pooled Variance	0.8052				
Number of Samples	2				
Confidence Level	95.00%				
	12	13			
ANOVA Sample Stats	Data Set #1	Data Set #1			
Sample Size	19	19			
Sample Mean	0.0157	0.6285			
Sample Std Dev	0.9041	0.8905			
Sample Variance	0.8174	0.7930			
Pooling Weight	0.5000	0.5000			
OneWay ANOVA Table	Sum of Squares	Degrees of Freedom	Mean Squares	F-Ratio	p-Value
Between Variation	3.5682	1	3.5682	4.4313	0.0423
Within Variation	28.9884	36	0.8052		
Total Variation	32.5566	37			
	Difference	No Correction			
Confidence Interval Tests	of Means	Lower	Upper		
12-13	-0.6129	**-1.203316939**	**-0.022406713**		

ANOVA Summary					
Total Sample Size	38				
Grand Mean	0.1755				
Pooled Std Dev	0.9275				
Pooled Variance	0.8603				
Number of Samples	2				
Confidence Level	95.00%				
	13	23			
ANOVA Sample Stats	Data Set #1	Data Set #1			
Sample Size	19	19			
Sample Mean	0.6285	-0.2775			
Sample Std Dev	0.8905	0.9631			
Sample Variance	0.7930	0.9276			
Pooling Weight	0.5000	0.5000			
OneWay ANOVA Table	Sum of Squares	Degrees of Freedom	Mean Squares	F-Ratio	p-Value
Between Variation	7.7985	1	7.7985	9.0645	0.0047
Within Variation	30.9721	36	0.8603		
Total Variation	38.7706	37			
	Difference	No Correction			
Confidence Interval Tests	of Means	Lower	Upper		
13-23	0.9060	**0.295709129**	**1.516356742**		

Dubos, Saleh, and Braun (2010) implicitly apply our Level 2 assumption when they look at the distribution of every TRL transition time. They found that TRLs' transition times have lognormal distributions and used that to propose average estimators and confidence intervals.

Because TRL data is typically scarce, and because of its high skewness, we should evaluate it using a more robust measure, such as the median instead of the average; however, it is harder to generate median estimators and confidence intervals. In such cases of asymmetrical data, typically when the data is truncated and skewed, which is the case for all maturity variables, Mooney and Duval (1993) recommend the bootstrap over classic parametric tests. The bootstrap is a resampling technique that generates an empirical distribution of the required statistic (the median in this case; Efron & Tibshirani, 1993). It is especially useful to make inference on small samples and access the (little) information contained in the sample without making parametric assumptions. To implement the median bootstrap, we resampled with replacement a large number of times; then we took the median of each of those resamples. From the resulting histogram of the medians, we obtained an empirical distribution of the median and used it to generate confidence intervals.

We used the iterated smoothed bootstrap for the median and mean of TRL transition times; then we saved the resulting mean and median distributions in two Excel user-defined functions, as shown in Figure 3. A user can easily access different estimates, standard deviations, and confidence intervals by typing the starting TRL, the ending TRL, and the required confidence level. We also created a bootstrap Excel function that generates the bootstrap distribution based on a data sample input.

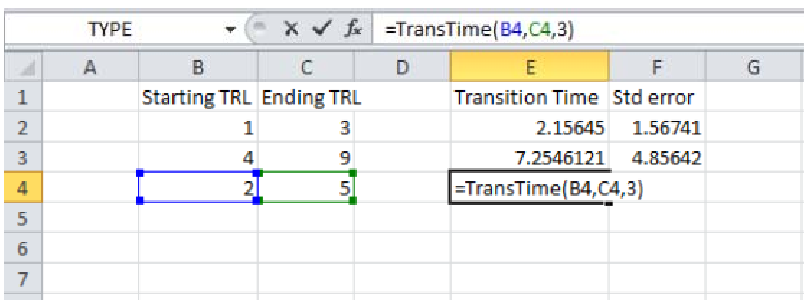

Figure 3. Snapshot of the Transition Time User-Defined Function in Excel

In summary, this Level 2 assumption is made whenever we want to study the statistics of maturity variables on each of the TRL transitions. Empirical data appears to confirm that the introduction of those TRL divisions increase the amount of significant information on the maturity variables. However, because the available data sets are small and skewed, the median bootstrap is a good method for making inferences on those transition-based maturity variables.

Level 3 Assumption

The Level 3 assumption is a stronger one. One of its major implications is that for one technology, we can use early transition information to make inferences on possible values of later transition variables.

One important question that we can answer under this assumption is this: If a technology is already developing quickly (or cheaply), does this mean that it is more likely to continue developing quickly (or cheaply)? The theoretical argument here might be that some technologies have overall properties that are independent of maturity or the TRL, such as, the technology is intrinsically harder to develop, the contractor is more expensive than the average, or the contract terms encourage late delivery. Another argument is that the cost risk or schedule risk during technology development always evolves in the same manner independently of the technology. For example, we might assume that cost risk always increases throughout the project, and hence initial values of cost overruns will give us an idea about future cost overruns.

Our analysis of empirical evidence using the NASA SAIC schedule data supports this assumption. Figure 4 shows log-transition times for all 18 technologies. By looking at each technology (each curve), we can see some similar trends before TRL 6, many values of zero (one year) for transition from TRL 2 to 3, a general tendency to increase after TRL 2–3 or after TRL 3–4, and then a general tendency to decrease after TRL 4–5; however, those trends seem to disappear after TRL 6–7 with either a constant value or a high amplitude oscillation, before converging at the last transition towards a value close to zero.

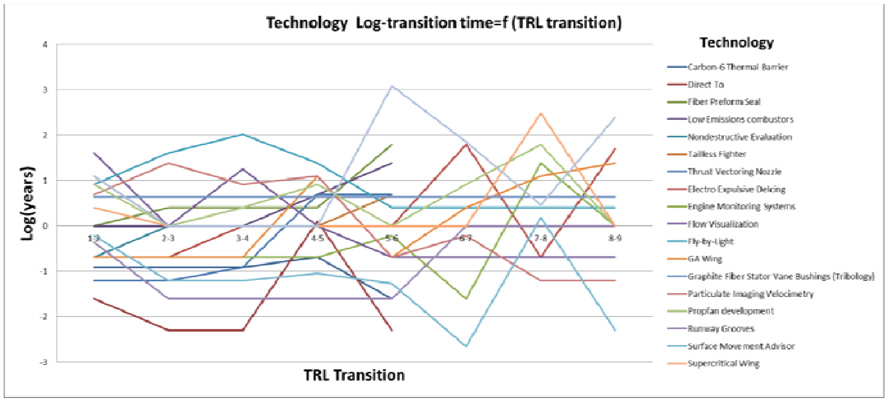

Figure 4. Log-Transition Times for NASA SAIC Data Technology

This graphical evidence of a well-behaving group before TRL 6 is confirmed by the cluster of positive correlations in Table 3. The data clearly shows a high-correlation cluster for transition times 1–2, 2–3, 3–4, and 4–5. Transition 6–7 is the least correlated with other transitions. One explanation of this lack of correlation is that the technology development responsibility changes from NASA laboratories and research agencies that have managed the research from TRL 1 to TRL 6 to large-scale programs that must integrate multiple technologies and that use contractors to complete the technology development.

Table 3. NASA SAIC Data Correlation for Log-Transition Times

Correlation Table	ln(12) log data	ln(23) log data	ln(34) log data	ln(45) log data	ln(56) log data	ln(67) log data	ln(78) log data	ln(89) log data
ln(12)	1.000	0.660	0.752	0.312	0.149	-0.074	-0.135	-0.606
ln(23)	0.660	1.000	0.905	0.673	0.385	0.043	-0.170	-0.350
ln(34)	0.752	0.905	1.000	0.639	0.351	0.113	-0.256	-0.265
ln(45)	0.312	0.673	0.639	1.000	0.490	0.344	0.006	0.073
ln(56)	0.149	0.385	0.351	0.490	1.000	0.325	0.331	0.307
ln(67)	-0.074	0.043	0.113	0.344	0.325	1.000	-0.092	0.633
ln(78)	-0.135	-0.170	-0.256	0.006	0.331	-0.092	1.000	0.180
ln(89)	-0.606	-0.350	-0.265	0.073	0.307	0.633	0.180	1.000

Figure 5 shows that the 6–7 transition time has the highest variance, which makes it very unpredictable. This confirms the GAO conclusion that technology risk is reduced after TRL 7, because finishing the 6–7 transition gets the technology past the step that has the highest variance.

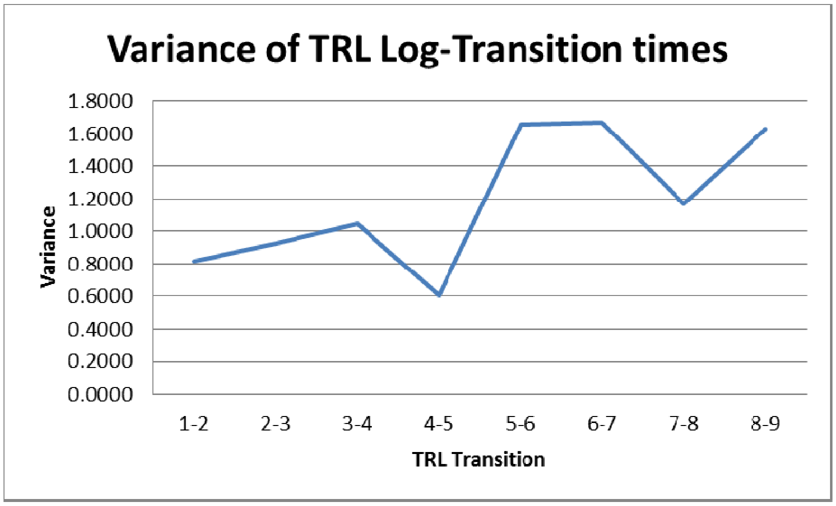

Figure 5. Variance of NASA SAIC's Technology Log-Transition Times

In simple terms, the correlation table (Table 3) implies that if a technology is maturing fast (resp. slow) at early stages, it is likely to keep its high (resp. low) maturation speed up until TRL 5. However, there are no significant correlations (positive or negative) after TRL 5.

We tried exploiting those effects by developing schedule forecasting models. In one approach, we modeled the transition times as an influence diagram (Shachter & Kenley, 1989) and did Bayesian updating to make forecasts. We also tried extrapolation techniques and some other statistical techniques, such as autoregression. The *autoregression* method consists of regressing the variable that is being forecasted against all the already known transitions by using the training data, then generating the forecast by applying the resulting linear function to the known transitions of the particular technology. For example, if we already know transitions 1–2, 2–3, and 3–4, and we want to forecast transition 7–8, we would use the training set to regress 7–8 against 1–2, 2–3, and 3–4; then we would get the forecast by multiplying the regression coefficients by the known values of 1–2, 2–3 and 3–4 transitions.

We then compared those estimates to those done by using fixed estimates (i.e., by simply using the mean or median transition time as an estimate, which is equivalent to limiting the assumptions to the Level 2 assumptions). Different measures of forecast errors were used, and we found, as expected, that most forecasting methods outperformed the fixed estimates on the early TRLs. However, only one of those techniques, the bounded autoregression, was able to consistently outperform the median and mean estimates of all TRL transitions for all of the forecasting error measures. The bounded autoregression method controls estimates through an upper bound to prevent the algorithm from predicting unrealistically high values.

This analysis is based on a small number of data points, and there is a risk that those forecasting techniques are overlearning the dataset. As long as we do not have enough testing data to properly validate this approach, it might be better to limit ourselves to Level 2 assumption and use fixed median estimates with bootstrapping.

Level 4 Assumption

The Level 4 assumption consists of mainly two separate assumptions: The first assumption is that TRL measures technology maturity and, by extension, the second assumption is that TRL also measures risk.

The first assumption that TRL measures maturity is easier to defend. A lot of factors outside of TRL contribute to technology maturity. Other factors that influence maturation variables and are not captured by TRL can be external, such as political risk (change of budget or requirements), technological risk (obsolescence due to disruptive technologies), bureaucracy (number of steps, amount of documentation, tests, demonstrations, and milestones required), contractor in charge and contracting structure and incentives, and availability of proper development labs and testing facilities. They can also be inherent technology properties, such as new technology vs. adaptation (or new use), type of technology (software vs. hardware vs. process), domain of technology (space vs. weapons), and scale of technology (component vs. system vs. integration). Nevertheless, this assumption stipulates that TRL captures a major part of the technology maturation process. After all, TRL was developed as a measure of technology maturity, and project managers can augment the model with those other factors whenever they are relevant for particular project; however, we did not find any historical data to support quantitative modeling.

The second assumption, that TRL measures risk, is not as easy to defend. As Nolte (2008) pointed out, TRL is a good measure of how far the technology has evolved, but it tells us very little about how difficult it will be to advance to the following steps. Advancement degree of difficulty (AD2) and research and development degree of difficulty (RD3) are two scales developed to eliminate uncertainty over the maturity variables of future steps.

Even if we accept the second assumption, Conrow (2009) pointed out another problem in this TRL-risk relationship. Cost, schedule, and performance risks can be decomposed into probability of occurrence and consequence of occurrence. Although TRL might be partially correlated with the probability of occurrence term, it has no correlation with the consequence of occurrence term.

Several cost uncertainty models (Lee & Thomas, 2003; Hoy & Hudak, 1994; Dubos & Saleh, 2008) are based on a regression of cost risk against TRL. Conrow (2009) mentioned a cardinality assumption made when these cost-risk models aggregate project TRLs through averaging (more specifically, cost-weighted TRLs). In our statistical modeling for schedule, we used only the ordinal properties of TRLs (the fact that TRL 2 is more mature than TRL 1, and TRL 3 more mature than TRL 2, etc.). Averaging TRLs implies that TRL values are not just placeholders, but that they are metric scales, and the distances between them can be compared (for example, this would mean that the difference of maturity between TRL 1 and TRL 4 is the same as that between TRL 5 and TRL 8). Conrow also points out other problems in this approach: There is no particular reason to weigh the TRL average by cost, and more importantly, there is no reason to believe that averages are the best way to get an aggregate system TRL. The maturity of a system (especially when we are looking at performance and schedule variables) depends heavily on the work breakdown structure (WBS). For example, if the WBS contains important parallel branches, then system maturity is better represented by taking the minimum of subsystem TRLs.

As a solution, Conrow (2009) proposed an AHP-based curve fit that calibrates the TRL scale making it cardinal. Although this approach is a good way of getting over the cardinality assumption and the assumption that TRLs measures maturity in general, there are a couple of caveats to the method:

On the one hand, the method factors in the subjectivities inherent to the AHP method that require asking and answering the 36 questions relating to how much more mature TRL scale level N is compared to TRL scale M.

On the other hand, the method factors in the subjectivities and definitional issues inherent to the term “maturity.” Let us assume the expert answered, “TRL 5 is 3.5 times more mature than TRL 2.” Were the experts thinking in terms of individual maturity variables (remaining work to be done, remaining cost, schedule, performance, program risk, or other variables related to program readiness), or were they doing some kind of mental average on some of those maturity variables?

Conrow’s (2009) technique would allow us to overcome all of the Level 4 assumptions except for the “TRL is a measure of remaining risk” assumption. Although TRL averaging is not encouraged, Conrow’s proposed method does offer a way of doing so without committing any potential cardinality-related errors, but it does add considerable subjectivity. The one major advantage of the three cost uncertainty models that do use regressions against TRL values is that they require only TRL as the input for predicting cost uncertainty and do not require the extensive subjective AHP comparison inputs suggested by Conrow (2009).

Part II: A New Framework for Cost and Schedule Joint Modeling

For joint cost and schedule modeling, we adopted a completely different (decision-based) approach of modeling TRLs and project management. The need for such a model emerges from the fact that the approaches presented so far present a major weakness when applied to joint cost and schedule models: They model cost and schedule risk separately and fail to take cost and schedule arbitrage into consideration. For example, let us assume that cost and schedule have normal distributions and that we want to generate a cost and schedule joint distribution. On the one hand, we cannot simply model them as independent variables because we would be neglecting cost and schedule interactions. On the other hand, even if we model the variables as a bivariate normal distribution, we would still be missing the cost and schedule dynamics because such a distribution reduces the interaction between cost and schedule to only one correlation factor. The relationship between cost and schedule is more than just a positive or a negative correlation: We might assume that the two variables are independent or positively correlated on some ranges, but such a relationship would not hold for the extreme values: If schedule slippage is going too high, the project manager would increase spending to reduce schedule slippage (or vice versa if cost is getting out of control). As a result, we should expect a cost and schedule arbitrage. The terms of this tradeoff would depend on factors such as (1) the available budget, (2) the reward structure of developing the technology, or (3) the relative attractiveness of other projects to where cost or schedule resources might get diverted.

First, it is not very realistic to consider a normal or lognormal budget that can take any value. Very often, the budget can take only certain discrete values and has some maximum allowable limit. More importantly, a single budget might have to be allocated across multiple projects.

A second factor that can impact the terms of the cost and schedule tradeoff is the expected payoff of developing the technology. Although some technologies generate higher utility if developed quickly, others give priority to cost considerations and generate utility as long as they don’t exceed budget.

Finally, cross-technology arbitrage is an important factor in shifting schedule and cost resources from one technology to another. For example, if a technology has already considerably matured and it is expected to have a high return if developed early, then this is a low-risk, high-return project, and it will get resources diverted to it from other higher risk projects where schedule is relatively less important (for example, projects that have low returns in the short term, but that were nevertheless started just to be kept alive). This is especially relevant if the budget is shared across different projects or technology components.

One way of capturing those variables and those project-management tradeoffs is through decision analysis using a simple finite-horizon dynamic programming model. Not only does the model incorporate those important factors, it also simulates the decision-process of a rational project manager controlling a portfolio of technologies and facing uncertainty. The proposed model works in two steps: First, we define a decision tree and solve it; second, we use the resulting policy matrix to generate the cost and schedule joint distribution. Because very little of the required data is readily available (especially joint cost and schedule data, and technology maturation utility data), what follows is a high-level description of the dynamic program and the methodology for generating the distribution (hence, the numbers that appear in the description are for illustrative purposes only).

The Dynamic Program

The model consists of nine periods (the nine TRLs of the technologies), where the manager has to manage a simple portfolio of two technologies (technology A, and technology B) initially at TRL 1. At each TRL, the decision-maker decides how to allocate the available budget on the two technologies. This decision will stochastically affect the state variable, which is the total cumulative development time of each of the two technologies. Once at TRL 9, each project will be rewarded based on its total development time.

Note that time here is the state variable and not the period of the model. We did this to avoid having TRL as a state variable. Having TRL as a state variable would generate (1) conceptual problems (it is a discrete variable that proceeds step by step; it would be unrealistic to have probabilities of a project jumping many TRLs in one period, or to have a project staying the same TRL without having its future probabilities being affected by that), and (2) practical problems (a very large number of periods would be needed if a project had low probabilities of transitioning to the next TRL). Those problems are avoided with TRL as the model's period. Hence, the total number of periods is fixed to nine; at each step the project has to progress by one TRL, and the state variable is the cumulative development time ("cumulative" because the final reward at the last step should account for the total development time of the technology).

The horizontal axis in Figure 6 shows the TRL level, which represents the nine periods of the model. Assume that we start at TRL 1 where the technologies A and B have already spent some time in development (indicated by the black bars, which represent the total development times so far for T_A and T_B). At this period, we have a lot of possible budget choices (C_A, C_B): We can either allocate the maximum possible budget for both technologies, allocate a maximum budget for one and a minimum budget for the other, or go for the minimal investment for both (represented by the colored bars under C_A and C_B).

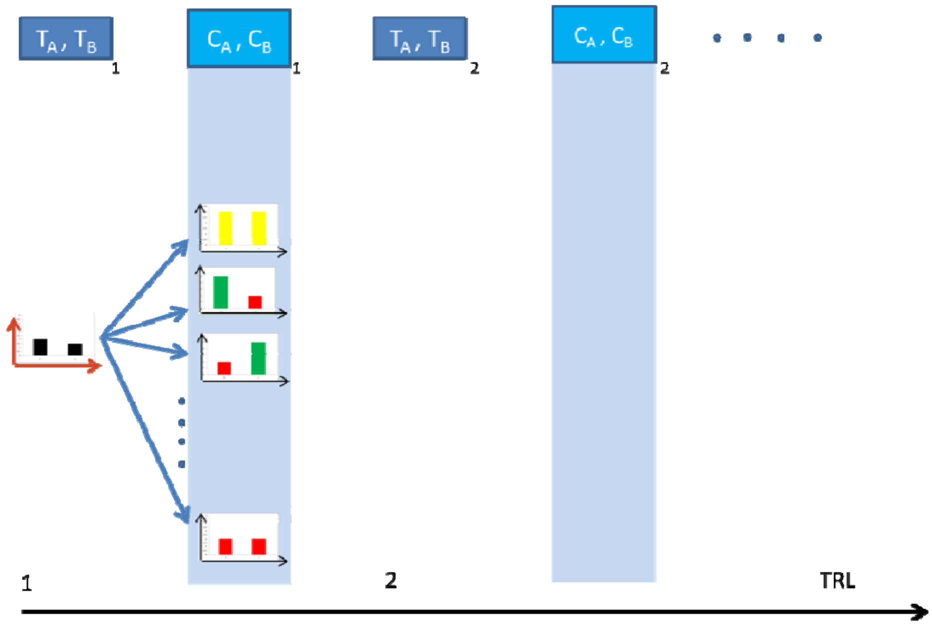

Figure 6. The State of the System at the First Period

For each of those budget decisions, there are multiple possible outcomes. When the decision-maker takes a certain budgetary decision, he or she does not know for sure how long the transition will take and what values T_A and T_B will have at Period 2 (the uncertainty in the outcomes is represented by the dotted blue lines in Figure 7).

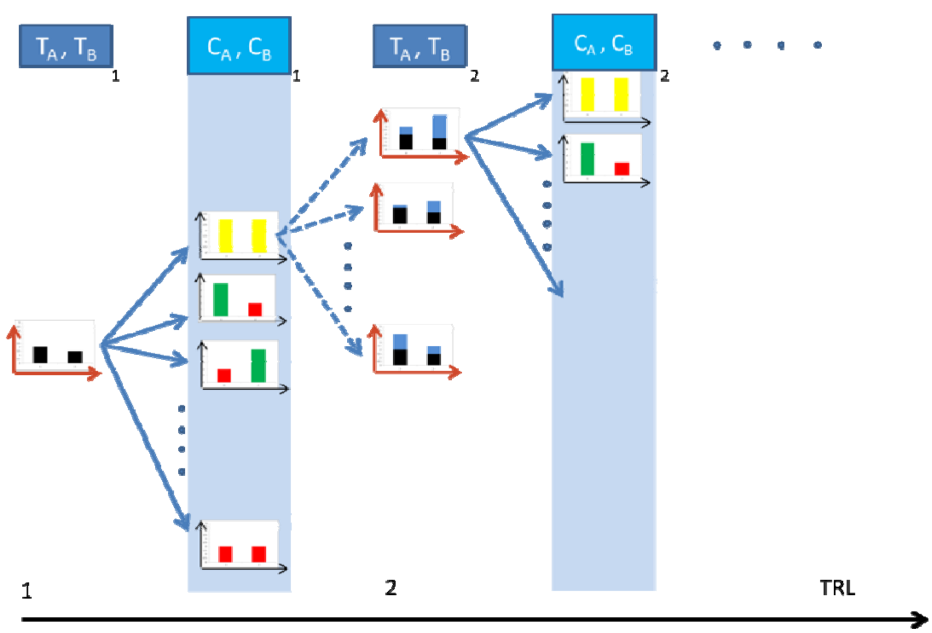

Figure 7. The Uncertain Outcomes After Allocating Budget at the First Period

The decision-maker can again make different budget allocations under each of those new states at Period 2. Finally, once TRL 9 is reached, we can calculate the terminal reward based on each of the technologies' total development time and total cost, as shown in Figure 8.

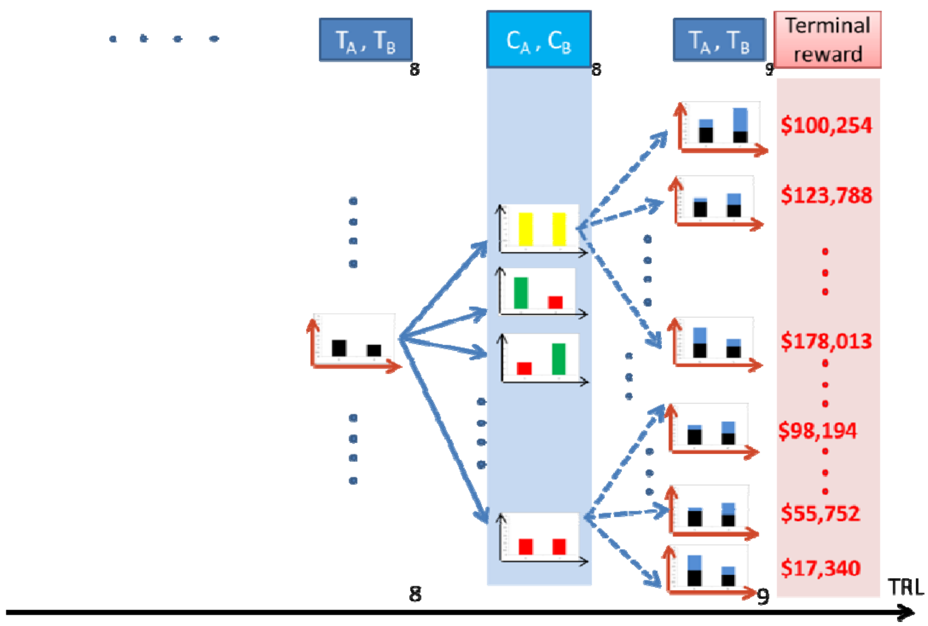

Figure 8. The Uncertain Outcomes After Allocating Budget at the First Period

Below is a more detailed description of the model parameters.

Decision Periods

T є {0,1,2,….,9} correspond to the 9 TRLs of the projects, which is when budget decisions are made.

State Variable

s є {1,2,….,45}x{1,2,….,45} the cumulative number of months of development of each of the two technologies.

Actions

In this case, we took a є {1,1.5,2,3.5,6}x{1,1.5,2,3.5,6} the budget (in million dollars) to allocate to each project at each period. Each cost corresponds (stochastically) to one of the following transition times in order {5,4,3,2,1}. This corresponds to the intuitive decreasing and convex Schedule = f(Cost) relation; it captures the cost and schedule tradeoff decisions that the manager has to make.

Those cost allocation pairs are bounded, however, by a total budget constraint $B \epsilon \mathbb{R}^9$ that limits the total budget that can be spent on every period. The aim of this vector is to put a budget constraint and to force the algorithm to take resources from one project to allocate them to another instead of investing the maximum in both technologies.

Rewards

Intermediary rewards: At each period, we incur the cost of developing the two technologies (cA_t and cB_t are the budgets allocated for projects A and B at period t).

$$R_t = -cA_t - cB_t \qquad (1)$$

Terminal reward: S_A and S_B being the state variables at T = 9 (i.e., the total technology development times), we used the following terminal reward function form

$$R = A * max(H_A - S_A, 0)^{\alpha} + B * max(H_B - S_B, 0)^{\beta} \qquad (2)$$

The difference terms ($H_A - S_A$, $H_B - S_B$) express the idea that the sooner a project is finished, the better (H_A and H_B represent the relevant time horizons over which utility is generated). Dubos and Saleh (2010) pointed out that considering useful time horizons is an important step in developing a value-centric design methodology (VCDM) for unpriced systems value (such as weapon and space systems). The max(…,0) means that the project's utility is zero if it finishes too late (i.e., if the technology is obsolete even before it matures). It also gives the option not to invest in one of the projects and allows focusing resources on only one project to maximize profits.

The utility function here was expressed as a power function (through α and β). Although it can take different functional forms, this form already expresses the fact that there are increasing marginal benefits of maturing a technology ahead of time. For example, higher values of α and β mean that the technology is badly needed and promises more future benefits if it is matured early. Hence, the utility of cross-technology arbitrage would be increased as the decision-maker has more interest in shifting one project's funds to the other in order to increase the expected final reward.

This transformation is also important because the costs and the time-related terminal rewards are eventually added to each other in the value function. A transformation is necessary in order to express the value of finishing early in dollar terms.

Similarly, A and B allow for more freedom to change the utility of the schedule terminal rewards relative to cost, and relative to each other.

Transition Function

$$s_{t+1} = s_t + w(a_t) \tag{3}$$

This means that the cumulative development time at t+1 is the total time at t plus a stochastic function of the budget decision.

$w(a_t)$ stochastically assigns one of the transition times (here {1,2,3,4,5}) to the action a_t taken by the decision-maker. Although transition time is uncertain, w assigns short transition times more often when a high budget is decided, and it assigns longer transition times more often when lower budget decisions are taken. $w(a_t)$ expresses the uncertainty in technology development; it will have a higher variance for transitions and technologies that are more uncertain. For some well-understood TRL transitions where the cost and schedule relationship is well known, the manager can use a more deterministic $w(a_t)$.

Generating the Distribution

Matlab can be used to solve this dynamic program by backward induction. As we solve the decision tree, we make sure we record the policy matrix X (a matrix that contains the best possible budget allocation for every single possible state in the tree). Once we know the best decisions all across the tree, we can redraw the tree without the decision nodes (the green dotted lines in Figure 9 indicate that the optimal decision is already taken, and the multiple tree scenarios are due only to the uncertainty of the cost and schedule relation).

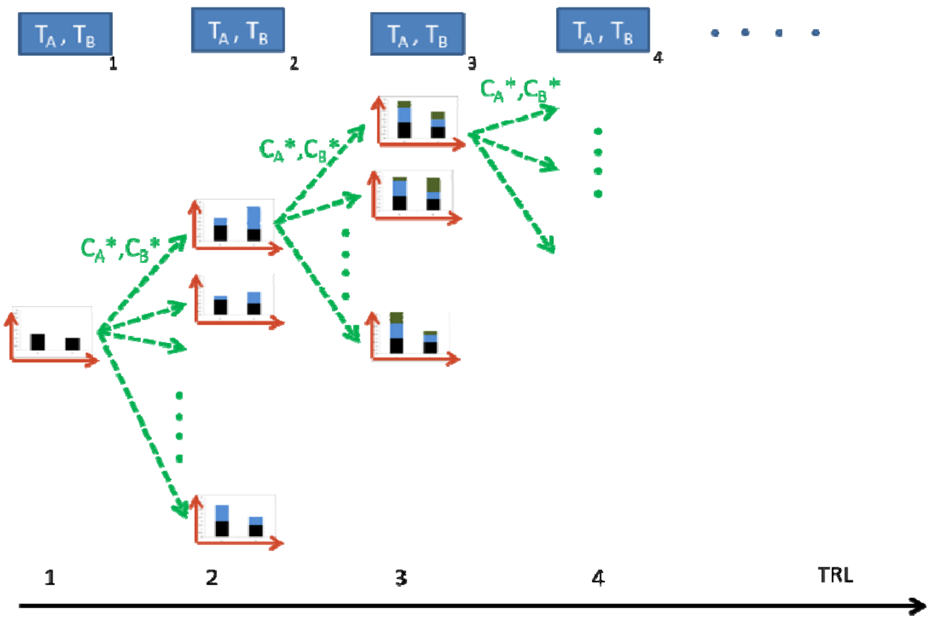

Figure 9. The Policy Diagram With Optimal Decisions

At this point, we can run a forward loop through the tree to compute the probability of getting to every possible final state, as well as the total cost when we reach this final state. In other words, assuming that a rational decision-maker is in charge of maximizing utility by managing the technology portfolio, we now have the three relevant variables of every single final state as shown in Figure 10: total development times, total expenditures, and the probability of arriving to this final state. These triplets $(C_i,T_i,p_i)_i$ define the joint cost and schedule distribution function of each of the two technologies, which can be calculated as shown in Figure 11.

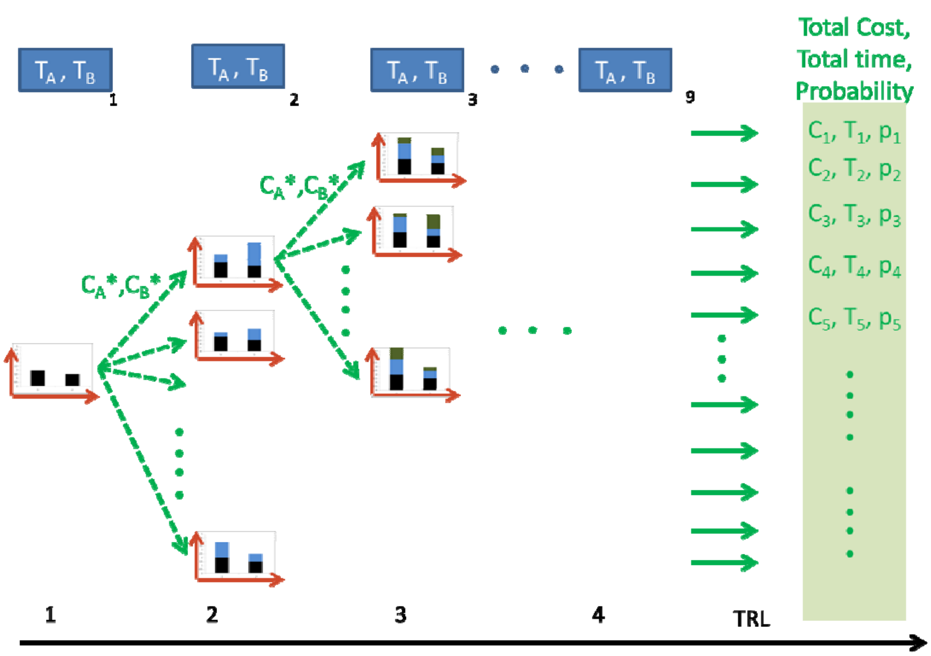

Figure 10. The Policy Diagram With Optimal Decisions and All Possible Outcomes

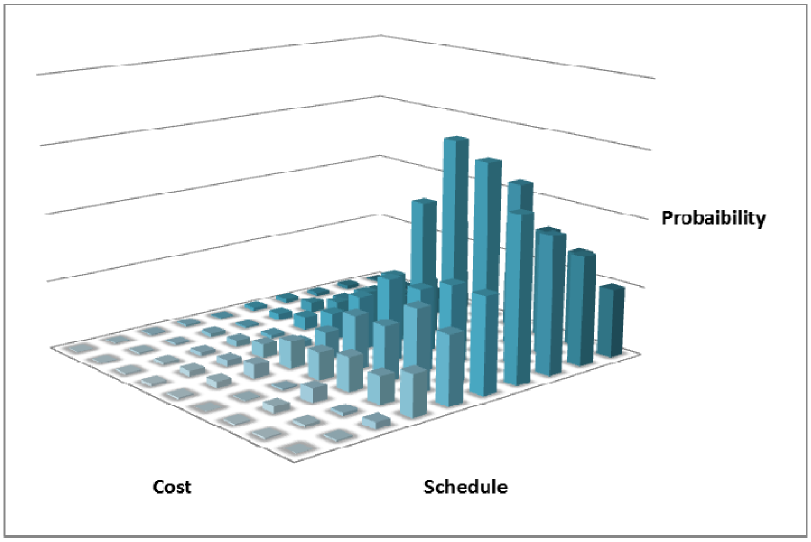

Figure 11. Example of a Resulting Cost and Schedule Empirical Joint Distribution

This model can be made more realistic by taking other factors into consideration, such as risk aversion and time discounting of money, by increasing the number of technologies in the portfolio, or by including cost in the state variable and defining a more suitable non-additive final utility function.

Conclusion

In this paper, we developed a TRL model taxonomy based on the increasing assumptions made by these models. The Level 1 assumption is irrefutable and validates the GAO's recommendations on technology transition risk. Although Level 2 and Level 3 assumptions seem to be empirically verified, the data shortage did not allow us to reliably validate the Level 3 approach, and it forced us to adopt bootstrap median estimation as a Level 2 approach instead. Finally, we are obliged at Level 4 to make at least the strong assumption that TRL is a measure of remaining risk. Although it is recommended to avoid TRL averaging and to adopt a WBS-based approach instead, we can still perform TRL averaging operations by using Conrow's (2009) calibration at the cost of the subjective inputs introduced by the AHP.

The decision-based joint cost and schedule model avoids most of the above assumptions, and it does make new assumptions on the dynamics of project management. The motivation behind this model was to capture cost and schedule arbitrage by considering different factors that affect the tradeoffs in the technology portfolio decision environment. The method theoretically generates a cost and schedule joint distribution that accounts for the decision process of the portfolio manager, but more data is needed to test and evaluate the predictive power of the model.

References

Conrow, E. H. (2009). *Estimating technology readiness level coefficients* (AIAA Paper No. 6727).

Dubos, G. F., & Saleh, J. F. (2010). Spacecraft technology portfolio: Probabilistic modeling and implications for responsiveness and schedule slippage. *Acta Astronautica, 68*(7–8), 1126–1146.

Dubos, G. F., Saleh, J. F., & Braun, R. (2008). Technology readiness level, schedule risk, and slippage in spacecraft design. *Journal of Spacecraft and Rockets, 45*(4), 836–842. doi:10.2514/1.34947

Efron, B., & Tibshirani, R. (1993). *An introduction to the bootstrap*. Boca Raton, FL: Chapman & Hall/CRC.

GAO. (1999). *Best practices: Better management of technology development can improve weapon systems outcome* (GAO/NSIAD-99-162). Washington, DC: Author.

GAO. (2009). *Defense acquisitions: Charting a course for lasting reform* (GAO-09-663T). Retrieved from http://www.gao.gov/new.items/d09663t.pdf

Hoy, K. L., & Hudak, D. G. (1994). *Advances in quantifying schedule/technical risk*. Presented at The Analytical Sciences Corporation, 28th DoD Cost Analysis Symposium, Xerox Document University, Leesburg, VA.

Lee, T. S., & Thomas, L. D. (2003). Cost growth models for NASA's programs. *Journal of Probability and Statistical Science, 1*(2), 265–279.

Mankins, J. C. (1995). *Technology readiness levels* (NASA Office of Space Access and Technology White Paper).

Mooney, C. Z., & Duval, D. R. (1993). Bootstrapping: A nonparametric approach to statistical inference. In Q*uantitative applications in the social sciences*.

Nolte, L. W. (2008). *Did I ever tell you about the whale? Or measuring technology maturity*. Information Age Publishing.

Peisen, D. J., & Schulz, L. C. (1999). *Task order 221 case studies: Time required to mature aeronautic technologies to operational readiness* (Draft report).

Shachter, R. D., & Kenley, C. R. (1989). Gaussian influence diagrams. *Management Science, 35*(5), 527–550.

Acknowledgement

This publication was developed under work supported by the Naval Postgraduate School (NPS) Acquisition Research Program Assistance Agreement No. N00244-10-1-0070 awarded by the U.S. Fleet Industrial Supply Center, San Diego (FISC-SD). It has not been formally reviewed by NPS. The views and conclusions contained in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the U.S. FISC-SD and NPS. The FISC-SD and NPS do not endorse any products or commercial services mentioned in this publication.

ACQUISITION RESEARCH PROGRAM
GRADUATE SCHOOL OF BUSINESS & PUBLIC POLICY
NAVAL POSTGRADUATE SCHOOL
555 DYER ROAD, INGERSOLL HALL
MONTEREY, CA 93943

www.acquisitionresearch.net