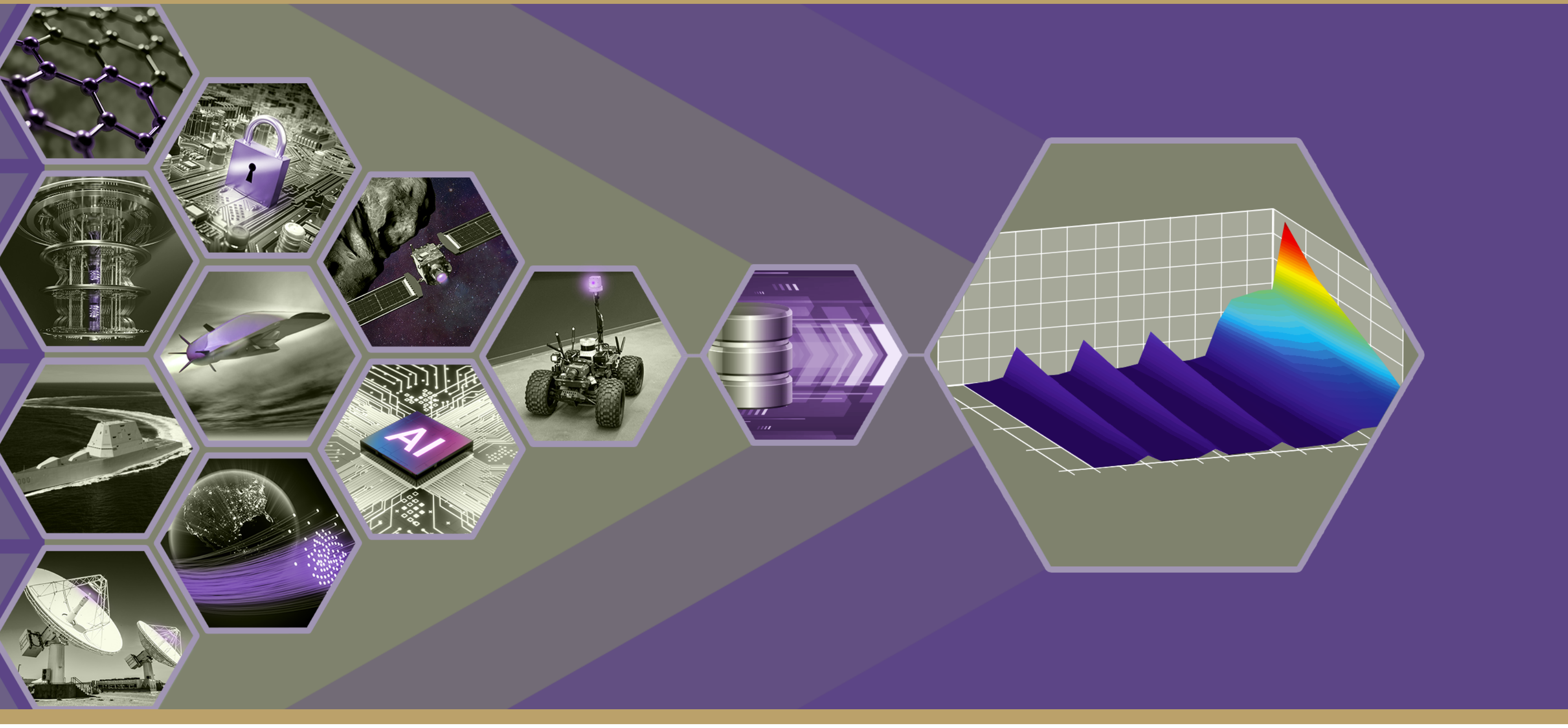
ADVANCED ESTIMATING METHODOLOGIES FOR CONCEPTUAL-STAGE DEVELOPMENT

National Security Report



Chuck Alexander

JOHNS HOPKINS
APPLIED PHYSICS LABORATORY

ADVANCED ESTIMATING METHODOLOGIES FOR CONCEPTUAL-STAGE DEVELOPMENT

Chuck Alexander

JOHNS HOPKINS
APPLIED PHYSICS LABORATORY

Contact Chuck Alexander at chuck.alexander@jhuapl.edu to request detailed model results.

This work was presented at the 2020 Military Operations Research (MORS) Emerging Techniques Forum and the 2021 NASA Cost and Schedule Symposium (NCSS). In addition, the International Cost Estimating and Analysis Association (ICEAA) selected the paper as part of its Distance Learning Series Online Cost Curriculum.

COMMS-22-03947

Contents

Figures

Figure credits:

Figure 4. GAO System Acquisition Uncertainty. From US Government Accountability Office, *GAO Cost Estimating and Assessment Guide: Best Practices for Developing and Managing Capital Program Costs (Supersedes GAO-07-1134SP)*, GAO-09-3SP (Washington, DC: GAO, March 2, 2009), 38, https://www.gao.gov/products/gao-09-3sp.

Figure 10. Relationship of Decision Points, Milestones, and Technical Reviews to MRLs and TRLs. From US Office of the Secretary of Defense Manufacturing Technology Program and the Joint Service/Industry MRL Working Group, *Manufacturing Readiness Level (MRL) Deskbook*, Version 2016 (Arlington, VA: US Department of Defense, 2016), 20, http://www.dodmrl.com/MRL_Deskbook_2016.pdf.

Figure 11. Level of Prototype Demonstrations, Venue, and Technology Maturity. From Edward Copeland, Thomas Holzer, Timothy Eveleigh, and Shahryar Sarkani, "The Effects of System Prototype Demonstrations on Weapon Systems," *Defense ARJ* 22, no. 1 (2015): 115, https://www.dau.edu/library/arj/_layouts/15/WopiFrame.aspx?sourcedoc=/library/arj/ARJ/ARJ72/ARJ-72_Copeland.pdf&action=default.

Figure 12. TRL Mapping to Prototype Demonstration Attributes. From Edward Copeland, Thomas Holzer, Timothy Eveleigh, and Shahryar Sarkani, "The Effects of System Prototype Demonstrations on Weapon Systems," *Defense ARJ* 22, no. 1 (2015): 117, https://www.dau.edu/library/arj/_layouts/15/WopiFrame.aspx?sourcedoc=/library/arj/ARJ/ARJ72/ARJ-72_Copeland.pdf&action=default.

Tables

Summary

Industry Challenges

In addition to software development, the greatest challenge in the cost estimating and economic analysis disciplines has long been developing a viable method for estimating conceptual-stage technology and system development. The industry, government, and institutional sectors have lacked effective methods and historical data with which to produce reliable cost and schedule estimates in this early-life-cycle phase. This uncertainty in estimating and forecasting can often lead to vastly inaccurate budgets and expected economic performance. As a result, cost overruns have long been a significant issue and concern across both government and industry, especially in highly technical environments traversed by the Department of Defense, the intelligence community, and civil agencies such as NASA. A range of realities drive these estimating difficulties, and major acquisition policy reforms and methodology changes over the years have improved general outcomes. At the same time, other factors, such as rapid development and acquisition strategies and techniques, have increased estimating challenges. Accurate technology and system development estimating has therefore remained elusive, except for in focused and mature technology areas where extensive technical, performance, and cost data are available.

First, the nature of new or immature technologies inherently suggests a lack of truly analogous systems from which to generate a basis of estimate. Traditional micro-parametric estimating models are also driven by fundamental engineering, design, or performance criteria that are generally unavailable in early design and development stages. Furthermore, these models typically focus on narrow technologies, functional areas, or environments and are often based on limited historical project data. These challenges are further exacerbated by the shortage of data due to the protected nature of development efforts, which often contain proprietary intellectual property, sensitive cost data, or classified information.

Methodology and Results

The research and results in this analysis overcome these formidable estimating difficulties through a comprehensive solution producing viable parametric cost models leveraging readily available metrics that reflect a full complement of primary technical, cost, and risk drivers and that can be applied across technology areas. This significantly magnifies and refines an initial 2017 investigation[1] that developed first-generation parametric cost and schedule models based on technology readiness level (TRL) and system hierarchy level (SHL) macro-parameters. The second-generation cost models developed in this extended analysis augment the base TRL improvement level (TIL) and SHL independent predictors with critical complementary macro-parameters, including research and development degree of difficulty (RD^3) and technology area (TA). This vastly expands and improves the available research, development, test, and evaluation (RDT&E) project landscape from a very rough twenty-five-data-point, two-parameter grid to a detailed four-dimensional high-resolution continuum of up to nine thousand possible SHL/TIL/RD^3/TA project configurations across the four independent macro-parameters.

1 Alexander, "Parametric Cost and Schedule Modeling for Early Technology Development."

The two greatest underlying drivers of cost, schedule, and risk for essentially any technology or system development are measures of project scale and complexity. Including these key measures that are directly associated with system function, intricacy, level of integration, development difficulty, and level of integration greatly enhances the estimating methodology by reflecting a more diverse and complete set of underlying cost and risk drivers. Model fidelity is further advanced by probability-based Bayesian uncertainty distributions, custom fit to independent macro-parameter levels. In addition to these enhancements, leading methods to aggregate these composite macro-parameter measures for multi-technology programs and system development efforts are also presented.

Further extending estimating capabilities, a standard development framework was constructed with which total development estimates can be broken down into major constituent activities and milestones for investment analysis and budget planning. This framework is based on common technology and systems development life-cycle stages, acquisition milestones, and standard research and development budget activities. A typical development cost profile is introduced relating cost expenditure levels with key acquisition milestones and TRL benchmarks. This profile is woven into the estimating framework, providing a function to break down or extrapolate costs across the primary development processes. Finally, it is also leveraged to refine the monolithic TIL parametric costs with discrete TRL start to TRL end adjustment factors.

These improvements fundamentally expand and transform the capabilities of baseline development cost models, capturing a substantially broader perspective of essential cost attributes. Forecasting power and precision is improved extensively by up to three hundred and sixty times from a coarse two-dimensional plane to a four-dimensional high-definition topography with uncertainty distributions available for each project configuration. An example probability density function (PDF) plot with a cumulative probability density (CPD) percentile curve for just one of these project data points is provided in Figure S-1. Collectively, these advancements provide a comprehensive, integrated estimating solution set for conceptual-phase technology and systems development.

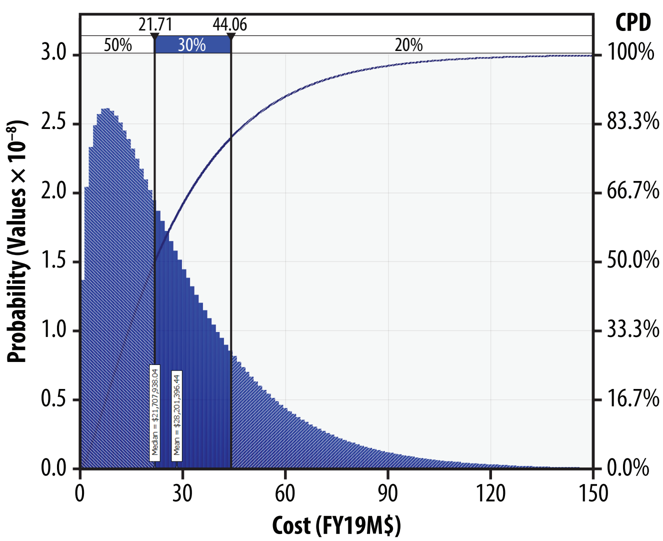


Figure S-1. PDF Uncertainty Output for a Project with SHL = 1, TIL = 4, RD^3 = 5, and TA = 4

Future Considerations

Federal agencies, research institutions, and industry technology leaders are beginning to more broadly endorse developing, measuring, and capturing standardized forms of macro-parameters to feed or integrate model-based systems engineering (MBSE) capabilities and enhance project planning, estimating, and performance measurement. Several research papers referenced in this report also focus on the need to use readiness and integration measures starting early in the development process. System readiness level, integration readiness level, manufacturing readiness level, and programmatic readiness level are each associated with various aspects of development maturity and system readiness. These measures all have potential to complement TRL-based macro-parametric forms of technology and system estimating. Complexity, as a primary cost driver, is affiliated with a variety of the underlying dimensions of TIL, SHL, RD^3, and TA macro-parameters.

Although these parametric measures hold substantial promise to improve and advance development-phase estimating, budgeting, and economic analysis capabilities, more project-level cost and schedule information will be needed for them to reach their full potential in resource planning and investment decision-making. Government policy-makers, acquisition officials, and program executives have the opportunity to implement data reporting standards and collection mechanisms to feed comprehensive MBSE processes and tools like these parametric models across the system life cycle. Through institutions and resources such as the Office of the Secretary of Defense (OSD) Cost Assessment and Program Evaluation (CAPE) and their Cost Assessment Data Enterprise (CADE) repositories and the Defense Contract Management Agency (DCMA) earned value management (EVM) contractor reporting standards, normalized and sanitized program life-cycle cost, schedule, and performance metrics can be captured to build MBSE databases large and diverse enough to support the full range of integrated life-cycle modeling capabilities.

Background

Viable cost and schedule estimating methods in the conceptual stages of systems development primarily involve analogous systems, macro-parametrics, and, to a lesser extent, micro-parametrics, as illustrated in Figure 1. Because of the general lack of analogous technologies; traditional technical, design, and performance micro-parameters; and related cost data, the research described in this report focused on estimating methodologies using macro-level predictor variables that are more readily available in this immature development phase. A previous investigation into estimating early technology development produced a series of preliminary cost and schedule models.[1] The two key input variables in the original analysis are the technology readiness level (TRL) improvement level (TIL)[2] from project start to completion and system hierarchy level (SHL) (refer to definitions in Appendix A). That research evaluated several hundred cost and schedule models traversing a full spectrum of forms, including a range of linear, nonlinear, simple and multiple regressions, and custom curve fits of the TIL and SHL independent predictors. An example of one of the higher-performing first-generation cost models is presented in Appendix B. This multiple-regression model is of the form Total Cost = $f\,[c_i + \text{TIL} + \text{SHL}]^2$, where c_i represents the regression constant intercept term. Model mean cost output results for the five SHL × five TIL matrix are displayed in Table 1.[3]

Attributes of the two initial independent variables relate directly to technology scale and maturity but have limited affiliation to other common cost drivers, such as technology and system complexity,

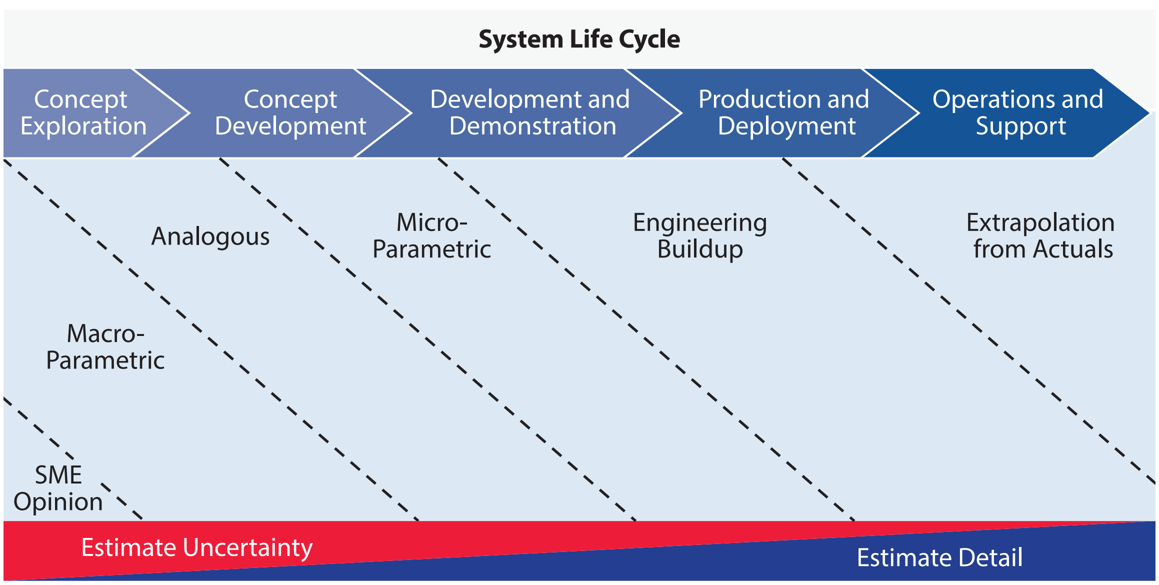


Figure 1. Estimating Methods over the System Life Cycle

[1] Alexander, "Parametric Cost and Schedule Modeling for Early Technology Development."

[2] TIL = TRL improvement from project start to completion, or project TRL end state less the TRL state at start ($\text{TRL}_{\text{End}} - \text{TR}_{\text{Start}}$).

[3] Project costs are from the NASA Technology Cost and Schedule Estimating (TCASE) tool, which defines total cost as total dollars required to complete a technology development project. This cost is provided by year and represents the total cost of labor, materials, travel, testing, equipment, and any needed facilities infrastructure investments made as part of the research project. Mean project costs shown in Table 1 are in fiscal year (FY) 2019 dollars (FY19$) converted from the initial analysis performed in FY15$.

Table 1. SHL/TIL Multiple-Regression Model Cost Output

TIL	Mean Project Cost (Calendar Year 2019 Dollars)				
	SHL 1	SHL 2	SHL 3	SHL 4	SHL 5
1	1,465,102	1,669,178	2,667,033	4,880,072	201,910,868
2	2,690,085	2,964,254	4,255,307	6,963,384	214,308,183
3	4,736,434	5,098,058	6,754,939	10,080,685	230,294,450
4	15,391,037	16,037,575	18,886,263	24,224,271	286,362,944
5	173,836,568	175,993,723	185,161,369	201,168,389	685,592,966

level of integration, development difficulty, and technology form or function. Therefore, one of the principal recommendations from the baseline analysis was to develop relationships for other prospective factors that could round out the full field of key cost drivers to capture the missing portions of the development estimating domain.

As shown in Table 1, there are significant cost escalations at both SHL 5 and TIL 5. Significant uncertainty growth is also reflected at these levels in the probability distribution results produced in the prior analysis. Driving this behavior for SHL are factors such as the aggregation of major components and subsystems and the exponential progression in the number of internal and external nodal interfaces and communication paths.[4] This includes internal hardware and software system modifications and interfaces as well as external legacy platforms or command, control, and communications system interfaces, each with potential nonlinear compound impacts on system complexity, engineering, design, integration, test, and demonstration. Possibly the greatest factor driving this cost growth, however, is the extremely broad range in the scope of SHL level 5 (i.e., the system level) that encompasses very large system-of-systems programs. This phenomenon suggests that segregating an SHL 6 for systems-of-systems development projects may be appropriate and worthy of investigation. TIL 5 similarly implies large, long-term, complex technology and systems development where costs can accelerate sharply at peak levels.

Introduction

This research examines parameters and techniques to vastly extend the capabilities and overall utility of previously developed technology and systems development estimating methodologies.[5] In-depth analysis resulted in significant improvements to the forecasting capacity, strength, precision, and reliability of preliminary TIL- and SHL-based models. More powerful custom solutions were produced via an array of advancements, including:

(1) Augmentation of first-generation cost models with supplemental macro-parameters tailored to reflect a more comprehensive set of common cost drivers. Original models are expanded from a limited twenty-five-point, two-dimensional project space to a four-dimensional macro-parametric composite topography of up to nine thousand available data points.

(2) Development of enhanced uncertainty models reflecting substantially larger and more diverse project data sets.

[4] Both of these relationships can grow at a rate approaching a theoretical limit of $(n^2 - n)/2$, where n represents the number of nodes. This second-order function parallels the second-order regression model demonstrative of one of the fundamental drivers of cost.

[5] Alexander, "Parametric Cost and Schedule Modeling for Early Technology Development."

(3) Construction of a standard technology and systems development framework integrated with key development activities, processes, acquisition milestones, and TRL achievement.

(4) Building of historical development cost benchmarks tied to the standard framework milestones and TRLs that

   (a) provide a method to segment total development estimates into a full range of common research and development (R&D) stages, milestones, and activities; and

   (b) refine uniform project TIL cost metrics with unique incremental TRL start (TRL_{Start}) and end (TRL_{End}) cost adjustment factors.

A variety of complementary independent macro-level predictors were assessed, and two complementary variables were selected to broaden and magnify the scope of limited baseline technology development models. The chosen variables, R&D degree of difficulty (RD^3) and technology area (TA), were examined, and estimating methods were explored to incorporate them into the analysis. Two primary techniques were identified, one involving formulation of mean cost factors (MCFs) and an alternative employing geometric means. These methods united the additional parameters with the first-generation SHL/TIL models, markedly improving forecasting performance and uncertainty analysis.

A standard technology development framework with associated cost benchmarks was then constructed. Common high-level research, development, test, and evaluation (RDT&E) milestones from the Department of Defense (DoD) acquisition management process and standard RDT&E budget activity (BA) category definitions were examined to develop a general development work breakdown structure (WBS). Cost estimating relationships aligned with the proposed development framework milestones and TRL achievement were then introduced based on an investigation of industry studies and historical R&D budget expenditure research. Research findings and historical budget metrics were collectively employed to produce a "typical" R&D cost benchmark landscape mapped to critical acquisition milestones and TRLs. This profile is applied to calculate and allocate costs for major elements of development, generate cost factors to refine uniform TILs into specific TRL start and end states, and also serve as an alternative cost estimate validation method.

Expanded Parametric Data Investigation

To pursue the research objective, additional project data from the NASA Technology Cost and Schedule Estimating (TCASE) database[6] were examined; these data were used for the initial research described in the 2017 report.[7] TCASE, a unique resource developed in early 2013 by the now-defunct NASA Headquarters Cost Analysis Division and SpaceWorks Enterprises, Inc., consists of a database of nearly three thousand development projects with integral user interface and query utility. The TCASE data were assessed for additional macro-parameters that would complement the preliminary TRL- and SHL-based independent variables and strengthen and enhance the base model's power and performance. Principal candidates identified from this exploration were the TCASE data fields for TA, RD^3, key performance parameters (KPPs), and advanced degree of difficulty (AD^2). These parameters showed potential to augment predictors in the first-generation model since they relate more directly to complementary cost drivers such as system complexity, performance, functionality, reliability, level of integration, and development difficulty.

These four parameters were each screened for viability as supplemental measures. AD^2 resulted in

[6] Wallace and Schaffer, *Technology Cost and Schedule Estimation (TCASE) Milestone 4*.

[7] Alexander, "Parametric Cost and Schedule Modeling for Early Technology Development."

Table 2. NASA Technology Areas (TAs)

No.	Description
TA01	Launch Propulsion Systems
TA02	In-Space Propulsion Technologies
TA03	Space Power and Energy Storage
TA04	Robotics, Telerobotics, Autonomous Systems
TA05	Communication and Navigation
TA06	Human Health, Life Support, Habitation Systems
TA07	Human Exploration Destination Systems
TA08	Science Instruments, Observatories, Sensor Systems
TA09	Entry, Descent, and Landing Systems
TA10	Nanotechnology
TA11	Modeling, Simulation, Information Tech
TA12	Materials, Structures, Mechanical Systems, Manufacturing
TA13	Ground and Launch Systems Processing
TA14	Thermal Management Systems
(+) 1	Aeronautics

Source: Cole et al., *Technology Estimating*, 48.

The list of space TAs and their supporting roadmaps were developed by NASA and reviewed and validated by the National Research Council.

Aeronautics, added to the original list as TA number (+) 1, has been labeled as TA15 for purposes of this analysis.

an insufficient population of projects to effectively apply. In addition, since the KPPs in TCASE are not standardized system attributes but rather manual entries of factors and qualities often specific to a particular project or technology, they were eliminated from contention.

Definitions for the remaining TA and RD^3 project characteristics are found in Tables 2 and 3. NASA breaks down R&D projects into fifteen standard TA categories.[8] Since NASA investigates and develops an extensive range of technologies going well beyond space and flight systems, it includes a diversity of relevant platforms, applications, and systems spanning the scientific, military, intelligence, and commercial sectors. RD^3 is the five-level qualitative scale of the degree of difficulty anticipated for a development project to achieve R&D objectives and the probability of success over the lifetime of the project.

Assessing TCASE projects for TA and RD^3 values revealed a broader, more diverse collection of project cost data than the data set included in the first-generation study. The original composite TIL and SHL parameter models contained two hundred and twenty-one project records. TA categories were found in well over seventeen hundred project records, and RD^3 measures were discovered in over four hundred. Because both parameters had substantial sample sizes as candidate predictor variables, associated project data containing each macro-parameter were extracted for analysis. However, an insufficient number of projects containing all four independent-variable measures (i.e., TIL, SHL, RD^3, and TA) was available to develop statistically significant multiple-regression models. Therefore, alternative techniques were explored to incorporate the additional parameters in the analysis; these techniques are presented in the next section.

Table 3. Research and Development Degree of Difficulty (RD^3)

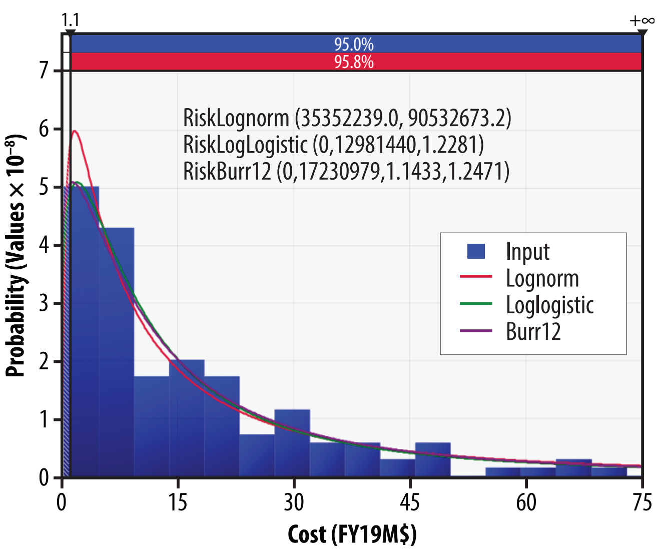
Level	Definition
5	The degree of difficulty anticipated in achieving R&D objectives for this technology is so high that a fundamental breakthrough is required ($P_{success}$ = 0.2).
4	A very high degree of difficulty is anticipated in achieving R&D objectives for this technology ($P_{success}$ = 0.5).
3	A high degree of difficulty is anticipated in achieving R&D objectives for this technology ($P_{success}$ = 0.8).
2	A moderate degree of difficulty should be anticipated in achieving R&D objectives for this technology ($P_{success}$ = 0.9).
1	A very low degree of difficulty is anticipated in achieving R&D objectives for this technology ($P_{success}$ = 0.99).

Source: Mankins, *Research & Development Degree of Difficulty (R&D3)*, 1.

8 Note that aeronautics, added to the original list as TA number (+) 1, has been labeled as TA15 for purposes of this analysis.

RD³ and TA Data Analysis and Cost Modeling

TIL, SHL, and RD³ independent variables each use progressive ordinal category-level definitions based on qualitative assessments. Therefore, they are not designed to follow any particular linear or nonlinear continuous mathematical algorithm. However, this does not mean that the ordinal or categorical costs cannot observe natural mathematical functions. Instead, it means there is no need to fit overall continuous functions, since the ordinal inputs are discrete integer values and interim fractional TRLs, SHLs, or RD³ values are both meaningless and unnecessary.[9] Similarly, TAs are not continuous variables but distinct independent categorical values that relate to cost through secondary effects such as system complexity and level of integration required.



The *y* axis represents a relative scale of probability similar to the frequency densities for a histogram. The *x* axis represents units in US dollars in FY and level shown (this is typical for all curve fit and PDF charts).

Figure 2. Example Project Cost Curve Fit PDFs for RD³ = 2 (FY19$M)

First, the RD³ and TA data were statistically analyzed to assess the viability of each parameter as a complementary independent variable. This evaluation included strong regression coefficient of determination (R^2 adjusted) cost response in the 0.7 to 0.8 range for RD³ and TA predictors (described in Appendix C). Multicollinearity, residual autocorrelation, and independence measures such as the Durbin Watson (DW) statistic, variance inflation factor (VIF), and low correlation coefficients between predictor variables also produced favorable results, as demonstrated in Appendix D. After this analysis, tailored cost curve fits for each RD³ ordinal level and selected TA categories were modeled. Example curve fit probability density function (PDF) plots for RD³ = 2 are provided in Figure 2. The highest-performing, or "best fit," of these functions for RD³ = 2 resulting in a lognormal function is shown in Figure 3 with cumulative probability distribution (CPD) and markers for a typical planning range (fiftieth to eightieth percentile). The best-fit PDF for each RD³ level and TA category was selected based on statistical selection criteria and guidelines using Palisade's @RISK software tool, as described in Appendix F.

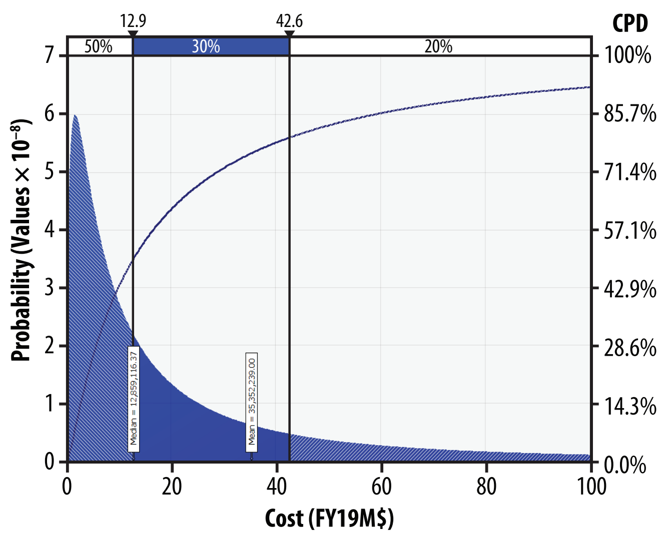


Figure 3. Project Cost Data Best-Fit PDF for RD³ = 2 (FY19$M)

9 For further discussion of ordinal, categorical, and other data types, refer to Appendix E.

The resulting RD^3 cost curve PDFs are generally highly right-skewed distributions with large standard deviations and relatively wide dispersion around the central measures. This form of uncertainty distribution is expected and appropriate for project resource, cost, and schedule data, especially for highly uncertain environments that accompany early-stage technology and systems development. This result is consistent with the US Government Accountability Office (GAO)-suggested uncertainty behavior in conceptual-stage development shown in Figure 4,[10] as well as with the right-skewed uncertainty distributions considered practical for cost estimating by the *Joint Agency Cost Schedule Risk and Uncertainty Handbook* (JACSRUH).[11] In this manner, Figure 3 effectively represents a probability-based vertical cross section of the GAO plot for $RD^3 = 2$. Numerous reasons drive this phenomenon, especially in conceptual stages, including cost growth due to the large range of unknowns, significant potential for requirements creep, technology and design changes, operational threat and environment changes, or organizational or staffing changeover. Supply chain disruptions; budget or resource priority realignments; legal, regulatory, or political environment changes; and poor management execution can also increase uncertainty and cost. Also underlying this uncertainty effect is the nature of cost—the bounded low end and the principally unbounded upper end.

With this statistical information, methods of aggregating the impact of the TA and RD^3 macro-parameters with the base-model SHL/TIL parametric cost models were explored. Two primary approaches were identified and assessed to incorporate the available RD^3 and TA project data: (1) a relative mean cost index (MCI) application method and (2) a technique merging the cost curve fit functions of the various independent predictors using a geometric mean. Both methods have the advantage of tailoring individual functions fit to each ordinal level or category, eliminating the constraint of an arbitrary forcing function across subjective ordinal parameter levels.

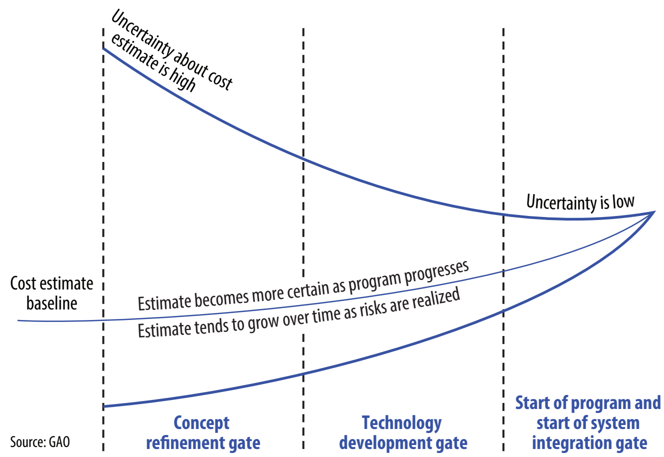


Figure 4. GAO System Acquisition Uncertainty

MCI Method

This estimating technique both extends and refines the results from the preliminary base regression cost model by establishing cost relationships between the SHL/TIL model cost data and each of the corresponding RD^3 and TA project data points. To accomplish this, the SHL/TIL project cost data and RD^3 and TA project data sets were first evaluated to establish that they are based on essentially equivalent mean project costs. In support of this premise, the three subject samples come from a common project population, each with sufficiently large and diverse sample sizes (SHL/TIL = 221, RD^3 = 425, TA = 1,730) with some project commonality and similar sample means.[12] Establishing a common sample equivalence, however, formally supports the practical application of a means-based cost relationship to model the relative impact of

[10] US Government Accountability Office, *GAO Cost Estimating and Assessment Guide.*

[11] Naval Center for Cost Analysis, *Joint Agency Cost Schedule Risk and Uncertainty Handbook.*

[12] SHL/TIL vs. RD^3 trimmed sample means fall within 0.25% of one another, and SHL/TIL vs. TA trimmed samples within 1.4% (Appendix G).

the additional RD^3 and TA macro-parameters. Therefore, equivalence was tested using the widely accepted two one-sided t-tests (TOST) and Welch's t-test (refer to the analysis in Appendix G). This analysis demonstrated practical sample equivalence between the overall SHL/TIL project data cost mean and the corresponding TA and RD^3 project data cost means.

MCI central point estimate values. MCI values relating the mean project costs for the RD^3 levels and TA categories to the SHL/TIL sample project mean were developed and assessed to determine the relative impact of both parameters on project development costs. To calculate the MCIs, RD^3 and TA project MCFs for each project were formulated. These cost adjustment factors are calculated as the ratio of the individual project cost to the SHL/TIL data set mean cost. The project data MCFs were then aggregated into summary statistical MCI measures (mean, median, standard deviation) for each RD^3 ordinal level and TA nominal category, as shown in Tables 4 and 5. These MCIs can be applied directly (i.e., multiplied) to the first-generation SHL/TIL parametric cost model outputs to refine results for RD^3 level and TA impacts.

Table 4. RD^3 Project Data MCI Statistics

RD^3 Level	RD^3 MCI		
	Mean	Median	Standard Deviation
1	0.4083	0.2352	0.4412
2	0.7759	0.3171	1.2473
3	1.0690	0.4770	2.6810
4	1.3620	0.6360	2.0470
5	1.9081	0.7929	1.7566

As with the RD^3 cost functions, the RD^3 MCI statistics demonstrate a progressive incremental relationship with RD^3 across all five levels. TA category MCIs similar to the TA cost curve fits in Appendix C produced reasonable MCI values for ten of the fifteen TAs, with the remaining five TA categories yielding questionable results, exhibiting very low or high MCI values (TAs 5, 8, 9, 10, and 15). As noted in Appendix C, these results are driven

Table 5. TA Project Data MCI Statistics

TA	TA MCI		
	Mean	Median	Standard Deviation
01 Launch Propulsion Systems	1.0940	0.0333	4.6480
02 In-Space Propulsion Technologies	0.8300	0.0416	2.5320
03 Space Power and Energy Storage	0.7940	0.0296	5.0520
04 Robotics, Telerobotics, Autonomous Systems	0.9603	0.4905	1.5894
05 Communication and Navigation	0.3125	0.0360	0.8966
06 Human Health, Life Support, Habitation Systems	1.9740	0.5900	3.2410
07 Human Exploration Destination Systems	1.8098	0.9807	2.3102
08 Science Instruments, Observatories, Sensor Systems	0.3310	0.0344	1.4660
09 Entry, Descent, and Landing Systems	13.2360	0.9640	24.8020
10 Nanotechnology	0.1025	0.0149	0.2023
11 Modeling, Simulation, Information Tech	1.4730	0.0552	6.5440
12 Materials, Structures, Mechanical Systems, Manufacturing	0.4390	0.0298	1.2510
13 Ground and Launch Systems Processing	1.8550	0.5010	4.6850
14 Thermal Management Systems	0.7125	0.0981	1.3793
15 Aeronautics	0.2186	0.0146	0.6291

by the nature of the broad uniform TA categories spanning the full range of project scale, complexity, and maturity in combination with limited sample sizes and in some instances TA inter-categorical project size concentrations. Since small sample sizes and a lack of project data diversity can result in biased statistical measures, these TA categories were therefore discarded and not applied for estimating purposes.

MCI uncertainty. In the same manner as for the RD^3 and TA cost curve fit PDFs, MCI curve fits were also produced for both parameters, and the best-performing overall function fits were selected. The resulting RD^3 level PDF @RISK functions are provided in Table 6. These PDFs are consistent with the lognormal, gamma, Weibull, and betaPERT PDFs commonly recommended by the JACSRUH[13] for estimating uncertainty. Continuing with the RD^3 level 2 example, cost curve fits from Figures 2 and 3, example MCI curve fit PDFs, and best-fit selection (i.e., lognorm) are provided in Figures 5 and 6. Appendix H also contains all TA and RD^3 and MCI PDF @RISK functions, including the corresponding plots for the other RD^3 levels (1, 3, 4, and 5). Similar to the RD^3 cost PDFs, the RD^3 MCI PDFs produced highly right-skewed distributions with relatively large standard deviations. As noted previously, these types of uncertainty distributions are expected and common for cost data, especially with the high level of unknowns and cost growth risk in early development stages.

Similar to the TA project cost data, several TA category MCIs also produced very large cost ranges and significant standard deviations, with most exhibiting very large coefficients of variation (CVs). As previously noted, this result is primarily due to the fact that each TA category spans a full range of project scale, complexity, and maturity and does not reflect graduated measurement with respect to cost. Therefore, the TA MCI PDFs contribute little

[13] Naval Center for Cost Analysis, *Joint Agency Cost Schedule Risk and Uncertainty Handbook.*

Table 6. RD^3 MCI Curve Fit PDFs

RD^3	PDF Type	@RISK PDF Formula
1	Gamma	=RiskGamma(0.59877,0.68192,RiskName ("RD3 Lvl 1 MCI"))
2	Lognorm	=RiskLognorm(0.84662,2.1681,RiskName ("RD3 Lvl 2 MCI"))
3	Pearson6	=RiskPearson6(1.1572,1.7721,0.71302, RiskName("RD3 Lvl 3 MCI"))
4	Gamma	=RiskGamma(0.71451,1.9062,RiskName ("RD3 Lvl 4 MCI"))
5	Gamma	=RiskGamma(1.3688,1.394,RiskName ("RD3 Lvl 5 MCI"))

value to uncertainty estimating in the analysis; they are shown in Table H-2 for demonstration purposes only and are not applied or recommended for modeling purposes. This does not create estimating limitations, however, since their central values still fall within reasonable ranges of overall population means, and project cost uncertainties are more effectively captured by RD^3 MCI PDFs. Attempting to model project cost uncertainty by compounding multiple perspectives (i.e., RD^3 and TA segmentations) of the same costs is also invalid since that artificially amplifies or "double counts" the impact of those uncertainties. Therefore, to avoid distorting cost risk, RD^3 MCI PDFs alone are suitable and

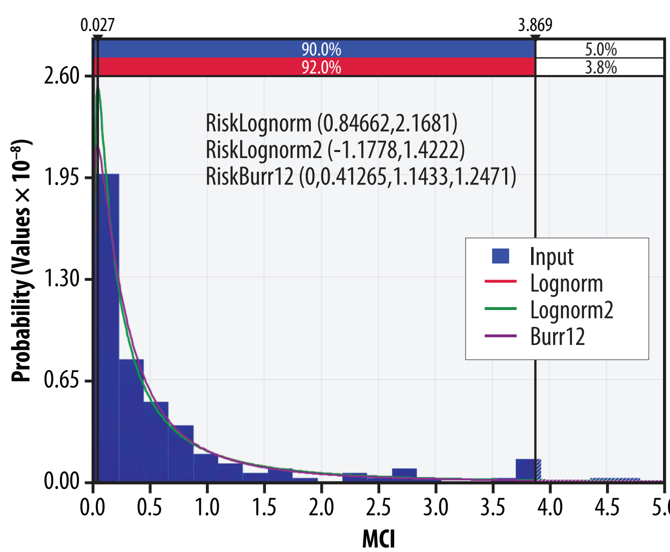


Figure 5. Example MCI PDF Curve Fits for RD^3 Level = 2

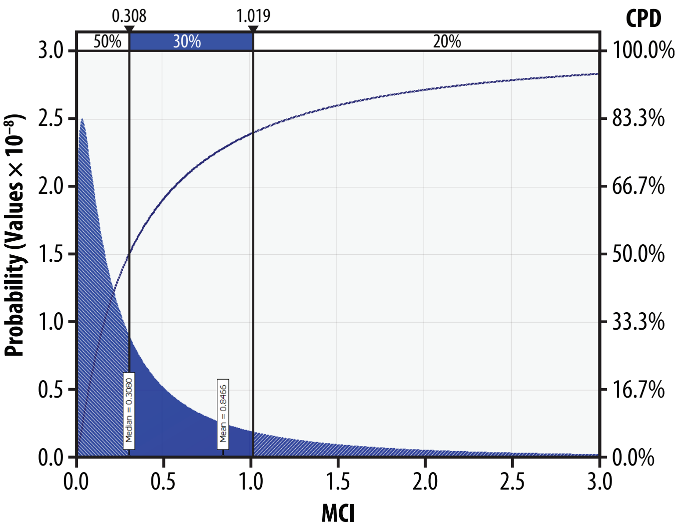


Figure 6. Selected Best-Fit MCI PDF for RD^3 Level = 2

effective for modeling total cost uncertainty. This approach provides central cost adjustment factors for applicable TA MCI categories but avoids redundant uncertainties caused by overlaying expansive TA MCI PDFs on top of the tailored RD^3 PDFs.

The resulting TA and RD^3 MCI stats in concert with the RD^3 uncertainty functions can therefore be applied directly to the range of first-generation SHL/TIL regression model variants developed in the initial research[14] to fine-tune them for the influence of the additional RD^3 and TA attributes.

Geometric Mean Curve Fit Method

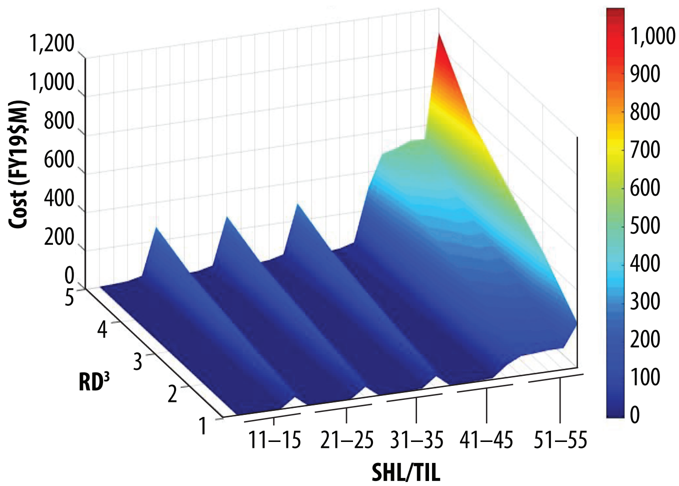
The geometric mean curve fit method involves creating composite functions of the independent variables by merging the uncertainty distributions of the selected predictor variables for each parametric combination. The average impact of individual custom cost curve fits for each independent parameter level is estimated by taking the geometric mean of their expected values (i.e., the root of their product) sampled from the individual PDFs in a Monte Carlo simulation. In a similar manner as the MCI technique, the best-performing SHL/TIL curve fits from the baseline research were applied. For this method, RD^3 and TA category project cost curve fits were applied instead of their respective static MCIs. Therefore, outputs represent the blended average of the three or four selected constituent macro-parameter groupings. This approach is fully delineated in Appendix I. However, results do not effectively capture the aggregate or compound impact of the independent parameters, and relatively low project costs were predicted with rather large residuals versus project actuals. Therefore, this method was determined to be a nonviable option for estimating purposes and abandoned.

[14] Alexander, "Parametric Cost and Schedule Modeling for Early Technology Development."

Cost Model Results

The MCI method enhances modeling capabilities, unifying the available RD^3 and TA project MCF data with legacy SHL/TIL parametric regression and curve fit models. Estimating power and precision improve, and the overall development project cost geometry grows extensively. Three- or four-parameter cost estimates can be generated as either multifaceted point estimates or composite functions with uncertainty. With a foundation of the highest-performing first-generation SHL/TIL regression model (from Appendix B), the RD^3 and TA MCI values from Tables 4 and 5 and uncertainty functions from Table 6 can be applied in product combinations to produce families of three- or four-parameter project estimates.

First, augmenting the base SHL/TIL models with just the RD^3 independent parameter MCI means from Table 4 results in 125 model configurations (25 SHL/TIL × 5 RD^3). The resulting project mean point estimates for these three-parameter models are displayed in Table J-1. An overall contour three-dimensional surface plot for the array of this SHL/TIL/RD^3 cost data is shown in Figure 7. As previously noted, each data point is a separate estimate for a unique parametric model combination. The two-digit numbers on the *x* axis represent the 25 distinct SHL/TIL-level (i.e., *xy*) combinations



The two-digit (SHL/TIL) ranges on the *x* axis represent the twenty-five distinct SHL/TIL-level (i.e., *xy*) combinations and not a continuous variable. For example, the value 23 represents a project with SHL = 2 and TIL = 3 and falls in the middle of the 21–25 range. Therefore, this plot does not represent a continuous function, but rather serves as a perspective illustration of the relative impact of the variables across their ranges versus cost scale.

Figure 7. SHL/TIL × RD^3 Mean Development Cost Model Topography Plot

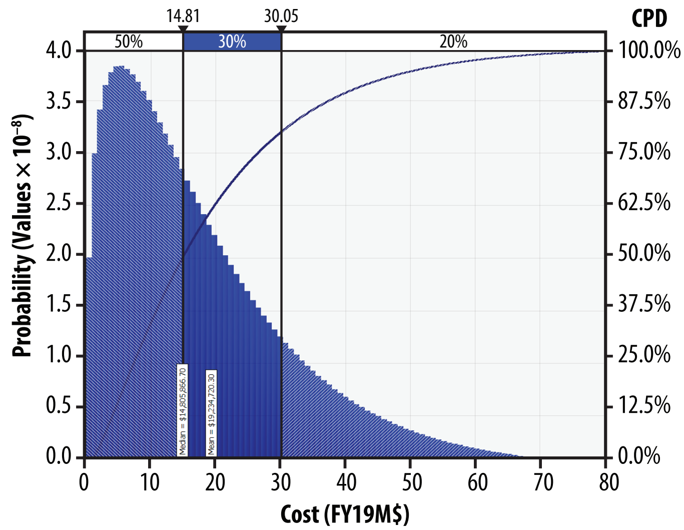


Figure 8. PDF Uncertainty Output for a Project with SHL = 4, TIL = 3, and RD^3 = 5

and not a continuous variable. For example, the value 42 represents a project with SHL = 4 and TIL = 2 and falls in the range for 41–45. Therefore, this plot does not represent a continuous function, but rather serves as an illustration of the relative impact of the variables' ranges on cost scale. Similarly specific three-dimensional SHL/TIL × RD^3 mean cost plots can be generated by TA. A detailed SHL/TIL/RD^3 PDF model for each project configuration can be also be produced by substituting the RD^3 MCI cost PDFs from Table 6 with the MCI mean values from Table 4 and running the resulting compound function through a Monte Carlo simulation. A resulting PDF estimate example for this type of three-parameter model is presented in Figure 8, representing the project attributes SHL = 4, TIL = 3, and RD^3 = 5 (i.e., model 4/3/5).

Similarly, 250 three-parameter SHL/TIL/TA model variants are produced by a product of the 25 SHL/TIL regression model output with the 10 TA category MCI cost data. The resulting mean costs for these three-parameter configurations are provided in Table J-2. Finally, the four independent macro-parameter products applied concurrently produce additional tailored composite functions reflecting the combined influence of all four variables. This expands the development cost topography to a full complement of 1,250 unique cost model variants (25 SHL/TIL × 5 RD^3 × 10 TA). These project estimates can be produced with the parametric cost data in Appendix B and Tables 4 and 5 but are too numerous to display in this report. Again, substituting the RD^3 MCI cost PDFs from Table 6 for the RD^3 MCI mean point estimate values from Table 4 within the compound functions and running the results through a Monte Carlo simulation produces a detailed multifactor model PDF for each SHL/TIL/RD^3/TA combination. An example resulting output PDF plot for one of these four-parameter model variants for project dimensions SHL = 1, TIL = 4, RD^3 = 5, and TA = 4 (model 1/4/5/4) is shown in Figure 9.

To summarize this analysis, establishing cost data relationships for the RD^3 and TA parameters advances and refines the estimating ability of the baseline parametric models and increases the segmentation of the technology development project space. Tailoring the analysis to a broader array of

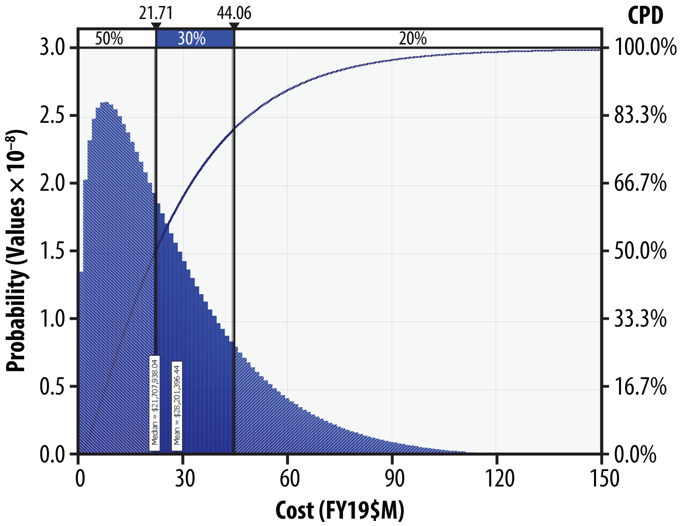


Figure 9. PDF Uncertainty Output for a Project with SHL = 1, TIL = 4, RD^3 = 5, and TA = 4

fundamental predictors with greater representation of primary technical, cost, and risk factors markedly improves the model's forecasting capacity and fidelity. RD^3 MCI PDFs also enrich the understanding of cost uncertainty as they embody much larger project data sets. Two methodologies that further extend these forecasting capabilities are developed in the following three sections. The first involves a technique to allocate total development costs to primary development activities and milestones via a standard framework tied to historical cost benchmarks. The second applies these development benchmarks to derive and refine the monolithic TIL estimates by actual TRL start and end allocation cost factors.

Standard Development Framework

The use of a product-oriented WBS is advantageous for systems acquisitions, but development processes can differ significantly from production processes with respect to the system architecture. Therefore, it is beneficial to use a development-phase WBS with common process breakdown spanning the major technology and systems development stages. An integrated framework for the total R&D phase using a standard set of "typical" development activities and milestones aligned with key macro-parameters can facilitate new technology and systems development investment scoping, estimating, and budgeting. Depending on the application and type of economic analysis, some macro-parameters that may be well suited for this purpose include TRL, system readiness level (SRL), integration readiness level (IRL), and manufacturing readiness level (MRL). Refer to Appendix A for industry definitions for SRL, IRL, and MRL.

Common development and demonstration activities related to the standard DoD acquisition process provide an extensible basis for this type of breakdown. Several authorities have linked standard R&D processes to acquisition phases and

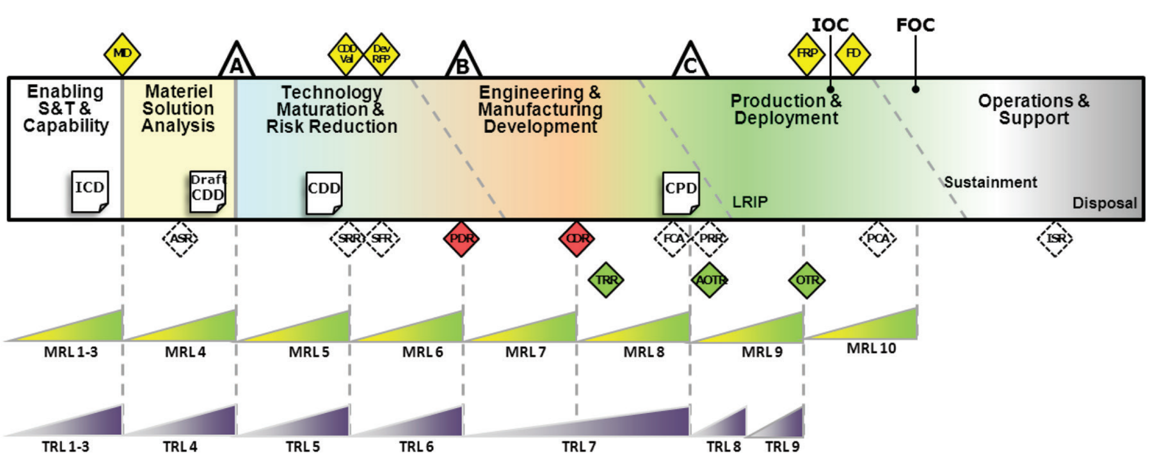


Figure 10. Relationship of Decision Points, Milestones, and Technical Reviews to MRLs and TRLs

milestones as well as general recommended TRLs, SRLs, and MRLs. Figures 10–12[15] illustrate this type of mapping. Figure 10 demonstrates the relationship of the acquisition milestone process to suggested TRLs and MRLs.[16] Figure 11 provides a similar yet slightly different perspective that includes incremental technology and system demonstrations. Figure 12 more descriptively characterizes the demonstration environments and state of technology versus suggested TRL progression. Table 7 defines the acronyms for the associated acquisition milestone and development process terms presented.

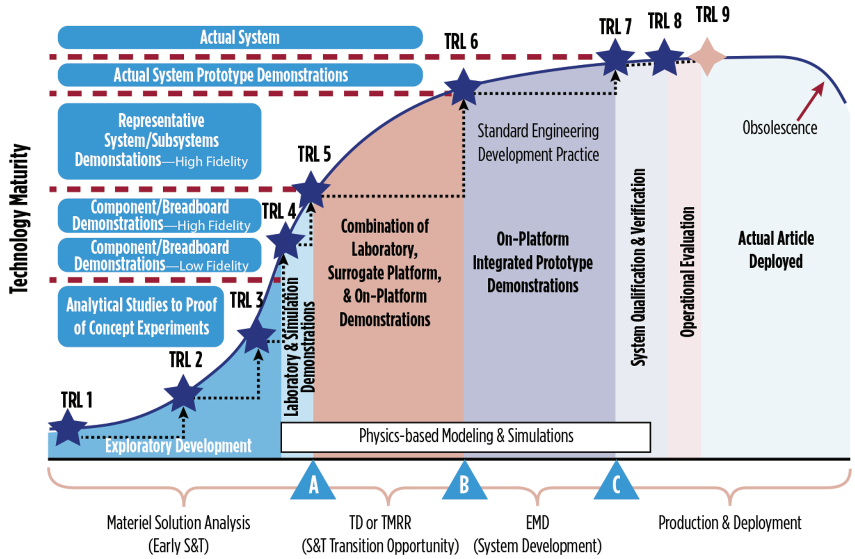


Figure 11. Level of Prototype Demonstrations, Venue, and Technology Maturity

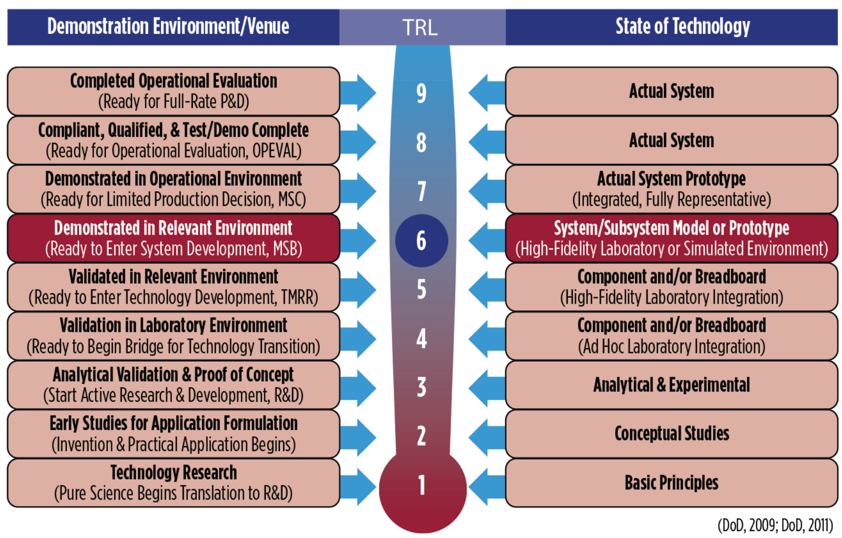


Figure 12. TRL Mapping to Prototype Demonstration Attributes

These constructs, along with the descriptions of processes and suggested technical achievements for the standard DoD RDT&E BAs,[17] vary slightly but are in general agreement for technology maturity at key acquisition milestones. Based on this general consensus, the high-level standard development framework proposed in Table 8 can be applied across a range of platforms, system architectures, and applications. This concept fuses WBS elements based on progressive development and demonstration processes or technical reviews with acquisition milestones and suggested TRLs and MRLs reached at the milestone or completion of a major activity.

Appendix L provides a more detailed four-level WBS for this framework containing a data dictionary and suggested element descriptions associated with corresponding RDT&E BAs.[18] This detailed WBS is not intended to be prescriptive but instead

[15] Copeland et al., "Effects of System Prototype Demonstrations on Weapon Systems."

[16] US Office of the Secretary of Defense Manufacturing Technology Program and the Joint Service/Industry MRL Working Group, *Manufacturing Readiness Level (MRL) Deskbook (2016).*

[17] US Office of the Under Secretary of Defense (Comptroller)/CFO, *Financial Management Regulation.*

[18] US Office of the Under Secretary of Defense (Comptroller)/CFO, *Financial Management Regulation.*

serves as general guidance in identifying the full range of activities in development, yet allowing for specific product orientation or system architectures to be threaded in where appropriate. This structure provides a comprehensive basis to help ensure that relevant design, development, integration, test, and demonstration requirements are effectively identified and captured for estimate development and budget planning.

Cost Benchmarks by Development Milestone and TRL

Ways to align cost with the development framework by bridging the progression of investment with development phases and acquisition milestones were investigated. This examination identified two conference papers that relate historical program costs with major acquisition milestones.

Table 7. Acquisition Milestone and Development Process Acronyms

Acronym	Definition	Acronym	Definition
CBA	Capability-Based Assessment	OPEval	Operational Evaluation
CDR	Critical Design Review	OT&E	Operational Test and Evaluation
DOE	Demonstrated in an Operational Environment	P&D	Production and Deployment
DRE	Demonstrated in a Relevant Environment	PDR	Preliminary Design Review
EMD	Engineering and Manufacturing Development	PoC	Proof of Concept
FOC	Full Operational Capability	S&T	Science and Technology
FRP	Full-Rate Production	SDD	Systems Development and Demonstration
ICD	Initial Capabilities Document	SRR	System Requirements Review
IOC	Initial Operating Capability	T&E	Test and Evaluation
LRIP	Low-Rate Initial Production	TMMR	Technology Maturation and Risk Reduction
MDD	Materiel Development Decision	VLE	Validation in a Laboratory Environment
MSA	Materiel Solution Analysis	VRE	Validation in a Relevant Environment
O&S	Operations and Support		

Table 8. Proposed Development Framework—Standard R&D WBS Activities vs. Acquisition Phases/Milestones and Suggested TRLs and MRLs

Development WBS	Application Phase	DoD Acquisition Milestone	TRL	MRL
1.1 Technology Development	Various			
1.1.1 Basic Research	Enabling S&T Capability	N/A	1	1
1.1.2 Technology Research	Enabling S&T Capability	CBA/ICD	2	2
1.1.3 Analytical PoC Validation	Enabling S&T Capability	MDD	3	3
1.1.4 VLE	MSA	A	4	4
1.1.5 VRE	TMRR	VRE/SRR	5	5
1.1.6 Prototype DRE	TMRR	B (PDR)	6	6
1.2 Systems Development	Various			
1.2.1 Systems Prototype DOE	EMD	CDR	6+	7
1.2.2 Full-Scale SDD	P&D	C (LRIP)	7	8
1.2.3 OpEval	P&D	IOC (FRP)	8	9
1.2.4 Operational Systems Development	O&S	FOC	9	10

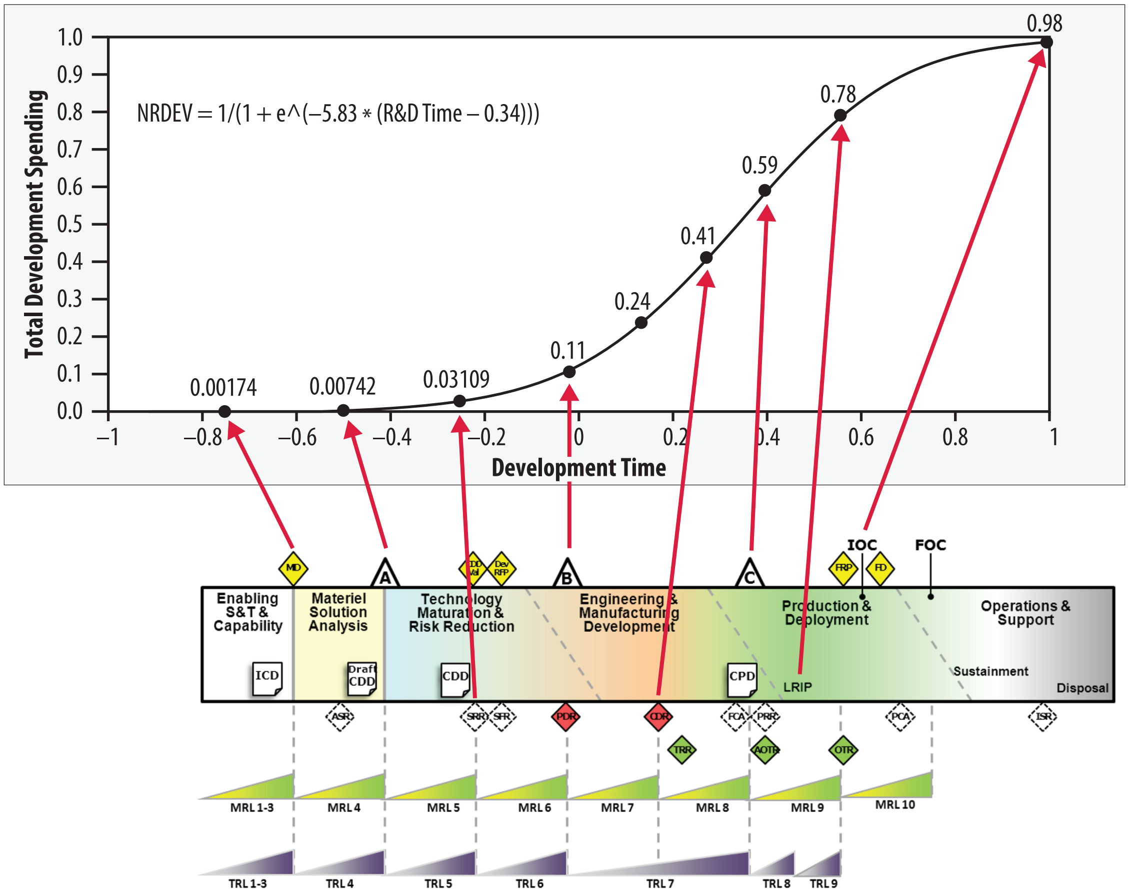


Figure 13. Development Spending Benchmarks vs. Development Milestones and TRL—Study 1

This arrangement provides historical cost profiles spanning the full development life cycle to produce progressive benchmarks at key milestones. In both analyses, corresponding functions were further fit to the resulting cost benchmarks. Expenditures can thereby be mapped to the corresponding general TRL and MRL macro-parameter achievement.

Cost Benchmark Study 1

The first study, "Methodology to Assess Cost and Schedule Impact Using System and Technology Readiness Level (SRL/TRL),"[19] was presented at the 2019 International Cost Estimating and Analysis Association (ICEAA) SoCal Workshop. This analysis applied historical selected acquisition report data from over one hundred forty major defense acquisition programs (MDAPs) to generate relative cost and schedule factors traversing the acquisition milestones. Cumulative nonrecurring development (NRDEV) spending cost benchmarks from this research were normalized (i.e., 0 to 1) over the full development life cycle and plotted across the timeline up through TRL 9. A cumulative NRDEV cost curve was fit to the equation NRDEV = 1/(1 + e^(–5.83 * (R&D Time – 0.34))). Figure 13 plots this exponential function and values for major acquisition and maps key development milestones to the progressive TRLs and MRLs in the acquisition milestone process from the Office of the Secretary of Defense (OSD) *Manufacturing Readiness Level (MRL) Deskbook*.[20]

[19] Sirirojvisuth, "Methodology to Assess Cost and Schedule."

[20] US Office of the Secretary of Defense Manufacturing Technology Program and the Joint Service/Industry MRL Working Group, *Manufacturing Readiness Level (MRL) Deskbook* (2016).

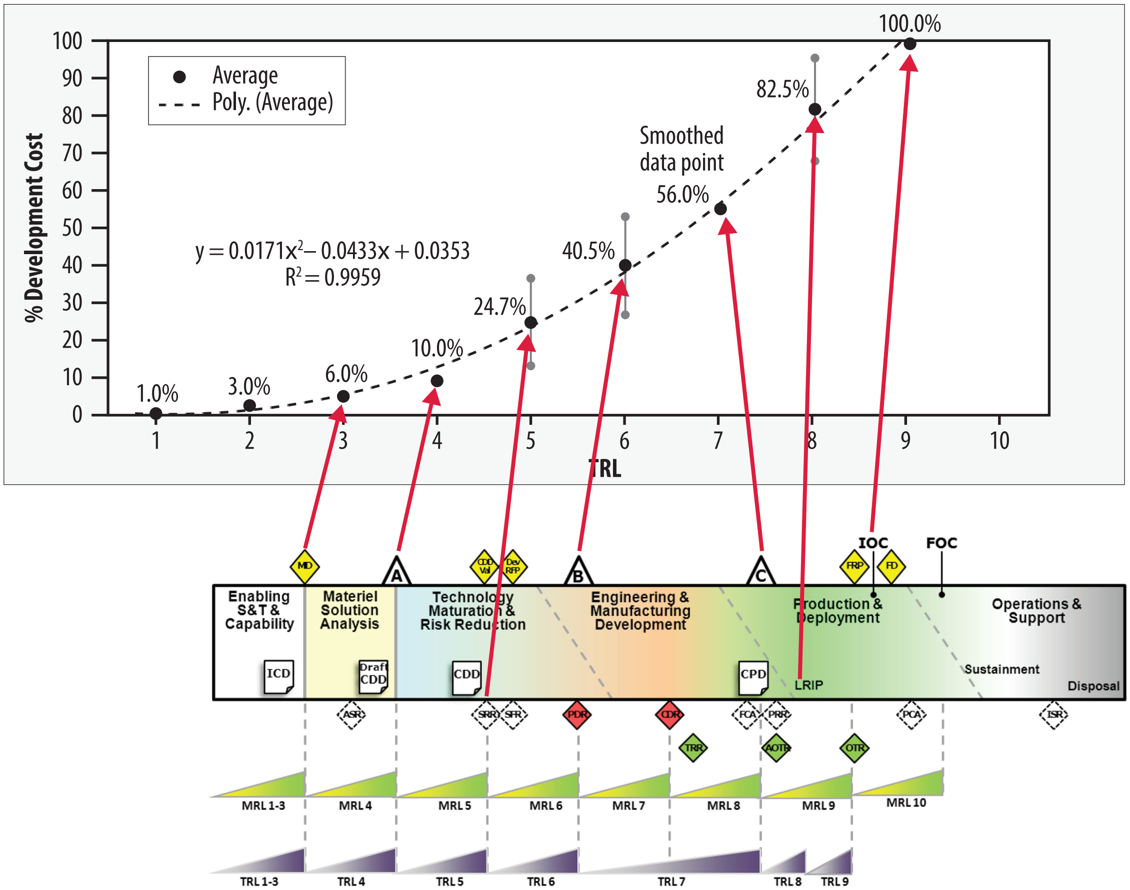


Figure 14. Development Spending Benchmarks vs. Development Milestones and TRL—Study 2

In the original analysis, TRL 8 and TRL 9 cost milestones were mapped slightly differently than suggested by the *Defense Acquisition Research Journal (ARJ)* prototype demonstration milestones, OSD *Deskbook* milestones,[21] and general RDT&E BA descriptions. To maintain consistency with the consensus of reference documentation and the resulting development framework, TRLs 8 and 9 are mapped to low-rate initial production (LRIP) and initial operating capability (IOC), respectively. As described in RDT&E BA 6.7 with development upgrades exceeding full-rate production and also demonstrated by the MRL/TRL/milestone relationship exhibit in Figure 10 from the OSD Deskbook, MRLs exceeding the extent of TRL 9 at full-rate production or IOC can continue through full operational capability (FOC). This implies that some development activities can occur past TRL 9 and mapping of TRL to milestones through FOC is needed. Some literature[22] addresses this shortfall, expressing the need for adding another TRL level to accommodate and capture post-IOC activities into extended operations.

Cost Benchmark Study 2

A second analysis, this one presented during a 2017 ICEAA conference,[23] similarly produced plots of percent development cost versus TRL (Figure 14). These results were based on a curve fit to data from approximately thirty programs primarily

[21] Copeland et al., "Effects of System Prototype Demonstrations on Weapon Systems."

[22] Straub, "In Search of Technology Readiness Level (TRL) 10."

[23] Linick, "Technology Development Level (TRL) vs. Percent Development Cost."

Table 9. Study 1 vs. 2 Cost Benchmarks by Development Milestone and TRL/MRL

Milestone at End	Macro-parameter		Sirirojvisuth Study		Linick Study	
	MRL	TRL	Total Development (%)	Cumulative Total Development (%)	Total Development (%)	Cumulative Total Development (%)
N/A	1	1	Negligible	Negligible	1.0	1.0
CBA/ICD	2	2	Negligible	Negligible	2.0	3.0
MDD	3	3	0.17	0.17	3.0	6.0
A	4	4	0.57	0.74	4.0	10.0
VRE/SRR	5	5	2.36	3.10	14.7	24.7
B (PDR)	6	6	20.4	23.5	15.8	40.5
CDR	7	6++	17.2	40.7	7.6	48.1
C (LRIP)	8	7	17.7	58.5	7.9	56.0
IOC (FRP)	8+	8	20.0	78.4	26.5	82.5
FOC	9	9	19.6	97.9	17.5	100.0

Negligible comparative costs occur in some early stages, while in other interim stages data were not available (N/A).

from an earlier paper by E. Conrow.[24] Conrow's research examined analytical hierarchy procedure–based TRL values using source data from a prior study.[25] The source data included cost information for programs in NASA's Resource Data Storage and Retrieval System (REDSTAR) database. The curve fit from this analysis similarly demonstrates a relationship between TRL and total development cost, normalized to a range of 0 to 1. A second-order function, $y = 0.0171x^2 - 0.0433x + 0.0353$, where x represents the current state of TRL, was fit to the data, producing a very strong coefficient of determination (R^2) exceeding 99%. Percent development cost points on the Figure 14 graph are the actual project empirical values and not calculation approximations from the curve fit function.

Table 9 compares development cost benchmark results from both studies versus development-phase acquisition milestones and associated TRLs and MRLs after application of a consistent methodology through IOC. Examination of the outcomes demonstrates that the two methods produced very similar outcomes for TRLs 7–9 but rather divergent results for TRLs 1–6.[26]

RDT&E Historical BA Data

To provide a third perspective on the progression of TRL-based development costs, another method, this one using historical RDT&E BA cost data, was employed. Key advantages of leveraging this data set are the huge data sample size and the application of BA categories characterizing the continuous sequential steps in the advancement of the development process that are already aligned within the proposed development framework, associated acquisition milestones, and TRL levels. Therefore, RDT&E expenditures were applied by BA. Twenty-three years of actual RDT&E R-1 budget exhibits by BA from FY1996 through FY2018 were analyzed to create historical BA cost profiles.[27] With so many consecutive years of data being used, the statistical summaries effectively represent the

[24] Conrow, "Estimating Technology Readiness Level Coefficients."

[25] Lee and Thomas, "Cost Growth Models for NASA's Programs."

[26] Sirirojvisuth's cumulative total development costs equal approximately 98% because some post-IOC development work was not included in the totals.

[27] US Office of the Under Secretary of Defense (Comptroller)/CFO, *DoD Budget Request*.

Table 10. Annual DoD % RDT&E Expenditure by BA (FY1996 to FY2018)

RDT&E BA	Milestone at End	Notional TRL at End	Annual Development	
			Average (%)	Cumulative Development-Phase Cost (%)
6.1 Basic Research	N/A	1	2.8	2.8
6.2 Applied Research	MDD	3	7.6	10.4
6.3 Advanced Technology Development (ATD)	VRE/SRR	5	9.2	19.6
6.4 Advanced Content Development and Prototypes (ACD&P)	B (PDR)	6	20.7	40.2
6.5 System Development and Demonstration (SDD)	C	7	23.1	63.3
6.7 Operational Systems Development	IOC (FRP)	9	36.7	100.0
Total Development			100.0	

development cycle of hundreds of DoD-wide programs of varying size and complexity. Statistics for percentage of total expenditures across BA categories 6.1 through 6.7 were created, characterizing the weighted average development costs spanning all development programs and stages.[28] In Table 10, summary statistics for these historical RDT&E R-1 BA expenditures corresponding to the completion of each major BA category funding phase over the historical time frame are linked with the acquisition milestones and suggested TRLs from the framework. Annual expenditures are fairly consistent over the twenty-three-year window, with CVs by BA falling primarily in the 8% to 13% range.

Although only six BA categories are available to map to the nine TRLs and corresponding development milestones, they align well with six milestones based on the description of activities and technical achievements at the completion of each BA.

Cost Benchmark Comparison

Table 11 charts the cost relationships, comparing the six shared or common RDT&E BA milestones and TRL-level cumulative percent development cost benchmarks versus the corresponding results from the two benchmark studies. TRL cumulative percent development cost benchmark data for the common milestones and TRLs demonstrate that the Linick research and analysis is fairly well aligned with the DoD-wide RDT&E results. In Sirirojvisuth's research, lower relative total development expenditures in early stages may be a reflection of the source data all being from large MDAP acquisition category I programs versus a more diverse range of programs for the R-1 exhibit BA data and the NASA project data from the Linick study. This lower relative early expenditure characteristic of MDAP program data may be the result of initial technology development efforts for very large, complex systems being a smaller portion of total development due to economies of scale similar to the spread of fixed or overhead cost pools over a larger base. More conservative existential technology selection to reduce overall developmental risk potential may also be an artifact of large investment programs.

Other factors underlying Sirirojvisuth's MDAP program development expenditure profile may include larger, more significant portions of MDAP program early-stage technology research (basic, fundamental, incubation) being captured or funded under separate organizational budgets or shouldered by a broader range of institutions. This could be the result of larger programs having the ability to leverage

[28] Expenditures include overseas contingency operations RDT&E funding. BA 6.6, RDT&E Management Support, was spread across the other six BA categories in proportion to annual expenditure amounts, so it did not alter or impact the effective percent development calculations.

Table 11. RDT&E BA Category Completion vs. Study 1 and 2 Milestone/TRL Cost Benchmarks

RDT&E BA	Milestone at End	Notional TRL at End	Cumulative Development-Phase Cost (%)		
			RDT&E R-1 Exhibit BA Budgets	Sirirojvisuth Study	Linick Study
6.1 Basic Research	N/A	1	2.8	1.0	Negligible
6.2 Applied Research	MDD	3	10.4	6.0	0.2
6.3 ATD	VRE/SRR	5	19.6	24.7	3.1
6.4 ACD&P	B (PDR)	6	40.2	40.5	11.1
6.5 SDD	C	7	63.5	56.0	58.5
6.7 Operational Systems Development	IOC (FRP)	9	100.0	100.0	97.9

Negligible comparative costs occur in some early stages, while in other interim stages data were not available (N/A).

more highly specialized skill sets and facilities or test ranges and assets needed. This strategy could potentially reduce or spread overall budget risk exposure for large program efforts by allocating portions of critical early-phase technology development to be more widely burdened by other organizations in the testing or science and technology communities. In this manner, the total associated development costs may not effectively be captured in MDAP selected acquisition reporting for programs that involve substantial efforts by bodies such as government labs, university research institutions, industry research groups, and key contractors or vendors with vested interests. An example is independent R&D investments made by large defense contractors. These factors and others could potentially contribute to the TRL 1–6 cost deviation in the MDAP program data used in the Sirirojvisuth study. Above TRL 6, however, the cumulative costs catch up and converge with the Linick findings as technology development transitions into broader overall systems development. As a result of these findings, the Linick results were applied for the parametric model TRL refinements introduced below.

Fine-Tuning TIL Estimates for Discrete TRL Start to End States

Another fundamental benefit to generating relative cost profiles across TRL levels is that they can be used to significantly enhance the fidelity and precision of the uniform TIL-based models. The incremental empirical cost benchmarks provide a means by which to calculate the relative size of all TRL start-to-end transitions. Homogeneous TIL costs from first-generation models[29] can thereby be fine-tuned to their discrete constituent project TRL_{Start} to TRL_{End} state costs via the relative cost adjustment weighting factors produced in Table 12. The second column in the upper section of this table shows the cumulative development cost (percent) at the TRL_{End} achieved for that row and is taken from the last column of Table 9 (Linick/Conrow analysis). These data were selected for application because they track well with results of the expansive DoD RDT&E BA program data, representing a diverse range of projects in terms of scale, complexity, difficulty, and uncertainty. The columns to the right show the incremental cost increase (percent) for the eight unitary TRL improvement transitions for TIL = 1 (e.g., 1 to 2, 2 to 3, 3 to 4, 4 to 5, 5 to 6, 6 to 7, 7 to 8, and 8 to 9). Percentages numbered 2 through 8 in the subsequent columns of the top section also represent the development cost increase to achieve the TRL_{End} (first column) starting from TRL_{Start} determined by: $TRL_{Start} = TRL_{End} - TIL$. TIL increases up through the maximum possible value

[29] Alexander, "Parametric Cost and Schedule Modeling for Early Technology Development."

Table 12. Cost Adjustment Weighting Factors—Discrete TRL Transition Start to End to TIL Average

TRL End	Cumulative Development Cost (%)	Total Development Cost between TRL Start and End (%)							
		TIL 1	TIL 2	TIL 3	TIL 4	TIL 5	TIL 6	TIL 7	TIL 8
1	1.0								
2	3.0	2.0							
3	6.0	3.0	5.0						
4	10.0	4.0	7.0	9.0					
5	24.7	14.7	18.7	21.7	23.7				
6	40.5	15.8	30.5	34.5	37.5	39.5			
7	56.0	15.5	31.3	46.0	50.0	53.0	55.0		
8	82.5	26.5	42.0	57.8	72.5	76.5	79.5	81.5	
9	100.0	17.5	44.0	59.5	75.3	90.0	94.0	97.0	99.0
	TIL Average % =	12.4	25.5	38.1	51.8	64.8	76.2	89.3	99.0
TRL End		Relative Cost Weighting to TIL Average							
		TIL 1	TIL 2	TIL 3	TIL 4	TIL 5	TIL 6	TIL 7	TIL 8
1									
2		0.16							
3		0.24	0.20						
4		0.32	0.27	0.24					
5		1.19	0.73	0.57	0.46				
6		1.28	1.20	0.91	0.72	0.61			
7		1.25	1.23	1.21	0.97	0.82	0.72		
8		2.14	1.65	1.52	1.40	1.18	1.04	0.91	
9		1.41	1.73	1.56	1.45	1.39	1.23	1.09	1.00
	TIL =	1	2	3	4	5	6	7	8

of 8 in the last column, for which only one possible transition exists (TRL 1 to 9).

The cost factors in green in the lower section of the table are simply the relative costs versus the TIL category average. They are calculated by taking the category percentage from the matching category cell in the upper table (i.e., the value for the combination TRL_{End} row and the TIL column from the upper table) divided by the overall average percentage for that TIL category found in the row named TIL Average % =. These relative weighting factors range from 0.16 to 2.14 and can be applied to tailor the uniform TIL costs from the first-generation TIL-based parametric models to arrive at discrete TRL_{Start} to TRL_{End} transition costs for all thirty-six possible transitions.[30] These cost adjustments further expand and refine the total development cost space to nine thousand possible data points (36 TRL start-end combinations × 5 SHLs × 5 RD³s × 10 TA categories).

TIL progressions above level 5 (i.e., in the 6 to 8 range) as part of one continuous project were found to be extremely rare in the large project population of the NASA TCASE database (approaching three thousand total project records). Therefore, there were inadequate data to model TILs above level 5 in the original TIL/SHL parametric models.

[30] The *n*th triangular number, or "termial function," of possible combinations for an interval range of 8 (i.e., 1 to 9) = $(n^2 + n)/2 = (64 + 8)/2 = 36$.

Even though it is unlikely that TILs 6 through 8 will occur or need to be calculated, this output also makes it possible to estimate these very large transitions by extrapolating beyond TIL 5 using the average TIL values as relative weighting factors. The two example estimates in Appendix K apply the discrete TRL adjustment factors to the three- and four-parameter combination (TIL/SHL, RD^3, and TA) examples described in the section on cost model results.

Composite Project or System Macro-Parametric Measures

Development projects may range from an individual technology development up to a system with multiple technology development efforts at varying states of maturity, scale, and development difficulty (i.e., TRL, SHL and RD^3). The scope of possible development projects can also extend to portions or potentially even all of the development life cycle demonstrated in Figures 10–12 and the development WBS in Table 8 and Appendix L. Projects that involve the advancement or progression of multiple technologies must include relevant integration, testing, and demonstration of those technologies up through the applicable TRL and development milestones at completion. Depending on the overall project SHL and phases of development involved in the TRL transition(s), this may include internal integration and test at the assembly, subsystem, and system levels. If project development progresses into broader systems development, it may also involve integration into external platforms, applications, networks, and command and control systems or processes up through operational testing and demonstration.

When applying a macro-parametric estimating approach to multi-technology developments that are part of one project or program, to the extent possible, each individual development should be estimated separately and then be rolled up or aggregated with progressive levels of integration, testing, and demonstration. However, if separate efforts are estimated together as one effort, an overarching SHL should be used to reflect the highest aggregate or predominant level of development. When aggregating the composite TRL and RD^3, independent macro-parameters must reflect the weighted average values across the overall project or system. Approaches have been proposed to calculate compound system or program TRL measures. For instance, Lee and Thomas[31] estimated a cost-weighted TRL, applying a component to total program percent cost allocation. Sophisticated multifactor TRL calculators and utilities have also been devised based on the weighted arithmetic or geometric mean of a range of attributes spanning TRLs. These include the Air Force Research Laboratory Transition Readiness Level Calculator[32] (refer to the paper by Nolte, Dziegiel, and Kennedy[33] and NASA's TRL Worksheet[34]). Alternative techniques applying scalars such as technical design (e.g., size, weight, and power requirements), performance, or complexity-related metrics could also individually or collectively be applied as relative weighting coefficients for calculating overall system or project TRL or RD^3 development parameters.

Sauser et al.[35] introduced a resourceful method to measure SRL as a function of TRL and IRL that deliberates both the technologies and integration elements along a numerical maturation scale to assess the maturity of the entire system. For this analysis, SRL is computed as a mathematical function using TRL and IRL matrices weighted on each technology within the system

[31] Lee and Thomas, "Cost Growth Models for NASA's Programs."

[32] Air Force Research Lab, Transition Readiness Level Calculator.

[33] Nolte, Dziegiel, and Kennedy, *Technology Readiness Calculator*.

[34] NASA Earth Science Technology Office, "ESTO TRL Worksheet Calculator."

[35] Sauser et al., *Development of Systems Engineering Maturity Models and Management Tools*.

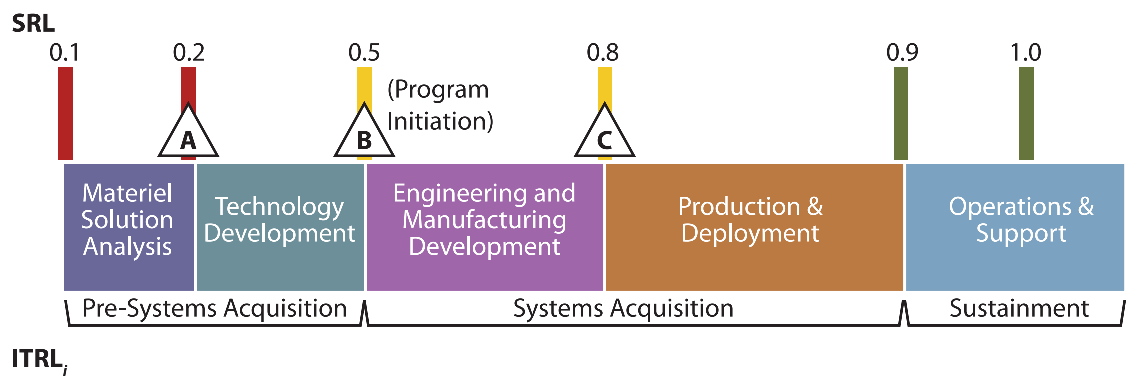


Figure 15. SRL Mapped against DoD Acquisition Life Cycle

according to all of its integrations at a "system" level: $[SRL]_{n\times 1} = [Norm]_{n\times n} \times [IRL]_{n\times n} \times [TRL]_{n\times 1}$, where, in the TRL and IRL matrices, the original (1,9) levels are normalized [*Norm*] to (0,1).[36] Like TRL, IRL is defined as a series of levels that relate to key maturity events for integration activities. Similar to TRL and MRL mapping, presented in the sections on the standard development framework and cost benchmarks, SRL is normalized across the DoD acquisition life cycle in this analysis, as shown in Figure 15. The Naval Sea Systems Command, with support from Northrop Grumman Corporation, has validated this SRL model to monitor development and integration progress in the Littoral Combat Ship Mission Modules Program.[37]

Results and Conclusions

Estimating conceptual-stage technology and systems development has long been the most uncertain, volatile, and challenging form of estimating for industrial, governmental, and institutional planning and investment decision analysis. This is primarily due to (1) the general lack of analogous systems; (2) unavailable micro-level technical, design, or performance-related parameters at this stage of development; and (3) the shortage or complete lack of applicable historical cost and technical data. This investigation leveraged several methods to build a comprehensive estimating solution applying diverse empirical project data with risk-based Bayesian techniques to fill this estimating methodology void in early-conceptual-stage development.

To avoid the problem with micro-level technical, design, or performance-related parameters generally unavailable at this stage of development, this solution set uniquely applies techniques using key macro-parametric cost and schedule drivers that are readily available or determinable in pre-design stages across TAs. The addition of the RD^3 and TA parameters substantially augments prior baseline TIL/SHL-based models, providing a more complete picture of the key drivers of technological and system scale, complexity, functionality, maturity, difficulty, and integration. Limited legacy technology development models based on a coarse two-dimensional TIL × SHL cost grid[38] are thereby transformed by up to three hundred and sixty times the predictor output using comprehensive four-dimensional TIL, SHL, RD^3, and TA solutions. High-level risks (known and unknown) associated with conceptual-stage technology development are also effectively captured by the generation of composite PDFs tailored to each specific project parametric configuration. Forecasting power, depth, and precision are all greatly enriched, reflecting a

[36] NASA, "Technology Readiness Levels (TRLs)."

[37] Sauser et al., *Development of Systems Engineering Maturity Models and Management Tools.*

[38] Alexander, "Parametric Cost and Schedule Modeling for Early Technology Development."

comprehensive set of primary technological, programmatic, and cost risk factors.

Results of cost models show that expected central costs for RDT&E projects can range from single-digit millions to billions of US dollars and spike exponentially at higher TIL and SHL levels. Cost uncertainty can vary as much as two to three times just within the typical planning range of the fiftieth (median) to eightieth percentile for a particular project configuration (TIL, SHL, RD^3, TA) and as much as 4.5 times across RD^3 levels for a fixed TIL and SHL configuration. This is expected, and there are valid reasons for this variability related to both known and unknown risks that are common in early-stage technology and system development.

A breakdown of common development processes into WBS elements linked to standard acquisition milestones and readiness levels was also created. This consensus framework was further related to cost benchmarks employing empirical studies and historical DoD RDT&E data. This contributes value to the macro-parametric cost modeling capabilities by developing a useful method with which to estimate cost at various interim stages of development. In addition, the relative TRL transition cost factors deliver a technique to refine broad uniform five-level TIL progressions into the full range of thirty-six discrete TRL_{Start} to TRL_{End} values. These improvements profoundly expand and transform the gross initial twenty-five-point TIL/SHL cost forecasts into a nine-thousand-point high-definition rendering of the R&D landscape.

The risk-adjusted parametric cost models, in concert with the proposed standard development framework with integrated milestone cost benchmarking, advance a unique estimating capability for conceptual-stage technology and system development that allows program executives, project managers, and investment decision-makers to better plan, forecast, and budget overall development-phase costs. These models can provide another estimating method as a check on other estimating techniques at a minimum, but they can also help long-term planning and potentially reduce the risk of cost overruns, milestone breaches, and wasteful program cancellations and the opportunity costs of valuable alternative resource investments.

The extensive enhancement to the fidelity of the first-generation development models, along with the development milestone cost benchmarking and other applied techniques from this research, yield improved capabilities for estimating conceptual-stage development.

Future Considerations

The expansion and enrichment of useful macro-parameters should continue to evolve early-stage development estimating. This could take many forms, including the addition of other TA categories and larger project data sets for all the key macro-parameters. Considerations for extending TRLs to level 10, mapped to MRL level 10 at FOC, as well as isolating and creating a "system-of-systems" level 6 in the SHL scale, also deserve consideration.

Composite system readiness and integration measures using as IRL, SRL, MRL, programmatic readiness level (PRL), and sustainment maturity level (SML) focused on various facets of maturity also hold potential to complement TRL-based macro-parametric forms of technology and system estimating. SRL and IRL measures, especially as modeled by Sauser et al.,[39] may add greater value to development-phase estimating since they consider both technology and broader systems development dimensions, including critical integration requirements.

The two largest underlying drivers of cost, schedule, and risk for any development are nearly always measures of project scale and complexity (i.e., both technological and overall system). Project scale is

[39] Sauser et al., *Development of Systems Engineering Maturity Models and Management Tools.*

effectively embodied by SHL, but a comprehensive measure of complexity might provide a dimension with the greatest potential to improve modeling utility. However, complexity is affiliated with a variety of the underlying dimensions and attributes of TIL, SHL, RD^3, and TA macro-parameters. These characteristics are therefore already inherent in the estimating methods included with this analysis, so a technique to remove redundant complexity in applying an explicit overall system measure would need to be developed and applied.

The DoD, civil agencies including NASA, major research institutions, and industry technology leaders are beginning to more broadly endorse developing, measuring, and capturing standardized forms of macro-parameters to enhance project planning, estimating, and performance measurement. Several important papers have expounded on the need to use readiness and integration measures early in the development process. Among these are two *Defense Acquisition Research Journal (ARJ)* papers.[40] Other significant work on this topic includes the Conference on Systems Engineering Research (CSER) *Procedia Computer Science* papers by Gove and Uzdzinski[41] and Atwater and Uzdzinski.[42] These various measures hold substantial promise to advance parametric estimating capabilities, but more project-level cost and schedule information will be needed for their application in resource planning and investment decision-making to reach its potential. Governmental, industrial, and institutional organizations with significant development program or project history may have the greatest opportunity to contribute to methodology advancements and impact progress of the estimating discipline through expanded sharing of project technical and cost data.

[40] Ross, "Application of System and Integration Readiness Levels"; and Eder, Mazzuchi, and Sarkani, "Beyond Integration Readiness Level (IRL)."

[41] Gove and Uzdzinski, "A Performance-Based System Maturity Assessment Framework."

[42] Atwater and Uzdzinski, "Wholistic Sustainment Maturity."

Appendix A Macro-Parameter Definitions

This appendix includes definitions for technology readiness level (TRL), system hierarchy level (SHL), system readiness level (SRL), integration readiness level (IRL), and manufacturing readiness level (MRL).

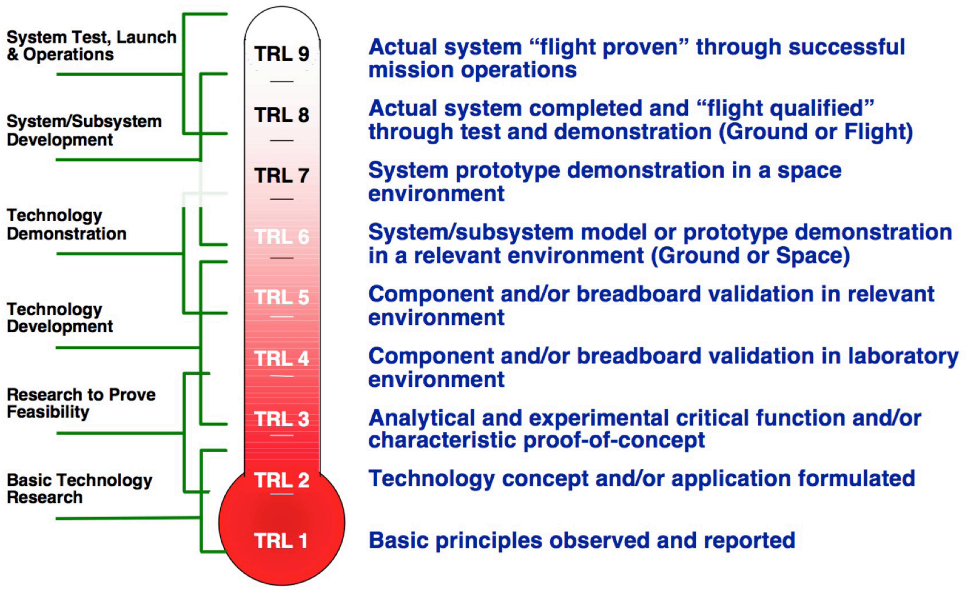


Figure A-1. NASA TRL Definitions

Table A-1. NASA SHL Definitions

No.	Tier	Definition	Example
5	System	An integrated set of constituent elements that are combined in an operational or support environment to accomplish a defined objective	A spacecraft or launch vehicle stage
4	Subsystem	A portion of a system	A satellite's propulsion system or launch vehicle's propulsion system
3	Assembly	A set of components (as a unit) before they are installed to make a final product	A satellite's thruster or launch vehicle's engine turbomachinery
2	Component/part	A portion of an assembly	A satellite's propellant valve or a launch vehicle's engine injector
1	Hardware/material	An item or substance used to form a component	Alloy, polymer, screws, bolts, pipes, semiconductor chips

Source: Adapted from Cole et al., *Technology Estimating.*

Numbers in the first column are inverted from the original table to correspond to the progressive ordinal numbers necessary to perform the analysis.

Table A-2. SRL Definitions

Level	Name	Definition
5	Operations and Support	Execute a support program that meets operational support performance requirements and sustains the system in the most cost-effective manner over its total life cycle.
4	Production and Development	Achieve operational capability that satisfies mission needs.
3	System Development and Demonstration	Develop a system or increment of capability; reduce integration and manufacturing risk; ensure operational supportability; reduce logistics footprint; implement human systems integration; design for producibility; ensure affordability and protection of critical program information; and demonstrate system integration, interoperability, safety, and utility.
2	Technology Development	Reduce technology risks and determine appropriate set of technologies to integrate into a full system.
1	Concept Refinement	Refine initial concept. Develop system/technology development strategy.

Source: Sauser et al., "From TRL to SRL."

Table A-3. IRL Definitions

Level	Definition
7	The integration of technologies has been ***verified and validated*** with sufficient detail to be actionable.
6	The integrating technologies can ***accept, translate, and structure information*** for its intended application.
5	There is sufficient ***control*** between technologies necessary to establish, manage, and terminate the integration.
4	There is sufficient detail in the ***quality and assurance*** of the integration between technologies.
3	There is ***compatibility*** (i.e., common language) between technologies to orderly and efficiently integrate and interact.
2	There is some level of specificity to characterize the ***interaction*** (i.e., ability to influence) between technologies through their interface.
1	An ***interface*** (i.e., physical connection) between technologies has been identified with sufficient detail to allow characterization of the relationship.

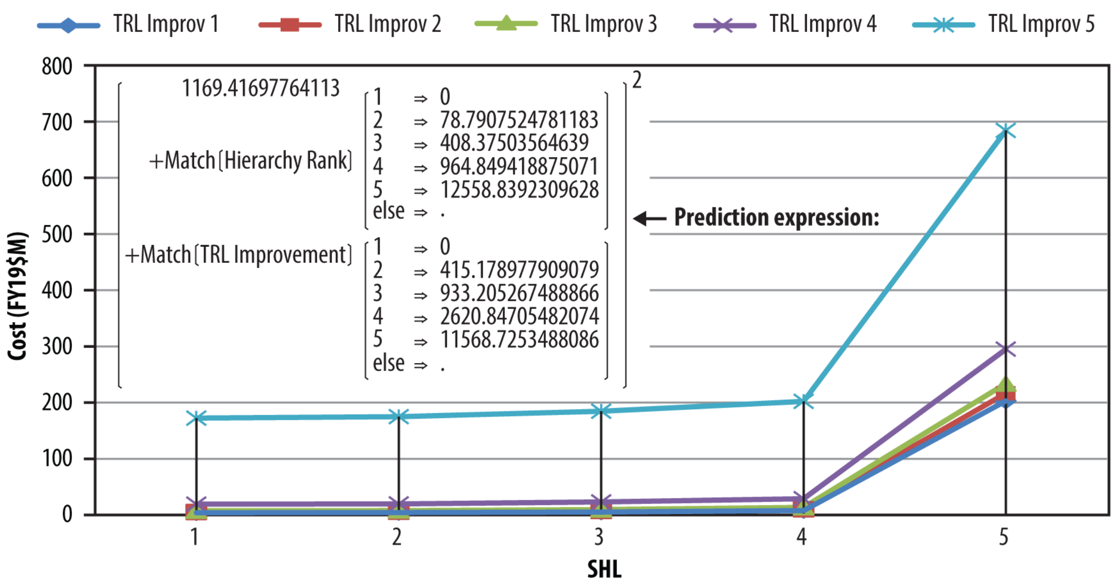
Table A-4. MRL Definitions

Level	Definition
1	Basic manufacturing implications identified.
2	Manufacturing concepts identified.
3	Manufacturing proof of concept developed.
4	Capability to produce the technology in a laboratory environment.
5	Capability to produce prototype components in a production-relevant environment.
6	Capability to produce a prototype system or subsystem in a production-relevant environment.
7	Capability to produce systems, subsystems, or components in a production-representative environment.
8	Pilot line capability demonstrated. Ready to begin low-rate production.
9	Low-rate production demonstrated. Capability in place to begin full-rate production.
10	Full-rate production demonstrated and lean production practices in place.

Source: Office of the Secretary of Defense Manufacturing Technology Program and the Joint Service/Industry MRL Working Group, *Manufacturing Readiness Level (MRL) Deskbook*.

Appendix B Highest-Performing First-Generation Cost Model

This appendix illustrates the highest-performing model—the technology readiness level (TRL) improvement (TIL)/system hierarchy level (SHL) cost model—from the previous study.[43] The study evaluated several hundred cost curve fit and regression models.



	1,465,102	1,669,178	2,667,033	4,880,072	201,910,868
	2,690,085	2,964,254	4,255,307	6,963,384	214,308,183
	4,736,434	5,098,058	6,754,939	10,080,685	230,294,450
	15,391,037	16,037,575	18,886,263	24,224,271	286,362,944
	173,836,568	175,993,723	185,161,369	201,168,389	685,592,966

Costs were escalated to fiscal year 2019 dollars. Color-coded TIL lines do not represent continuous functions but are shown to illustrate the progression of costs within and across SHLs and TILs. Uncertainty probability density functions (PDFs) for each TIL/SHL project data point were also constructed using lognorm functions in the original study.

Figure B-1. First-Generation Model No. 9—Mean Total Cost vs. *f* [TIL + SHL]

43 Alexander, "Parametric Cost and Schedule Modeling for Early Technology Development."

Appendix C Relationship Screening

Total project cost data were parsed into the range of research and development degree of difficulty (RD^3) levels and technology area (TA) categories for the two data sets to perform initial data relationship screening between the RD^3 and TA predictors and the cost response variable. Some outliers from the samples (1.1% of TA project data and 2.3% of RD^3 data) were filtered, and the remaining project cost statistics for the deconstructed RD^3 level and TA categorical data are shown in Tables C-1 and C-2.[44] Linear regressions of cost versus RD^3 produce adjusted coefficient of determination (R^2 Adj.) values over 0.7 and cost versus TA regression R^2 Adj. exceeding 0.8, both implying that relatively durable relationships may exist.

Table C-1. RD^3 Total Project Data Cost Statistics

Level	No. Records	Cost (FY19$)		
		Mean	Median	Standard Deviation
1	17	18,072,037	9,799,623	18,436,153
2	153	32,399,635	13,242,734	52,082,945
3	174	44,543,794	19,864,101	111,674,939
4	76	56,739,467	26,485,469	85,282,868
5	6	79,677,118	57,605,894	73,348,093

Table C-2. TA Total Project Data Cost Statistics

TA	No. Records	Cost (FY19$)		
		Mean	Median	Standard Deviation
1 Launch Propulsion Systems	159	29,482,594	896,999	125,232,312
2 In-Space Propulsion Technologies	111	22,420,479	1,122,812	68,386,702
3 Space Power and Energy Storage	229	21,455,560	800,408	136,454,438
4 Robotics, Telerobotics, Autonomous Systems	73	25,936,013	13,246,144	42,926,634
5 Communication and Navigation	182	8,439,804	972,011	24,215,606
6 Human Health, Life Support, Habitation Systems	224	53,192,277	15,891,281	87,320,195
7 Human Exploration Destination Systems	59	48,878,481	26,485,469	62,394,548
8 Science Instruments, Observatories, Sensor Systems	123	8,934,078	926,115	39,299,914
9 Entry, Descent, and Landing Systems	15	356,640,735	25,965,543	668,318,491
10 Nanotechnology	24	2,762,815	401,754	5,452,029
11 Modeling, Simulation, Information Tech	95	39,777,986	1,491,313	176,746,995
12 Materials, Structures, Mechanical Systems, Manufacturing	229	11,845,815	803,508	33,782,225
13 Ground and Launch Systems Processing	23	50,093,679	13,529,154	126,535,384
14 Thermal Management Systems	85	19,242,667	2,648,547	37,251,256
15 Aeronautics	99	5,904,203	393,329	16,990,139

[44] Costs presented in this report for NASA Technology Cost and Schedule Estimating (TCASE) source data and first-generation model results have been escalated to fiscal year 2019 dollars using the Research, Development, Test, and Evaluation (RDT&E) Appropriation TY$ indices from the current Naval Center for Cost Analysis Joint Inflation Calculator.

Although two of the RD^3 categories contained somewhat limited project sample sizes, project cost statistics demonstrate a direct and progressive incremental relationship to RD^3 across all five levels. Cost statistics for TA categories produced mixed results, with five categories being apparent outliers. TAs 5, 8, 9, 10, and 15 exhibited very low or high mean cost values versus the overall TA project sample; TAs 1, 3, 9, 11, and 13 contained extensive cost ranges with very significant standard deviations; and most TAs also contained very large coefficients of variation (CVs). The large cost ranges and variability across categories is primarily due to the fact that each TA category spans a full range of project scale, complexity, and maturity and does not reflect any graduated measurement levels with respect to cost. A closer examination of the underlying project data reveals that some of low and high central value behavior can also be largely attributed to limited sample sizes and a focus of similar small- or large-scoped projects in some categories. The reason for these project size concentration anomalies is unclear, but they are possibly related to repetitive-type development efforts, project budgeting, or execution policies or practices for particular technical areas. They may also simply reflect the way cost data were reported, captured, or characterized by individuals providing historical project information for certain TA categories in the TCASE database.

As noted for the data investigation described in the main report, there were not enough projects containing all four variables to produce comprehensive multiple-regression models; however, other techniques were explored to leverage cost impacts from the additional parameters. To effectively apply these techniques, further screening tests and analysis were first conducted to look for multicollinearity and residual autocorrelation among the TIL, SHL, RD^3, and TA independent variables. First, regression analysis between combinations of the four independent parameters versus the cost response was performed. These tests produced favorable results, with a Durbin Watson (DW) statistic in the range of 1.83 to 2.13 for all regressions and about 92% of parameter category levels possessing a variance inflation factor (VIF) less than 4 and an average overall category-level VIF of ~2.4 across models. For more detail on the DW and VIF statistics and specific results of tests conducted, refer to Appendix D.

Last, the absolute value of correlation coefficients assessed between independent parameters fell to under 0.1 for 85% of the category combinations, between 0.1 to 0.2 for 13% of cases, and between 0.2 to 0.4 for the remaining 2% of cases. Therefore, all three indicators—DW statistic, VIF, and correlation coefficients—suggest no noteworthy residual autocorrelation or multicollinearity between the four predictor variables. This supports their independence, and bringing the two additional parameters into the analysis should therefore not introduce any significant common influential effects or overlapping causal factors with respect to cost.

Appendix D Independent-Variable Multicollinearity and Residual Autocorrelation Testing

Variance inflation factors (VIFs) and regression correlation coefficients (CCs) between independent predictor variables were assessed as indicators of potential multicollinearity. To check for autocorrelation among regression residuals used in independent-variable screening, the Durbin Watson (DW) statistic was also evaluated. A VIF detects multicollinearity in regression analysis. Multicollinearity occurs when there is correlation between predictors (i.e., independent variables) in a model, and its presence can adversely affect regression results. The VIF estimates how much the variance of a regression coefficient is inflated due to multicollinearity in the model. VIFs begin at 1 and move upward. The numerical value for VIF indicates (in decimal form) what percentage the variance (i.e., the standard error squared) is inflated for each coefficient. A rule of thumb for interpreting the VIF is, in general, a VIF of 1 is not correlated, a VIF between 1 and 5 is moderately correlated, and a VIF greater than 5 is highly correlated.[45]

The DW statistic tests the null hypothesis that the residuals from an ordinary least-squares regression are not autocorrelated. The statistic ranges in value from 0 to 4. A value near 2 indicates non-autocorrelation, a value toward 0 indicates positive autocorrelation, and a value toward 4 indicates negative autocorrelation. A rule of thumb is that test statistic values in the range of 1.5 to 2.5 are relatively normal, and even those in the 0.5 to 3 range are generally considered acceptable.[46]

To test for multicollinearity in the model forecasts of cost response to the four independent-variable terms,[47] multiple-regression CCs and VIFs were assessed. Multiple-regression models were formulated to perform these tests between the RD^3/TIL/SHL, TA/TIL/SHL, and RD^3/TA independent cost variables.

For the RD^3/TIL/SHL independent cost variable multiple regressions, of the thirty-two independent-variable term combinations, twenty CCs (63%) fell in the –0.1 to 0.1 range, eleven (34%) fell in the –0.2 to –0.1 or 0.1 to 0.2 range, and one (3%) fell in the 0.3 to 0.4 range. VIFs for the various RD^3/TIL/SHL terms ranged from 1.1 to 3.5, with an average of 1.68, 67%, falling under 2.0. For the multiple regressions between the TA/TIL/SHL independent cost variables, of the one hundred forty-four independent-variable combinations, one hundred twenty-six CCs (88%) fell in the –0.1 to 0.1 range, fifteen (1 %) fell in the –0.2 to –0.1 or 0.1 to 0.2 range, and three (2%) fell in the 0.2 to 0.3 or –0.2 to –0.3 range. A total of 95% of VIFs for the various TA/TIL/SHL terms fell under 4.0, with 1 TIL term at 5.3 and an average VIF of 2.51. Finally, to test for multicollinearity between the two new variables introduced, RD^3 and TA, using multiple regressions, for the forty-eight independent-variable cost term combinations, forty-four CCs (92%) fell in the –0.1 to 0.1 range, three CCs (6%) fell in the –0.2 to –0.1 or 0.1 to 0.2 range, and one CC (2%) in fell the 0.2 to 0.4 range. A total of 81% of VIFs for the various RD^3/TA terms fell under 4.0, with an average of 3.22.

For the same cost regressions, these tests produced DW statistic values of 2.1 for RD^3/SHL/TIL variables, 2.1 for TA/SHL/TIL variables, and 1.8 for the RD^3/TA variables. These results suggest that no autocorrelation issues are evident.

[45] Penn State Eberly College of Science, "What is a Variation Inflation Factor?"

[46] Glen, "Durbin Watson Test & Test Statistic."

[47] Technology readiness level (TRL) improvement level (TIL) and system hierarchy level (SHL) from the original parametric models with the newly introduced research and development degree of difficulty (RD^3) and technology area (TA) parameters.

Appendix E Data Types

The development cost response variable applied in this analysis is a continuous quantitative variable. Technology readiness level (TRL) improvement level (TIL), system hierarchy level (SHL), and research and development degree of difficulty (RD^3) predictor variables are discrete ordered categorical values. Technology area (TA) is simply a list of categorical class values. Categorical variables that have two or more incremental levels are often measured on an ordinal scale so that the characteristic or property described by the category levels or class (i.e., 1 through *K*) can be considered as ordered, but not as equally spaced. This is the case with TRL, SHL, and RD^3, as determination of those levels can involve various subjective criteria that span a wide range of scale and complexity both between and within categories.

Traditional linear regression models, however, make no distributional assumptions about the independent predictor variables. Consequently, ordinal variables must be interpreted carefully when attempting to fit a continuous function, especially if large or random interval variance is possible between class rankings. Fortunately, statistical analysis tools, such as SAS JMP used for the first-generation TIL/SHL models,[48] solve this potential issue by employing a regression technique that leverages response to the ordinal *interval* values. Further, since the dependent cost variable response in this analysis is being assessed at the discrete ordinal levels only and not as continuous functions, that completely neutralizes any concerns that a possible lack of a natural ordinal interval size structure could impact results.

Historically, ordinal *response* variables have been substantially investigated in regression modeling, but less research has been performed on ordinal *predictors*. Anderson[49] notes two major types of ordinal categorical predictor variables: "grouped continuous variables" and "assessed ordered categorical variables." Various techniques to model ordinal predictor variables have been suggested (e.g., quadratic penalization regression, ridge reroughing, and five-point Likert scales),[50] but no definitive method or approach has been identified in the literature.

Ordinal qualitative measures nevertheless are ordered, and for technologies, this progression can be driven by certain underlying development structure, known or unknown, such as architecture, functionality, complexities, common development processes, and support activities. As a result, a quantitative relationship can exist, and it may be modeled between an ordinal scale (or the variability in such a scale) and continuous numeric parameters. Since this relationship is not necessarily or even likely to be linear in nature, data transformations, coefficient/correction/adjustment factors, and nonlinear functions are often applied to normalize ordinal values to account for the variability in cost and schedule modeling.[51]

48 Alexander, "Parametric Cost and Schedule Modeling for Early Technology Development."

49 Anderson, "Regression and Ordered Categorical Variables."

50 Stauner, "Effect of Two Demographic IVs"; Gertheiss and Tutz, "Penalized Regression with Ordinal Predictors"; and Berry, *Understanding Regression Assumptions.*

51 Malone et al., "Application of TRL Metrics to Existing Cost Prediction Models"; Smoker and Smith, "System Cost Growth Associated with Technology-Readiness Level"; and Conrow, "Estimating Technology Readiness Level Coefficients."

Appendix F Curve Fit Methodology

All probability density function (PDF) cost curve fits for this analysis were produced using Palisade's @RISK software. Sample data and calculated distribution data values were "fit" to a library of possible probability-based distribution functions using the tool's distribution fitting utility and standard fit measurement techniques. More than twenty functions (or families of functions) are assessed, including beta, chi-square, Erlang, exponential, gamma, inverse Gaussian, Lévy, loglogistic, lognorm, Pareto, Pearson, program evaluation and review technique (PERT), Raleigh, triangular, uniform, Weibull, and several others. The distribution fit utility is applied to down-select higher-performing functions using the following commonly applied goodness-of-fit statistical significance methods/techniques:

- Akaike information criterion
- Bayesian information criterion
- Kolmogorov–Smirnov
- Anderson–Darling
- Chi-square tests

A lower bound of zero and unlimited upper bound were input as search range criteria to best replicate the highly right-skewed cost functions involved, which are common to cost and schedule behavior and related early-life-cycle estimating methodologies. Functions with best-result consensus across these techniques are selected, considering key statistical metrics versus the sample data, such as fit of the estimate mean, a commonly applied budget planning and forecast range between the fiftieth (i.e., median), seventieth, and eightieth percentile, the standard deviation, and distribution shape characteristics (skewness, kurtosis, etc.). The curve fits produced appropriately reflect the highly uncertain environments; they have relatively wide dispersion and large standard deviations around the central datum, which is expected because of the high level of unknowns in conceptual stages of development.

Appendix G Investigation of Project Data Sample Equivalence

A data relationship between either the research and development degree of difficulty (RD^3) or technology area (TA) sample project cost data and the system hierarchy level (SHL)/technology readiness level improvement level (TIL) project cost data can be established via means translations (i.e., factor of the sample means). In addition to the data groups coming from a common population with a small difference in sample means and other empirical evidence described below, equivalence tests were also applied to demonstrate a degree of sample equivalence. These equivalence tests include the two one-sided test (TOST) and the Welch's t-test.

In a classical hypothesis test, the goal is to reject the null hypothesis of equality. As part of an equivalence test, however, the goal is to validate the equivalence between two samples. TOST is a test of equivalence based on the classical t-test used to test the hypothesis of equality between two means. Therefore, equivalence tests differ from standard t-tests in that the null and alternative hypothesis are reversed:

- Null hypothesis (H_0)—The difference between the means is outside your equivalence interval. The means are not equivalent.
- Alternative hypothesis (H_1)—The difference between the means is inside your equivalence interval. The means are equivalent.

The TOST equivalence test can be used to validate the equivalence of the means of two groups by demonstrating that they do not differ by more than a specified margin. When the sample sizes and variances of two groups are unequal (nonparametric), such as with the SHL/TIL, RD^3, and TA data samples being compared, Welch's t-test for unequal variance (also known as the Satterthwaite's test, the Smith/Welch/Satterthwaite test, the Aspin–Welch test, or the unequal variances t-test) is also commonly used to test sample equivalence.[52]

Welch's t-test is more robust than the Student's t-test and maintains type I error rates that are close to nominal for unequal variances and for unequal sample sizes,[53] as is the case for this analysis. Welch's t-test also remains robust for skewed distributions and large sample sizes,[54] again present in this investigation. When group sizes are unequal, a small proportion of outlying observations is commonly trimmed to alleviate problems related to the skewness in underlying distributions. This was first proposed by Tukey and McLaughin (1963) and later combined with Welch's test by Yuen (1974).[55] The resulting trimmed Welch's test is resistant to outliers and alleviates some of the problems that occur because of skewness in the underlying distributions. In applying this method, G represents the percent of data trimmed, which generally less than 25% and often in the 5% to 10% range.[56]

---

52 NCSS, "Equivalence Tests for Two Means (Simulation)"; Ruxton, "Uunequal Variance *t*-Test"; and Lakens, "Equivalence Tests."

53 Ruxton, "Uunequal Variance *t*-Test"; and Lakens, "Equivalence Tests."

54 Fagerland, "t-Tests, Non-parametric Tests, and Large Studies."

55 NCSS, "Equivalence Tests for Two Means (Simulation)."

56 NCSS, "Equivalence Tests for Two Means (Simulation)."

Sample equivalence between the cost means for the SHL/TIL data set and the corresponding RD^3 and TA parameter samples was tested using both the trimmed TOST and Welch's trimmed t-test, assuming unequal variances, in the SAS JMP software. The three data sets involved with the analysis each contain sufficiently large sample sizes with a raw number of observations (n_i) = 221, 425, and 1,750 each for the SHL/TIL sample (no. 1), RD^3 sample (no. 3), and TA sample (no. 2), respectively. All extracted data come from a common development project database population (the NASA Technology Cost and Schedule Estimating, or TCASE, database) and include a degree of individual project commonality or overlap. The extreme cost data ranges and variance (coefficients of variation, or CVs, in the 1.7 to 3.8 range) within the project data can make equivalence testing more challenging. The actual G values for the final samples tested both fell well within the acceptable range: G = 2.8% (=55/1,971 for n = 221 + 1,750 = 1,971) for the SHL/TIL versus TA stacked project sample and G = 8.4% (=55/656 for n = 221 + 425 = 656) for the SHL/TIL versus RD^3 stacked project sample. Because of the extremely large overall project population cost variance, the sample data equivalence tests were performed at an alpha level of 0.10. Results of the trimmed TOST and Welch's trimmed test, along with other evidence like sample density plot overlays, are provided in Figures G-1 and G-2.

For both sample data set comparisons, TOST test p-values are smaller than alpha (0.1). Therefore, the difference in population means is located within the lower and upper confidence thresholds, and the sample means are practically equivalent. Both Welch tests also indicate that the null hypothesis can be rejected as the F-ratio is small and Prob > F is high, and therefore no significant differences in the samples are detected.[57] Other empirical evidence, such as the very small percentage difference in sample means (only 0.25% variability between the SHL/TIL versus RD^3 trimmed samples and 1.4% for the SHL/TIL versus TA trimmed samples) and the comparison and composition of density plots, also provides rational support to demonstrate that the parameter sample means are similar enough to practically represent the same population. Based on this preponderance of evidence, it is therefore reasonable to extend RD^3 and TA influence on the SHL/TIL parametric models by applying statistical index values between sample means.

[57] JMP Statistical Discovery, "Description of the Welch's Test Report"; GraphPad Software, "Interpreting Results"; and mdawson69, "How Do You Interpret Welch's Test Results?"

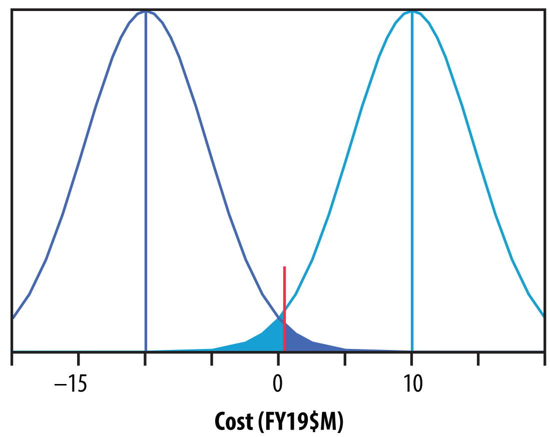
Means and Standard Deviations

Level	No.	Mean	Standard Deviation	Standard Error of the Mean	Lower 90%	Upper 90%
1 (SHL/TIL)	186	29,304,292	75,771,209	5,555,814.6	20,119,798	38,488,786
3 (RD^3)	416	29,717,115	37,905,496	1,858,469.9	26,653,365	32,780,865

Test	F-Ratio	DFNum	DFDen	p-Value
Bartlett	136.0031	1	0	<0.0001
F test 2-sided	3.9958	185	415	<0.0001

Practical Equivalence between RD^3 (No. 3) and SHL/TIL (No. 1) Samples

Null Hypothesis	DF	t-Ratio	p-Value
Mean difference ≥ 1,000,000	600	−2.0674	0.0196
Mean difference ≤ −1,000,000	600	2.245442	0.0126
Max over both			0.0196*

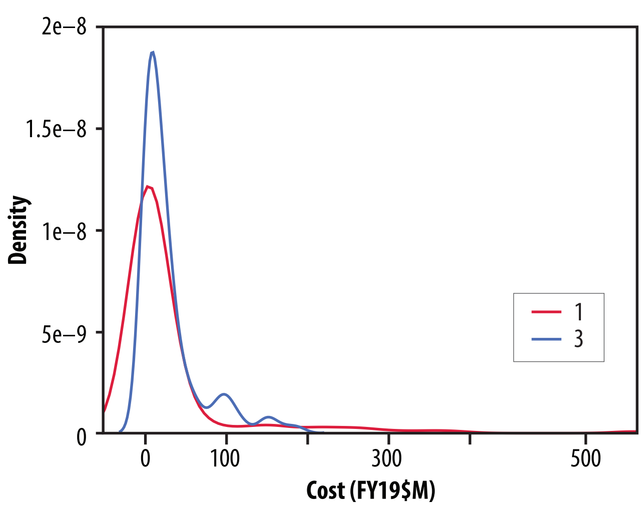


Welch's Test

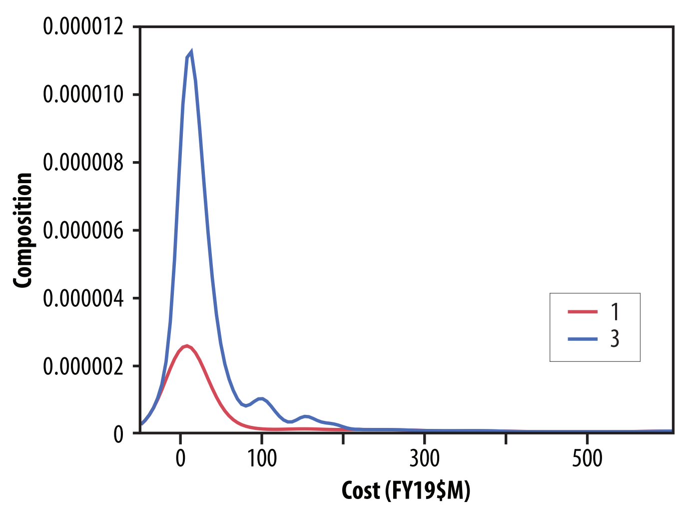
F-Ratio	DFNum	DFDen	Prob > F
0.0050	1	227.45	0.9439

Welch's ANOVA testing; means equal, allowing standard deviations not equal

Compare Densities



Composition of Densities



Den, denominator; DF, degrees of freedom; Num, numerator.

Figure G-1. Trimmed Equivalence Tests of SHL/TIL vs. RD^3 Sample Mean Cost Data

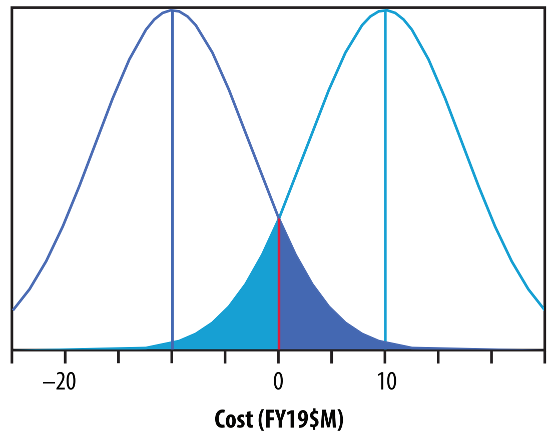
Means and Standard Deviations

Level	No.	Mean	Standard Deviation	Standard Error of the Mean	Lower 90%	Upper 90%
1 (SHL/TIL)	221	27,052,654	77,816,153	5,234,480.9	18,406,290	35,699,018
2 (TA)	1,703	27,059,297	104,306,493	2,527,574.8	22,899,542	31,219,052

Test	F-Ratio	DFNum	DFDen	p-Value
Bartlett	28.7506	1	0	<0.0001
F Test 2-sided	1.7967	1,702	220	<0.0001

Practical Equivalence between TA (No. 2) and SHL/TIL (No. 1) Samples

Null Hypothesis	DF	t-Ratio	p-Value
Mean difference ≥ 1,000,000	1,922	−1.37535	0.0846
Mean difference ≤ −1,000,000	1,922	1.377176	0.0843
Max over both			0.0846

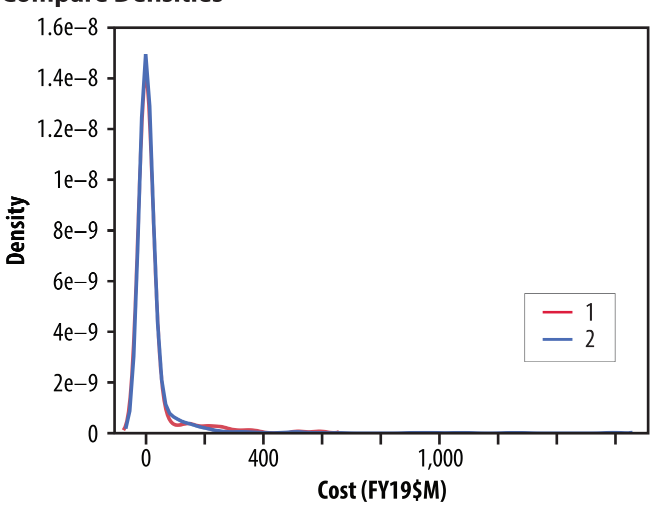


Welch's Test

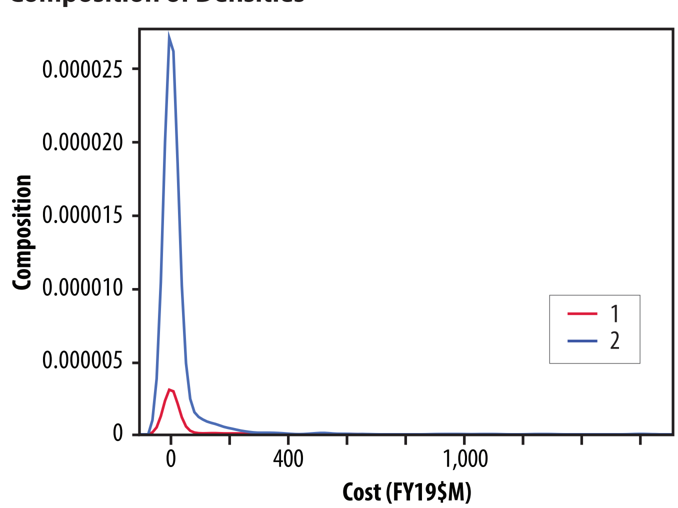
F-Ratio	DFNum	DFDen	Prob > F
0.0000	1	332.22	0.9991

Welch ANOVA testing; means equal, allowing standard deviations not equal

Compare Densities



Composition of Densities



Den, denominator; DF, degrees of freedom; Num, numerator.

Figure G-2. Trimmed Equivalence Tests of TIL/SHL vs. TA Sample Mean Cost Data

Appendix H Mean Cost Index Curve Fit Probability Density Function Formulas and Plots

This appendix includes mean cost index (MCI) curve fit probability density functions (PDFs) for research and development degree of difficulty (RD^3) and technology area (TA).

Tables H-1 and H-2 PDFs are consistent with the right-skewed lognormal-, gamma-, Weibull-, and betaPERT-type PDFs commonly recommended for estimating uncertainty in the *Joint Agency Cost Schedule Risk and Uncertainty Handbook* (JACSRUH).[58] Table H-2 is provided for analysis demonstration purposes only; these PDFs are not recommended for application in modeling for reasons explained in the section on the MCI method in the main report.

Table H-1. RD^3 MCI Curve Fit PDFs

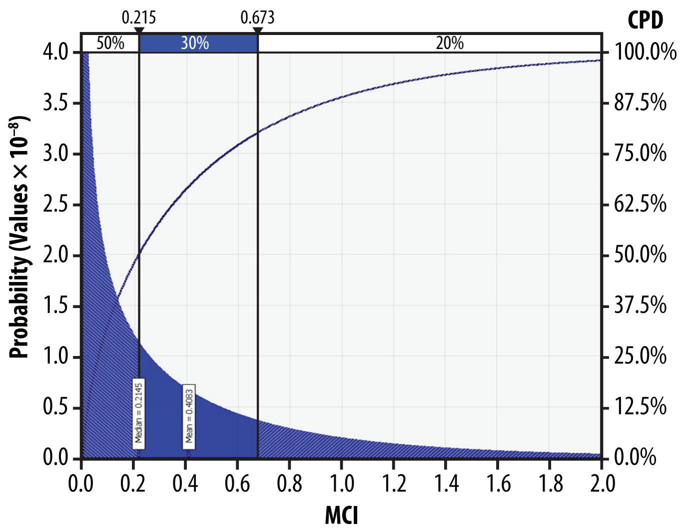
Level	PDF Type	@RISK PDF Formula
1	Gamma	=RiskGamma(0.59877,0.68192,RiskName(“RD3 Lvl 1 MCI”))
2	Lognorm	=RiskLognorm(0.84662,2.1681,RiskName(“RD3 Lvl 2 MCI”))
3	Pearson6	=RiskPearson6(1.1572,1.7721,0.71302,RiskName(“RD3 Lvl 3 MCI”))
4	Gamma	=RiskGamma(0.71451,1.9062,RiskName(“RD3 Lvl 4 MCI”))
5	Gamma	=RiskGamma(1.3688,1.394,RiskName(“RD3 Lvl 5 MCI”))

Table H-2. TA MCI Curve Fit PDFs

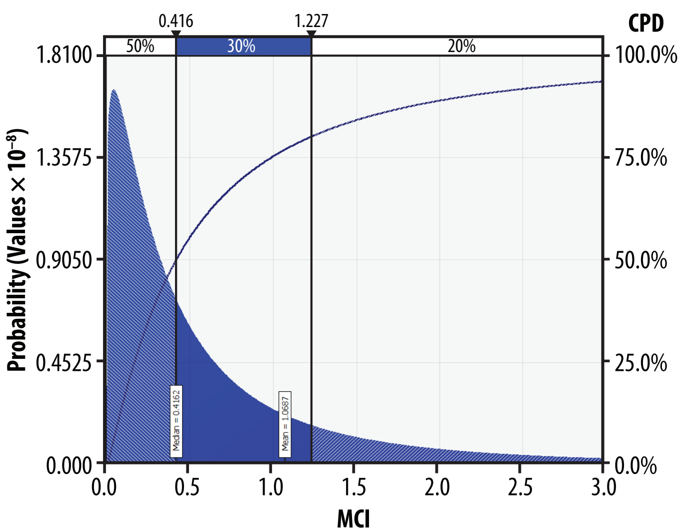
TA	PDF Type	@RISK PDF Formula
1 Launch Propulsion Systems	Fréchet	=RiskFrechet(0,0.016039,0.60073,RiskName(“TA1 Mean Cost Index”))
2 In-Space Propulsion Technologies	Lognorm	=RiskLognorm(1.0673,18.846,RiskName(“TA2 Mean Cost Index”))
3 Space Power and Energy Storage	Fréchet	=RiskFrechet(0,0.014939,0.63461,RiskName(“TA3 Mean Cost Index”))
4 Robotics, Telerobotics, Autonomous Systems	Gamma	=RiskGamma(0.33743,2.8459,RiskName(“TA4 Mean Cost Index”))
5 Communication and Navigation	Lognorm	=RiskLognorm(0.32008,2.0834,RiskName(“TA5 Mean Cost Index”))
6 Human Health, Life Support, Habitation Systems	Weibull	=RiskWeibull(0.57905,1.2756,RiskName(“TA6 Mean Cost Index”))
7 Human Exploration Destination Systems	Gamma	=RiskGamma(0.50991,3.5492,RiskName(“TA7 Mean Cost index”))
8 Science Instruments, Observatories, Sensor Systems	Loglogistic	=RiskLoglogistic(0,0.030355,0.8796,RiskName(“TA8 Mean Cost Index”))
9 Entry, Descent, and Landing Systems	Lévy	=RiskLevy(0,0.55536,RiskName(“TA9 Mean Cost Index”))
10 Nanotechnology	Lévy	=RiskLevy(0,0.0070262,RiskName(“TA10 Mean Cost index”))
11 Modeling, Simulation, Information Tech	Lognorm	=RiskLognorm(1.156,16.677,RiskName(“TA11 Mean Cost index”))
12 Materials, Structures, Mechanical Systems, Manufacturing	Fréchet	=RiskFrechet(0,0.015598,0.59836,RiskName(“TA12 Mean Cost index”))
13 Ground and Launch Systems Processing	Pareto2	=RiskPareto2(0.49689,1.1179,RiskName(“TA13 Mean Cost index”))
14 Thermal Management Systems	FatigueLife	=RiskFatigueLife(0,0.11763,3.1833,RiskName(“TA14 Mean Cost Index”))
15 Aeronautics	Invgauss	=RiskInvgauss(0.21861,0.008011,RiskName(“TA15 Mean Cost index”))

[58] Naval Center for Cost Analysis, *Joint Agency Cost Schedule Risk and Uncertainty Handbook.*

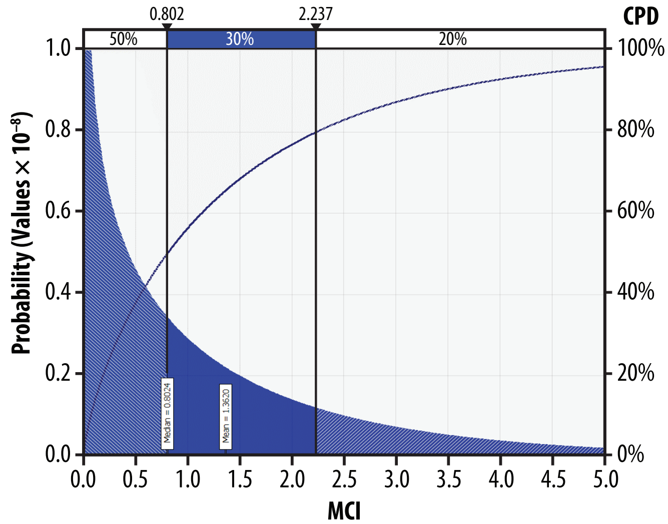
The following RD³ MCI curve fit PDFs for levels 1, 3, 4, and 5 are plots of the continuous functions, with the *x* axis representing the MCI values. These functions can express larger concentrations as they approach 0; however, they are used because they closely replicate the typical range of interest in the sample data. This area of interest is the planning range between the fiftieth to eightieth percentiles generally applied in budgeting and investment decision-making. For a more detailed explanation of the curve fit methodology, refer to Appendix F.



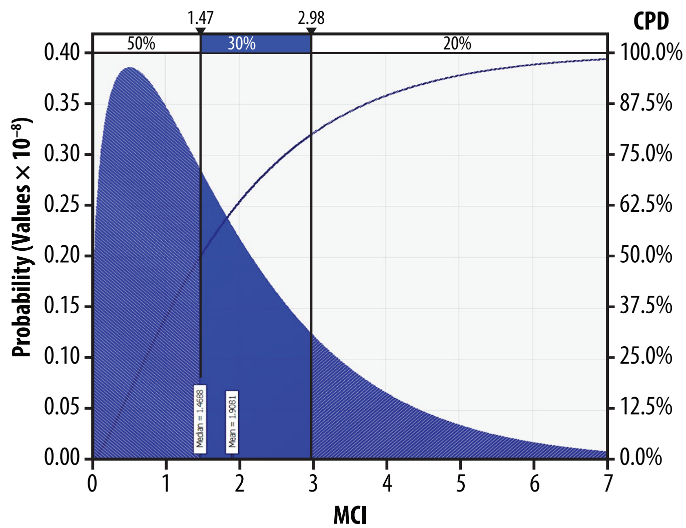
Figures H-1. MCI Curve Fit PDFs for RD³ Level 1



Figures H-2. MCI Curve Fit PDFs for RD³ Level 3



Figures H-3. MCI Curve Fit PDFs for RD^3 Level 4



Figures H-4. MCI Curve Fit PDFs for RD^3 Level 5

Appendix I Geometric Mean Curve Fit Method

Geometric means (GMs) are preferable over arithmetic means because the GM form tends to minimize or dampen the influence of extreme data points, such as large or small data values in the generally highly skewed data distributions predominant in the technology project sample data.[59] The GM formula[60] is represented as follows, where a series of n data (a_1, ... a_n) are compounded taking the nth root of the product:

$$\text{Geometric mean} = \left(\Pi_{i=1}^{n} a_i\right)^{1/n} = \sqrt[n]{a_1 a_2 a_3 \ldots \ldots a_n}\,.$$

Much of the study data includes skewed lognormal or "lognormal-like" distributions common to early-life-cycle cost data. With a true lognormal data set, the median and GM are identical, so with highly skewed data, they can provide a substantially better indication of central tendency than the arithmetic mean.[61]

This technique involves creation of composite functions of the independent variables by merging the uncertainty distributions of the selected predictor variables for each parametric combination in a GM. The blended impacts of individual tailored probability density function (PDF) cost curve fits for each independent parameter level are aggregated in a product (i.e., the GM) of their expected values sampling their individual values in Monte Carlo simulation. The highest-performing system hierarchy level (SHL) and technology readiness level (TRL) improvement level (TIL) PDF cost curve fits from the initial study[62] were used, along with newly developed research and development degree of difficulty (RD^3) level and technology area (TA) category cost PDF curve fits (refer to Table I-1 for the RD^3 cost curve fit PDFs). The GM of the project development cost included combinations of the four parameters is

$$\left(PDF_{SHL} \times PDF_{TIL} \times PDF_{RD3} \times PDF_{TA}\right)^{1/n},$$

where n represents the number of independent parameters actually applied (four in this equation). Any combination of two to four of the parameters can be modeled in simulation applying the $1/n$ root power. Monte Carlo simulation runs calculating the expected GM for the full range of curve fit PDF combinations across the four independent variables (SHL/TIL/RD^3/TA) were performed. Output from the simulations therefore represent a blended average of the three selected constituent macro-parameters.

Table I-1. RD^3 Project Sample Cost Data Curve Fit Functions (FY19$)

Level	PDF Form	PDF Formula
1	Gamma	RiskGamma(0.81109,22281302,RiskName("RD3 Lvl 1 (FY19$)"))
2	Lognorm	RiskLognorm(35352239,90532673.2,RiskName("RD3 Lvl 2 (FY19$)"))
3	Burr12	RiskBurr12(0,27822346,1.1144,1.4889,RiskName("RD3 Lvl 3 (FY19$)"))
4	Weibull	RiskWeibull(0.78281,48452157,RiskName("RD3 Lvl 4 (FY19$)"))
5	Erlang	RiskErlang(1,79677117,RiskName("RD3 Lvl 5 (FY19$)"))

59 Hu, "Simple Mean, Weighted Mean, or Geometric Mean?"; and Clark-Carter, "Geometric Mean."

60 Roenfeldt, "Better than Average."

61 McChesney, "You Should Summarize Data with the Geometric Mean."

62 Alexander, "Parametric Cost and Schedule Modeling for Early Technology Development."

Results, however, did not effectively capture the compound or aggregate impact of the independent parameters and predicted relatively low project costs with rather large residuals. This method was therefore abandoned as a viable option for estimating purposes.

Appendix J Estimating Methodology

This appendix illustrates the estimating methodology for the system hierarchy level (SHL)/technology readiness level (TRL) improvement level (TIL)/research and development degree of difficulty (RD^3) and the TIL/system hierarchy level (SHL)/technology area (TA) composite models.

In Table J-1, values for the three-parameter estimates represent the expected mean point estimate costs (in fiscal year 2019 dollars, FY19$) for the 125 possible three-parameter (SHL/TIL/RD^3) model combinations. Probability density (PDF) uncertainty distributions for each model are also available by running Monte Carlo simulation for the product of TIL/SHL regression model output (Table 1 and Appendix B) × the applicable RD^3 mean cost index (MCI) PDF functions in Table 6. To produce models including all four parameters, simply include another factor for the applicable TA MCI mean value from Table 5 in the product in the simulation (e.g., TIL/SHL mean × RD^3 PDF × TA MCI). This results in 1,250 possible four-parameter model variants (25 TIL/SHL × 5 RD^3s × 10 TAs). Finally, to adjust for actual TRL start and end states, use the adjustment factors found in the lower section of Table 12 to produce up to 9,000 possible model variants (36 TRL Start-End × 5 SHL × 5 RD^3 × 10 TA).

In Table J-2, values for the three-parameter estimates represent the expected mean point estimate costs (FY19$) for the 250 possible three-parameter (SHL/TIL/TA) model combinations. Costs are the product of TIL/SHL regression model output (Table 1 and Appendix B) × the applicable TA MCI values in Table 5. To produce models including all four parameters, simply include another factor for the applicable RD^3 MCI PDF from Table H-1 in the product in Monte Carlo simulation (e.g., TIL/SHL mean × TA MCI × RD^3 PDF). This results in 1,250 possible four-parameter model variants (25 TIL/SHL × 10 TAs × 5 RD^3s). Finally, to adjust for actual TRL start and end states, use the adjustment factors found in the lower section of Table 12 to produce up to 9,000 possible model variants (36 TRL Start-End × 5 SHL × 5 RD^3 × 10 TA).

Table J-1. SHL/TIL/RD^3 Composite Model Mean Project Costs (FY19$)

Model No. (SHL / TIL / RD^3)	Mean Project Pt. Estimate Cost (FY19$)
1/1/1	$ 598,201
1/1/2	$ 1,136,772
1/1/3	$ 1,566,194
1/1/4	$ 1,995,468
1/1/5	$ 2,795,560
2/1/1	$ 681,525
2/1/2	$ 1,295,115
2/1/3	$ 1,784,351
2/1/4	$ 2,273,421
2/1/5	$ 3,184,959
3/1/1	$ 1,088,950
3/1/2	$ 2,069,351
3/1/3	$ 2,851,059
3/1/4	$ 3,632,499
3/1/5	$ 5,088,966
4/1/1	$ 1,992,533
4/1/2	$ 3,786,448
4/1/3	$ 5,216,797
4/1/4	$ 6,646,658
4/1/5	$ 9,311,665
5/1/1	$ 82,440,208
5/1/2	$ 156,662,643
5/1/3	$ 215,842,718
5/1/4	$ 275,002,603
5/1/5	$ 385,266,128
1/2/1	$ 1,098,362
1/2/2	$ 2,087,237
1/2/3	$ 2,875,701
1/2/4	$ 3,663,896
1/2/5	$ 5,132,952
2/2/1	$ 1,210,305
2/2/2	$ 2,299,965
2/2/3	$ 3,168,787
2/2/4	$ 4,037,314
2/2/5	$ 5,656,093
3/2/1	$ 1,737,442
3/2/2	$ 3,301,693
3/2/3	$ 4,548,923
3/2/4	$ 5,795,728
3/2/5	$ 8,119,551
4/2/1	$ 2,843,150
4/2/2	$ 5,402,890
4/2/3	$ 7,443,858
4/2/4	$ 9,484,130
4/2/5	$ 13,286,834
5/2/1	$ 87,502,031
5/2/2	$ 166,281,719
5/2/3	$ 229,095,448
5/2/4	$ 291,887,745
5/2/5	$ 408,921,444

Model No. (SHL / TIL / RD^3)	Mean Project Pt. Estimate Cost (FY19$)
1/3/1	$ 1,933,886
1/3/2	$ 3,674,999
1/3/3	$ 5,063,248
1/3/4	$ 6,451,023
1/3/5	$ 9,037,589
2/3/1	$ 2,081,537
2/3/2	$ 3,955,583
2/3/3	$ 5,449,824
2/3/4	$ 6,943,554
2/3/5	$ 9,727,604
3/3/1	$ 2,758,042
3/3/2	$ 5,241,157
3/3/3	$ 7,221,030
3/3/4	$ 9,200,227
3/3/5	$ 12,889,099
4/3/1	$ 4,115,944
4/3/2	$ 7,821,604
4/3/3	$ 10,776,253
4/3/4	$ 13,729,894
4/3/5	$ 19,234,956
5/3/1	$ 94,029,224
5/3/2	$ 178,685,464
5/3/3	$ 246,184,767
5/3/4	$ 313,661,041
5/3/5	$ 439,424,841
1/4/1	$ 6,284,160
1/4/2	$ 11,941,905
1/4/3	$ 16,453,018
1/4/4	$ 20,962,592
1/4/5	$ 29,367,637
2/4/1	$ 6,548,142
2/4/2	$ 12,443,555
2/4/3	$ 17,144,168
2/4/4	$ 21,843,177
2/4/5	$ 30,601,297
3/4/1	$ 7,711,261
3/4/2	$ 14,653,851
3/4/3	$ 20,189,415
3/4/4	$ 25,723,090
3/4/5	$ 36,036,878
4/4/1	$ 9,890,770
4/4/2	$ 18,795,612
4/4/3	$ 25,895,745
4/4/4	$ 32,993,457
4/4/5	$ 46,222,331
5/4/1	$ 116,921,990
5/4/2	$ 222,189,008
5/4/3	$ 306,121,987
5/4/4	$ 390,026,330
5/4/5	$ 546,409,134

Model No. (SHL / TIL / RD^3)	Mean Project Pt. Estimate Cost (FY19$)
1/5/1	$ 70,977,471
1/5/2	$ 134,879,793
1/5/3	$ 185,831,291
1/5/4	$ 236,765,406
1/5/5	$ 331,697,555
2/5/1	$ 71,858,237
2/5/2	$ 136,553,530
2/5/3	$ 188,137,290
2/5/4	$ 239,703,451
2/5/5	$ 335,813,623
3/5/1	$ 75,601,387
3/5/2	$ 143,666,706
3/5/3	$ 197,937,504
3/5/4	$ 252,189,785
3/5/5	$ 353,306,409
4/5/1	$ 82,137,053
4/5/2	$ 156,086,553
4/5/3	$ 215,049,008
4/5/4	$ 273,991,345
4/5/5	$ 383,849,403
5/5/1	$ 279,927,608
5/5/2	$ 531,951,583
5/5/3	$ 732,898,881
5/5/4	$ 933,777,620
5/5/5	$ 1,308,179,939

Table J-2. TIL/SHL/TA Composite Model Mean Project Costs (FY19$)

Model No. (SHL/TIL/TA)	Mean Project Cost (FY19$)	Model No. (SHL/TIL/TA)	Mean Project Cost (FY19$)	Model No. (SHL/TIL/TA)	Mean Project Cost (FY19$)
1/1/1	$ 1,602,821	2/1/1	$ 1,826,081	3/1/1	$ 2,917,735
1/1/2	$ 1,216,034	2/1/2	$ 1,385,418	3/1/2	$ 2,213,638
1/1/3	$ 1,163,291	2/1/3	$ 1,325,327	3/1/3	$ 2,117,625
1/1/4	$ 1,406,937	2/1/4	$ 1,602,912	3/1/4	$ 2,561,152
1/1/6	$ 2,892,111	2/1/6	$ 3,294,958	3/1/6	$ 5,264,724
1/1/7	$ 2,651,541	2/1/7	$ 3,020,879	3/1/7	$ 4,826,797
1/1/11	$ 2,158,095	2/1/11	$ 2,458,699	3/1/11	$ 3,928,540
1/1/12	$ 643,180	2/1/12	$ 732,769	3/1/12	$ 1,170,828
1/1/13	$ 2,717,763	2/1/13	$ 3,096,325	3/1/13	$ 4,947,347
1/1/14	$ 1,043,885	2/1/14	$ 1,189,289	3/1/14	$ 1,900,261
1/2/1	$ 2,942,953	2/2/1	$ 3,242,894	3/2/1	$ 4,655,306
1/2/2	$ 2,232,771	2/2/2	$ 2,460,331	3/2/2	$ 3,531,905
1/2/3	$ 2,135,928	2/2/3	$ 2,353,618	3/2/3	$ 3,378,714
1/2/4	$ 2,583,289	2/2/4	$ 2,846,573	3/2/4	$ 4,086,371
1/2/6	$ 5,310,228	2/2/6	$ 5,851,437	3/2/6	$ 8,399,976
1/2/7	$ 4,868,516	2/2/7	$ 5,364,707	3/2/7	$ 7,701,255
1/2/11	$ 3,962,495	2/2/11	$ 4,366,346	3/2/11	$ 6,268,067
1/2/12	$ 1,180,947	2/2/12	$ 1,301,307	3/2/12	$ 1,868,080
1/2/13	$ 4,990,108	2/2/13	$ 5,498,691	3/2/13	$ 7,893,595
1/2/14	$ 1,916,686	2/2/14	$ 2,112,031	3/2/14	$ 3,031,906
1/3/1	$ 5,181,658	2/3/1	$ 5,577,275	3/3/1	$ 7,389,903
1/3/2	$ 3,931,240	2/3/2	$ 4,231,388	3/3/2	$ 5,606,599
1/3/3	$ 3,760,728	2/3/3	$ 4,047,858	3/3/3	$ 5,363,422
1/3/4	$ 4,548,397	2/3/4	$ 4,895,665	3/3/4	$ 6,486,768
1/3/6	$ 9,349,720	2/3/6	$ 10,063,566	3/3/6	$ 13,334,250
1/3/7	$ 8,571,998	2/3/7	$ 9,226,465	3/3/7	$ 12,225,089
1/3/11	$ 6,976,767	2/3/11	$ 7,509,439	3/3/11	$ 9,950,025
1/3/12	$ 2,079,294	2/3/12	$ 2,238,047	3/3/12	$ 2,965,418
1/3/13	$ 8,786,084	2/3/13	$ 9,456,897	3/3/13	$ 12,530,412
1/3/14	$ 3,374,709	2/3/14	$ 3,632,366	3/3/14	$ 4,812,894
1/4/1	$ 16,837,794	2/4/1	$ 17,545,107	3/4/1	$ 20,661,572
1/4/2	$ 12,774,560	2/4/2	$ 13,311,187	3/4/2	$ 15,675,598
1/4/3	$ 12,220,483	2/4/3	$ 12,733,835	3/4/3	$ 14,995,693
1/4/4	$ 14,780,013	2/4/4	$ 15,400,883	3/4/4	$ 18,136,478
1/4/6	$ 30,381,906	2/4/6	$ 31,658,173	3/4/6	$ 37,281,483
1/4/7	$ 27,854,698	2/4/7	$ 29,024,803	3/4/7	$ 34,180,359
1/4/11	$ 22,670,997	2/4/11	$ 23,623,348	3/4/11	$ 27,819,465
1/4/12	$ 6,756,665	2/4/12	$ 7,040,495	3/4/12	$ 8,291,069
1/4/13	$ 28,550,373	2/4/13	$ 29,749,702	3/4/13	$ 35,034,018
1/4/14	$ 10,966,114	2/4/14	$ 11,426,772	3/4/14	$ 13,456,462
1/5/1	$ 190,177,205	2/5/1	$ 192,537,133	3/5/1	$ 202,566,538
1/5/2	$ 144,284,351	2/5/2	$ 146,074,790	3/5/2	$ 153,683,937
1/5/3	$ 138,026,235	2/5/3	$ 139,739,016	3/5/3	$ 147,018,127
1/5/4	$ 166,935,256	2/5/4	$ 169,006,772	3/5/4	$ 177,810,463
1/5/6	$ 343,153,385	2/5/6	$ 347,411,609	3/5/6	$ 365,508,543
1/5/7	$ 314,609,421	2/5/7	$ 318,513,440	3/5/7	$ 335,105,046
1/5/11	$ 256,061,265	2/5/11	$ 259,238,754	3/5/11	$ 272,742,697
1/5/12	$ 76,314,253	2/5/12	$ 77,261,244	3/5/12	$ 81,285,841
1/5/13	$ 322,466,834	2/5/13	$ 326,468,356	3/5/13	$ 343,474,340
1/5/14	$ 123,858,555	2/5/14	$ 125,395,528	3/5/14	$ 131,927,476

(continues)

Table J-2 (*continued*)

Model No. (SHL/TIL/TA)	Mean Project Cost (FY19$)	Model No. (SHL/TIL/TA)	Mean Project Cost (FY19$)
4/1/1	$ 5,338,798	5/1/1	$ 220,890,490
4/1/2	$ 4,050,459	5/1/2	$ 167,586,021
4/1/3	$ 3,874,777	5/1/3	$ 160,317,230
4/1/4	$ 4,686,333	5/1/4	$ 193,895,007
4/1/6	$ 9,633,261	5/1/6	$ 398,572,054
4/1/7	$ 8,831,954	5/1/7	$ 365,418,290
4/1/11	$ 7,188,346	5/1/11	$ 297,414,709
4/1/12	$ 2,142,351	5/1/12	$ 88,638,871
4/1/13	$ 9,052,533	5/1/13	$ 374,544,661
4/1/14	$ 3,477,051	5/1/14	$ 143,861,494
4/2/1	$ 7,617,942	5/2/1	$ 234,453,152
4/2/2	$ 5,779,609	5/2/2	$ 177,875,792
4/2/3	$ 5,528,927	5/2/3	$ 170,160,697
4/2/4	$ 6,686,938	5/2/4	$ 205,800,148
4/2/6	$ 13,745,721	5/2/6	$ 423,044,354
4/2/7	$ 12,602,333	5/2/7	$ 387,854,950
4/2/11	$ 10,257,065	5/2/11	$ 315,675,954
4/2/12	$ 3,056,926	5/2/12	$ 94,081,292
4/2/13	$ 12,917,078	5/2/13	$ 397,541,680
4/2/14	$ 4,961,411	5/2/14	$ 152,694,581
4/3/1	$ 11,028,270	5/3/1	$ 251,942,129
4/3/2	$ 8,366,969	5/3/2	$ 191,144,394
4/3/3	$ 8,004,064	5/3/3	$ 182,853,794
4/3/4	$ 9,680,482	5/3/4	$ 221,151,761
4/3/6	$ 19,899,273	5/3/6	$ 454,601,245
4/3/7	$ 18,244,025	5/3/7	$ 416,786,896
4/3/11	$ 14,848,850	5/3/11	$ 339,223,725
4/3/12	$ 4,425,421	5/3/12	$ 101,099,264
4/3/13	$ 18,699,672	5/3/13	$ 427,196,205
4/3/14	$ 7,182,488	5/3/14	$ 164,084,796
4/4/1	$ 26,501,352	5/4/1	$ 313,281,061
4/4/2	$ 20,106,145	5/4/2	$ 237,681,244
4/4/3	$ 19,234,071	5/4/3	$ 227,372,178
4/4/4	$ 23,262,567	5/4/4	$ 274,994,335
4/4/6	$ 47,818,710	5/4/6	$ 565,280,452
4/4/7	$ 43,841,085	5/4/7	$ 518,259,657
4/4/11	$ 35,682,351	5/4/11	$ 421,812,617
4/4/12	$ 10,634,455	5/4/12	$ 125,713,333
4/4/13	$ 44,936,022	5/4/13	$ 531,203,262
4/4/14	$ 17,259,793	5/4/14	$ 204,033,598
4/5/1	$ 220,078,217	5/5/1	$ 750,038,705
4/5/2	$ 166,969,763	5/5/2	$ 569,042,162
4/5/3	$ 159,727,701	5/5/3	$ 544,360,815
4/5/4	$ 193,182,004	5/5/4	$ 658,374,926
4/5/6	$ 397,106,399	5/5/6	$ 1,353,360,516
4/5/7	$ 364,074,550	5/5/7	$ 1,240,786,151
4/5/11	$ 296,321,037	5/5/11	$ 1,009,878,439
4/5/12	$ 88,312,923	5/5/12	$ 300,975,312
4/5/13	$ 373,167,361	5/5/13	$ 1,271,774,953
4/5/14	$ 143,332,477	5/5/14	$ 488,484,989

Appendix K Macro-Parametric Model Project Estimating Examples

Project 1

The first sample project estimate is for one of the 125 three-parameter system hierarchy level (SHL)/ technology readiness level (TRL) improvement level (TIL)/research and development degree of difficulty (RD^3) models with a project configuration of SHL = 4, TRL_{Start} = 4, TRL_{End} = 7 (TIL = 3), and RD^3 = 5 (model no. 4/4/7/5 representing SHL/TRL_{Start}/TRL_{End}/RD^3/technology area, or TA). The methodology starts with the SHL/TIL multiple-regression model output for SHL = 4 and TIL = 3 from Table 1 and Appendix B, which results in a mean cost of $10,080,685 (in fiscal year, or FY, 2019 thousands of dollars). This mean project value is adjusted to discrete TRL start and end states of 3 and 7 (for a TIL = 7 – 4 = 3) applying the cost factor from Table 12 of 1.21 (rounded from 1.20788) and further refined by the RD^3 mean cost index (MCI) value = 1.9081 from Table 4, producing a project mean point estimate of ~$23,233,500. To provide a perspective of expected cost with uncertainty, however, a Monte Carlo simulation was run, substituting the probability density function (PDF) for the RD^3 = 5 MCI from Table H-1 (@RISK formula = RiskGamma (1.3688,1.394,RiskName ("RD^3 Lvl 5 MCI")) and Figure H-4 of Appendix H, producing the project cost uncertainty distribution shown in Figure K-1. The resulting fiftieth to eightieth percentile cost planning range for these project attributes is ~$18 million to ~$36 million with a seventieth percentile of $28.3 million, as illustrated in the PDF plot and table. Generating curve fits for this PDF produces and optimal function in @RISK of =RiskGamma(1.3689,16972745,RiskName("4/4/7/5/Project Cost PDF Curve Fits (FY19$)").

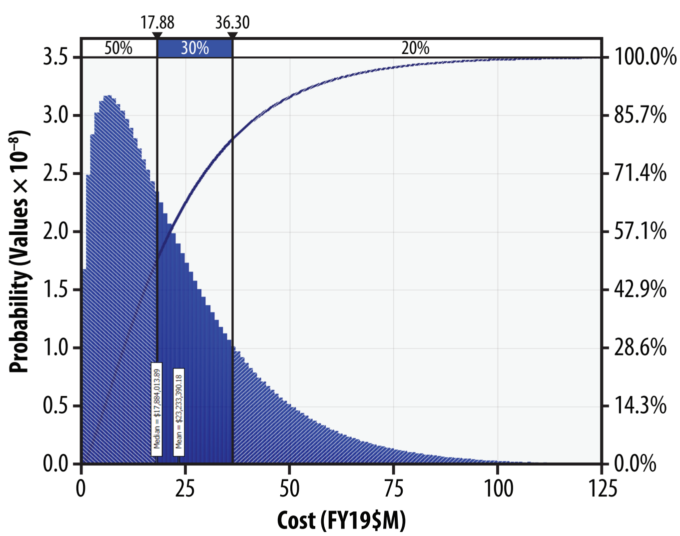


Figure K-1. Project 1 Uncertainty PDF: SHL = 4, TRL_{Start} = 4, TRL_{End} = 7 (TIL = 3), and RD^3 = 5

Project 2

Similarly, a four-parameter SHL/TIL/RD^3/TA macro-parametric model estimate is demonstrated for a hypothetical project characterized by SHL = 1, TRL_{Start} = 3, TRL_{End} = 7 (TIL = 4), RD^3 = 5, and TA = 4 (i.e., robotics, telerobotics, autonomous systems) (model no. 1/3/7/5/4 representing SHL/TRL_{Start}/TRL_{End}/RD^3/TA). This estimate is calculated starting with a base TIL/SHL macro-parametric regression model and then fine-tuned by the discrete TRL start/end cost factor and both the RD^3 MCI and TA MCI estimate values. Again, to provide a perspective of estimate uncertainty, the inputs are run in a Monte Carlo simulation, replacing the RD^3 MCI point estimate with the corresponding RD^3 MCI PDF. The SHL/TIL regression model returns a mean point estimate of $15,391,037 (FY 2019 dollars), from Table 1 and Appendix B. This mean project value is adjusted by a TRL start/end (=3/7) to TIL (=4) average cost factor of 0.97 (rounded from 0.96525) from Table 12, an RD^3 MCI value of 1.9081 from Table 4, and a TA MCI = 0.9603 from Table 5, producing a project mean point estimate value of ~$27,221,700. To develop the overall expected cost with uncertainty, a Monte Carlo simulation is run using the PDF for RD^3 = 5 MCI from Table H-1 (@RISK formula = RiskGamma (1.3688,1.394,RiskName ("RD^3 Lvl 5 MCI")) and Figure H-4 of Appendix H, producing the project cost uncertainty PDF shown in Figure K-2. The resulting median to eightieth percentile cost planning range for these project characteristics is ~$21 million to ~$42.5 million with a seventieth percentile of $33.2 million, as illustrated in the PDF plot and table.

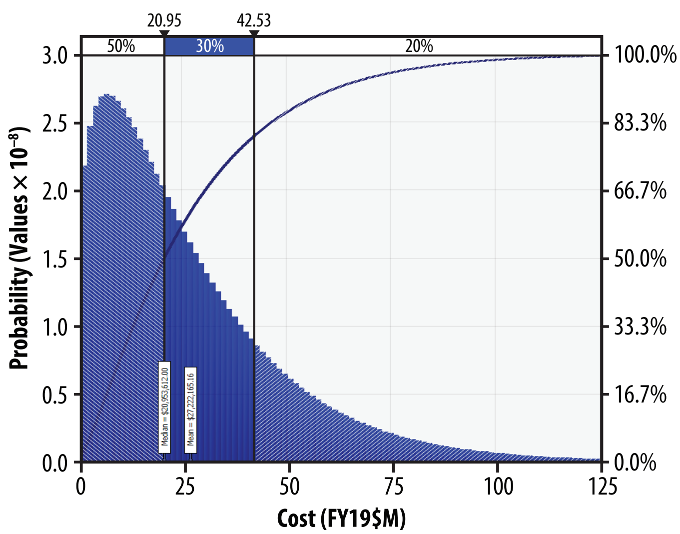


Figure K-2. Project 2 Uncertainty PDF: SHL = 1, TRL_{Start} = 3, TRL_{End} = 7 (TIL = 4), RD^3 = 5, and TA = 4

Appendix L Detailed Standard Development Framework Work Breakdown Structure Elements

This appendix provides a detailed four-level work breakdown structure (WBS) for containing a data dictionary and suggested element descriptions associated with corresponding research, development, test, and evaluation (RDT&E) budget activities (BAs).[63] This detailed WBS is not intended to be prescriptive but instead serves as general guidance in identifying the full range of processes in development, yet allowing for specific product orientation or system architectures to be threaded in where appropriate. This structure provides a comprehensive basis to help ensure that relevant design, development, integration, test, and demonstration requirements are effectively identified and captured for estimate development and budget planning.

Table L-1. Four-Level WBS

WBS No.	WBS Name	WBS Description[a]
1.0	DEVELOPMENT	Technology and systems development advancing and transitioning technology from conceptual scientific investigation through full systems development and demonstration in an operational environment to full operational capability (FOC).
1.1	Technology Development	Proof of concept (PoC) or feasibility demonstration in simulation and laboratory environment.
1.1.1	Basic Research	Basic research is systematic study directed toward greater knowledge or understanding of the fundamental aspects of phenomena and of observable facts without specific applications toward processes or products in mind. Generally performed by others outside of government programs at government or commercial labs, research universities, or industry independent research and development (IRAD) efforts.
1.1.2	Technology Research	Incubation-stage scientific investigation with translation to basic principles and early exploratory development during pre-materiel solution analysis (pre-MSA).
1.1.3	Analytical PoC Validation	Analytical PoC or feasibility demonstrated in a simulated environment establishing initial practicality of proposed solutions to technological requirements.
1.1.3.1	Development Nonrecurring Systems Engineering (NRE)	Development NRE including security considerations.
1.1.3.2	Systems Hardware	Systems hardware development, modifications or purchases (commercial off-the-shelf, or COTS), needed for this phase of demonstration.
1.1.3.3	Systems Software	Systems software development, modifications, or purchases (COTS), needed for this phase of demonstration.
1.1.3.4	Systems Integration	System integration activities including internal and external interfaces needed for this phase of demonstration,
1.1.3.5	Testing	Testing including any applicable test labor, equipment, labs/ranges, or platform costs and certification requirements, etc. needed for this phase of demonstration.
1.1.3.6	Project Management (PM)	Project planning, management, and oversight activities.
1.1.3.7	Support Services	Other support services may include logistics support, configuration management, facilities, IT, security, etc.
1.1.3.8	Other Direct Costs (ODCs)	ODCs may include applicable subcontract services, network/communications costs, travel, etc.
1.1.4	Validation in a Laboratory Environment (VLE)	Component or breadboard validation or ad hoc demonstration testing in a laboratory environment (VLE).
1.1.4.1	Development NRE	Development NRE including security considerations.
1.1.4.2	Systems Hardware	Systems hardware development, modifications, or purchases (COTS), needed for this phase of demonstration.
1.1.4.3	Systems Software	Systems software development, modifications, or purchases (COTS), needed for this phase of demonstration.
…	…	…

(continues)

[63] US Office of the Under Secretary of Defense (Comptroller)/CFO, *Financial Management Regulation.*

(continued)

1.1.5	VRE	Component or breadboard high-fidelity PoC validation or demonstration in a laboratory or relevant environment (VRE) (around SRR).
1.1.5.1	Development NRE	Development NRE including security considerations.
1.1.5.2	Systems Hardware	Systems hardware development, modifications, or purchases (COTS), needed for this phase of demonstration.
1.1.5.3	Systems Software	Systems software development, modifications, or purchases (COTS), needed for this phase of demonstration.
…	…	…
1.1.6	Prototype Demo in Relevant Environment (DRE)	Prototype system/subsystem technology design, integration build, test, and checkout for DRE.
1.1.6.1	Prototype System Design	Design of prototype architecture functional product breakdown of primary hardware, software, and all internal and external interfaces.
1.1.6.2	Vendor NRE	Vendor NRE.
1.1.6.3	Prototype System Build(s)	Build of prototype architecture functional product breakdown of primary hardware, software, and all internal and external interfaces.
1.1.6.4	Support Platform(s)/ Systems Modification Design	Platforms like sea/air/land/space assets and communications systems that require modifications to support concepts of operations.
1.1.6.5	System Integration, Assembly, Test, and Checkout (IAT&C)	Prototype IAT&C.
1.1.6.6	Systems Data	Prototype data and documentation including vendor system specs, drawings/diagrams and operations manuals as well as government purchase of intellectual data property rights.
1.2	Systems Development	Advancing technology from prototype to full-scale system functional integration, test and demonstration with operational system through IOC to FOC; includes prototype system integration and test and technology demonstration in an operational environment.
1.2.1	Systems Prototype Demo in Operational Environment (DOE)	Systems Prototype DOE.
1.2.2	Full-Scale Systems Development and Demonstration (SDD)	System test and evaluation (T&E)—functional or operational system test and demonstration.
1.2.2.1	Full-Scale System (FSS) Design	Design of full-scale architecture functional product breakdown of primary hardware, software, and all internal and external interfaces.
1.2.2.2	FSS Vendor NRE	Vendor NRE.
1.2.2.3	FSS Low-Rate Initial Production (LRIP) Build(s)	Build of LRIP full-scale systems including primary hardware, software, and all internal and external interfaces.
1.2.2.4	FSS Support Platform(s)/ Systems Modification Design	Platform modification and integration design and including sea/air/land/space assets and communications, command, control and intelligence (C3I) systems to support concepts of operations.
1.2.2.5	FSS IAT&C	FSS IAT&C.
1.2.2.6	FSS Data	FSS data and documentation including vendor system specs, drawings/diagrams, and operations manuals as well as government purchase of intellectual data property rights.
1.2.2.7	FSS Test Labor	Government (military and civilian) and contractor personnel to plan and perform the operational system field tests.
1.2.2.8	FSS Test Equipment	Procurement or lease of all necessary FSS test equipment.
1.2.2.9	FSS Test Support Organizations and Ranges	Costs for use of all test facilities, labs, ranges and associated ODCs.
1.2.2.10	FSS Test Platforms	Procurement, lease, or usage fees for test support platforms including sea/air/land/space assets and C3I systems that are part of the operational concept of operations.
1.2.2.11	FSS Pre-Test Certification	Costs associated with certification/approval to integrate development systems with operational systems for testing.
1.2.2.12	FSS Demonstration Test	System T&E/demonstration testing.
1.2.2.13	PM	Project planning, management, and oversight activities.
1.2.3	Operational Systems Evaluation (OPEval)	Full-system operational evaluation (OPEval) through full-rate production (RFP) approval, concluding with initial operational capability (IOC).
1.2.4	Operational Systems Development	Development efforts such as engineering or design modifications to resolve manufacturing or production issues for fielded systems up to FOC.

[a] General WBS guidance only and not intended as prescriptive; tailor WBS to system architecture and project requirements.

Bibliography

Air Force Research Lab, Transition Readiness Level Calculator. Version 2.2. May 17, 2004. https://www.dau.edu/cop/stm/Lists/Tools/DispForm.aspx?ID=4&Source=https%3A%2F%2Fwww%2Edau%2Eedu%2Fcop%2Fstm%2Flists%2Ftools%2Fallitems%2Easpx&ContentTypeId=0x010051CD3A8A368D5347A6595B3B5385510700AA867A0513EB0A44B7EC5897985C3BB8.

Alexander, Chuck. "Parametric Cost and Schedule Modeling for Early Technology Development." Laurel, MD: Johns Hopkins University Applied Physics Laboratory, 2018. https://www.jhuapl.edu/Content/documents/ParametricCostScheduleModeling.pdf.

Anderson, J. A. "Regression and Ordered Categorical Variables." *Journal of the Royal Statistical Society Series B* 46, no. 1 (1984): 1–30.

Atwater, Brad, and Joe Uzdzinski. "Wholistic Sustainment Maturity: The Extension of a System Readiness Methodology across All Phases of the Lifecycle of a Complex System." *Procedia Computer Science* 28 (2014): 601–609. https://doi.org/10.1016/j.procs.2014.03.073.

Berry, William D. *Understanding Regression Assumptions.* Newbury Park, CA: SAGE, 1993. https://dx.doi.org/10.4135/9781412986427.

Clark-Carter, David. "Geometric Mean." *Encyclopedia of Statistics in Behavioral Science*. Edited by B. S. Everitt and D. C. Howell. Wiley, 2005. https://doi.org/10.1002/0470013192.bsa376.

Cole, Stuart K., John D. Reeves, Julie A. Williams-Byrd, Marc Greenberg, Doug Comstock, John R. Olds, Jon Wallace, Dominic DePasquale, and Mark Schaffer. *Technology Estimating: A Process to Determine the Cost and Schedule of Space Technology Research and Development*. NASA Technical Report NASA/TP–2013-218145. Washington, DC: December 2013. https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20140005476.pdf.

Conrow, Edmund H. "Estimating Technology Readiness Level Coefficients." AIAA 2009-6727. Presented at the AIAA Space Conference & Exposition. AIAA: Pasadena, CA: 2009. https://doi.org/10.2514/6.2009-6727.

Copeland, Edward, Thomas Holzer, Timothy Eveleigh, and Shahryar Sarkani. "The Effects of System Prototype Demonstrations on Weapon Systems." *Defense ARJ* 22, no. 1 (2015): 106–134. https://www.dau.edu/library/arj/_layouts/15/WopiFrame.aspx?sourcedoc=/library/arj/ARJ/ARJ72/ARJ-72_Copeland.pdf&action=default.

Eder, Clarence, Thomas A. Mazzuchi, and Shahram Sarkani. 2017. "Beyond Integration Readiness Level (IRL): A Multidimensional Framework to Facilitate the Integration of System of Systems." *Defense ARJ* 24, no. 3 (2017): 488–533. https://www.dau.edu/library/arj/_layouts/15/WopiFrame.aspx?sourcedoc=/library/arj/ARJ/ARJ82/ARJ82_Article%205%20-%2016-766%20Eder.pdf.

Fagerland, Morten W. "t-Tests, Non-parametric Tests, and Large Studies—A Paradox of Statistical Practice?" *BMC Research Methodology* 12, no. 78 (2012): 1–7. https://doi.org/10.1186/1471-2288-12-78.

Gertheiss, Jan, and Gerhard Tutz. "Penalized Regression with Ordinal Predictors." *International Statistical Review* 77, no. 3 (2009): 345–365. https://doi.org/10.1111/j.1751-5823.2009.00088.x.

Glen, Stephanie. "Durbin Watson Test & Test Statistic." StatisticsHowTo.com: Elementary Statistics for the Rest of Us! Accessed September 27, 2022. https://www.statisticshowto.com/durbin-watson-test-coefficient/.

GraphPad Software, Inc. "Interpreting Results: One-Way ANOVA." In *GraphPad Statistics Guide*. Accessed October 5, 2019. https://www.graphpad.com/guides/prism/7/statistics/f_ratio_and_anova_table_(one-way_anova).htm?toc=0&printWindow.

Gove, Ryan, and Joe Uzdzinski. "A Performance-Based System Maturity Assessment Framework." *Procedia Computer Science* 16 (2013): 688–697. https://doi.org/10.1016/j.procs.2013.01.072.

Hu, Shu-Ping. "Simple Mean, Weighted Mean, or Geometric Mean?" Presented at the International Society of Parametric Analysts (ISPA)/Society of Cost Estimating and Analysis (SCEA) Professional Development and Training Workshop. San Diego, CA: ICEAA, 2010. https://www.iceaaonline.com/wp-content/uploads/2017/09/MET08A-Hu.pdf.

JMP Statistical Discovery. "Description of the Welch's Test Report." JMP.com. March 19, 2020. https://www.jmp.com/support/help/14-2/tests-that-the-variances-are-equal-report.shtml.

Lakens, Daniël. "Always Use Welch's t-Test Instead of Student's t-Test." *The 20% Statistician* (blog). January 26, 2015. http://daniellakens.blogspot.com/2015/01/always-use-welchs-t-test-instead-of.html.

———. "Equivalence Tests: A Practical Primer for *t* Tests, Correlations, and Meta-Analyses." *Social Psychological and Personality Science* 355-362. https://doi.org/10.1177/1948550617697177.

Lee, Tze-San, and Dale Thomas. "Cost Growth Models for NASA's Programs: A Summary." *Journal of Probability and Statistical Science* 8, no. 4 (2017): 265–279.

Linick, James. "Technology Development Level (TRL) vs. Percent Development Cost." Presented at the International Cost Estimating and Analysis Association (ICEAA) Professional Development & Training Workshop, Portland, OR, 2017. http://www.iceaaonline.com/wp-content/uploads/2017/07/DA05-PPT-Linick-Technology-Readiness-Level.pdf.

Malone, Patrick, Roy Smoker, Henry Apgar, and Lawrence Wolfarth. "The Application of TRL Metrics to Existing Cost Prediction Models." *Proceedings of the 2011 Aerospace Conference*, Big Sky, MT. https://doi.org/10.1109/AERO.2011.5747634.

Mankins, John C. *Research & Development Degree of Difficulty (R&D3): A White Paper*. Washington, DC: NASA Headquarters Office of Space Flight, Advanced Projects Office, March 10, 1998. http://www.economicswebinstitute.org/essays/nasadiff.pdf.

McChesney, Jasper. "You Should Summarize Data with the Geometric Mean." *Medium*, May 5. 2019. https://medium.com/@JLMC/understanding-three-simple-statistics-for-data-visualizations-2619dbb3677a.

mdawson69. September 30, 2015. Comment on jeff_kolton1, "How Do You interpret Welch's Test Results?" JMP User Community. https://community.jmp.com/t5/Discussions/How-do-you-interpret-Welch-s-Test-results/td-p/14531.

NASA. “Technology Readiness Levels (TRLs).” Accessed September 27, 2022. https://esto.nasa.gov/trl/.

NASA Earth Science Technology Office. “ESTO TRL Worksheet Calculator.” Microsoft Excel file. November 30. 2010. https://esto.nasa.gov/files/TRL_Worksheet_11-30-10.xls.

Naval Center for Cost Analysis. *Joint Agency Cost Schedule Risk and Uncertainty Handbook*. NCCA: March 12, 2014. https://www.ncca.navy.mil/tools/csruh/JA_CSRUH_16Sep2014.pdf.

———. “Naval Cost Division (FMB-6) Inflation Indices and Joint Inflation Calculator.” https://www.ncca.navy.mil/tools/inflation.cfm.

NCSS. “Equivalence Tests for Two Means (Simulation).” PASS Documentation, chap. 465. NCSS.com. Accessed August 2018. https://ncss-wpengine.netdna-ssl.com/wp-content/themes/ncss/pdf/Procedures/PASS/Equivalence_Tests_for_Two_Means-Simulation.pdf.

Nolte, William L., Roger J. Dziegiel, and Brian C. Kennedy. “Technology Readiness Calculator.” Presented at the National Defense Industrial Association (NDIA) 6th Annual Systems Engineering Conference, San Diego, CA, October 20, 2003. https://ndiastorage.blob.core.usgovcloudapi.net/ndia/2003/systems/nolte2.pdf.

Penn State Eberly College of Science. “What is a Variation Inflation Factor?” STAT 462, Applied Regression Analysis, online course materials. Lesson 10: Regression Pitfalls. Accessed September 27, 2022. https://online.stat.psu.edu/stat462/node/180/#:~:text=How%20do%20we%20interpret%20the,is%20not%20inflated%20at%20all.

Roenfeldt, Kimberly. “Better than Average: Calculating Geometric Means Using SAS.” Paper 56-2018. *SAS Proceedings: Western Users of SAS Software (WUSS)*, 2018. https://www.lexjansen.com/wuss/2018/56_Final_Paper_PDF.pdf.

Ross, Sean. “Application of System and Integration Readiness Levels to Department of Defense Research and Development.” *Defense ARJ*, 23, no. 3 (2016): 248–273. https://www.dau.edu/library/arj/ARJ/ARJ78/ARJ78-Ross_Article01.pdf?Web=1.

Ruxton, G. D. “The Unequal Variance t-Test Is an Underused Alternative to Student’s t-Test and the Mann–Whitney U Test.” *Behavioral Ecology* 17, no. 4 (2006): 688–690. https://doi.org/10.1093/beheco/ark016.

Sauser, Brian J., Jose E. Ramirez-Marquez, David Nowicki, Abhi Deshmukh, and Matin Sarfaraz. *Development of Systems Engineering Maturity Models and Management Tools*. Final Technical Report 2011-TR-014. Hoboken, NJ: Systems Engineering Research Center, 2011. https://apps.dtic.mil/sti/pdfs/ADA546788.pdf.

Sauser, Brian, Jose Ramirez-Marquez, Dinesh Verma, and Ryan Gove. “From TRL to SRL: The Concept of Systems Readiness Level.” Paper 126. Presented at the Conference on Systems Engineering Research (CSER), Los Angeles, CA, April 7–8, 2006. https://www.researchgate.net/publication/228652562_From_TRL_to_SRL_The_concept_of_systems_readiness_levels.

Shermon, Dale, and Catherine Barnaby. "Macro-Parametrics and the Applications of Multi-colinearity and Bayesian to Enhance Early Cost Modeling." Presented at the International Cost Estimating and Analysis Association (ICEAA) 2015 Professional Development & Training Workshop, San Diego, CA, June 9–12, 2015. http://www.iceaaonline.com/wp-content/uploads/2015/06/PA08-Paper-Shermon-Macro-Parametrics.pdf.

Sirirojvisuth, Nate. "Methodology to Assess Cost and Schedule Impact Using System and Technology Readiness Level (SRL/TRL)." Presented at the International Cost Estimating and Analysis Association (ICEAA) SoCal Workshop, El Segundo, CA, March 20, 2019. http://www.iceaaonline.com/wp-content/uploads/2019/03/Sirirojvisuth-Methodology-to-Access-Cost-032019.pdf.

Smoker, Roy E., and Sean Smith. "System Cost Growth Associated with Technology-Readiness Level." *Journal of Parametrics* 26, no. 1 (2007): 17–32. https://doi.org/10.1080/10157891.2007.10462276.

Stauner, Nick. "Effect of Two Demographic IVs on Survey Answers (Likert Scale)." Cross Validated. February 17, 2014. http://stats.stackexchange.com/questions/86923/effect-of-two-demographic-ivs-on-survey-answers-likert-scale.

Straub, Jeremy. "In Search of Technology Readiness Level (TRL) 10." *Aerospace Science and Technology* 46 (2015): 312–320. https://doi.org/10.1016/j.ast.2015.07.007.

US Government Accountability Office. *GAO Cost Estimating and Assessment Guide: Best Practices for Developing and Managing Capital Program Costs (Supersedes GAO-07-1134SP)*. GAO-09-3SP. Washington, DC: GAO, March 2, 2009. https://www.gao.gov/products/gao-09-3sp.

US Office of the Secretary of Defense Manufacturing Technology Program and the Joint Service/Industry MRL Working Group. *Manufacturing Readiness Level (MRL) Deskbook*. Version 2016. Arlington, VA: US Department of Defense, 2016. http://www.dodmrl.com/MRL_Deskbook_2016.pdf.

US Office of the Under Secretary of Defense (Comptroller)/CFO. Data for years 1996–2018. "DoD Budget Request." Accessed Oct 13, 2019. https://comptroller.defense.gov/Budget-Materials/.

———. *Financial Management Regulation*. DoD 7000.14-R, Volume 2B, Chapter 5. Washington, DC: Department of Defense, November 2017. https://comptroller.defense.gov/Portals/45/documents/fmr/Volume_02b.pdf.

Wallace, Jon, and Schaffer, Mark. *Technology Cost and Schedule Estimation (TCASE) Milestone 4—Final Report*. Atlanta, GA: SpaceWorks, September 19, 2015. https://ntrs.nasa.gov/search.jsp?R=20160000761.

Acknowledgments

Many thanks to John Ivancovich, Daniel Nicoloso, Matthew Schaffer, and Rodney Yerger of APL's National Security Analysis Department for their constructive reviews and inputs that helped focus key aspects of this report. Thanks also to Doug Comstock, Marc Greenberg, and Eric Plumer of NASA Headquarters and Kirk Cole of the NASA Langley Research Center Research Directorate for providing access to and background for the TCASE database, which was pivotal to this effort.

About the Author

Chuck Alexander is a member of the Principal Professional Staff at APL with over forty years of professional experience involving engineering, economic and cost analysis, and technology and business operations management across a range of technologies, applications, systems, and operational environments. He is a licensed professional engineer (PE) and a certified public accountant (CPA), and he holds a BS in mechanical engineering from Clarkson University and an MBA in finance from the William & Mary Mason School of Business.

JOHNS HOPKINS
APPLIED PHYSICS LABORATORY