AIAA SPACE 2007 Conference & Exposition
18 - 20 September 2007, Long Beach, California
AIAA 2007-6020

AIAA Space 2007
18-20 September 2007, Long Beach, CA
AIAA-2007-6020

Technology Readiness Level, Schedule Risk and Slippage in Spacecraft Design: Data Analysis and Modeling

Gregory F. Dubos*, **Joseph H. Saleh**†, **Robert Braun**‡

Georgia Institute of Technology, Atlanta, Georgia 30332

Schedule slippage plagues the space industry, and is antinomic with the recent emphasis on space responsiveness. The Government Accountability Office has repeatedly noted the difficulties encountered by the Department of Defense in keeping its acquisition of space systems on schedule, and identified the low Technology Readiness Level (TRL) of the system/payload under development as a principal culprit driving schedule risk and slippage.

In this paper, we analyze based on data from past space programs the relationship between technology uncertainty and schedule risk in the acquisition of space systems, and propose an analytical framework to identify appropriate schedule margins for mitigating the risk of schedule slippage. We also introduce the TRL-schedule-risk curves to help program managers make risk-informed decisions regarding the appropriate schedule margins for a given program, or the appropriate TRL to consider should the program's schedule be exogenously and rigidly constrained. We recommend based on our findings, that the industry adopts and develops schedule risk curves (instead of single schedule point estimates), 2) that these schedule risk curves be made available to policy- and decision-makers in acquisition programs; and 3) that adequate schedule margins be defined according to an agreed upon and acceptable schedule risk level.

Nomenclature

$env(RSS)$	=	envelope of RSS
$f(.)$	=	probability density function
FTD	=	Final Total schedule Duration
IDE	=	Initial schedule Duration Estimate
LB	=	lower bound of the RSS envelope
R^2	=	coefficient of determination in a regression analysis
RSS	=	Relative Schedule Slippage
sm	=	schedule margin
TRL	=	Technology Readiness Level
UB	=	upper bound of the RSS envelope
$WTRL$	=	Weighted average TRL
Z	=	standard normal random variable
$\langle RSS \rangle$	=	sample average of RSS
$\langle \overline{RSS} \rangle$	=	modeled average of RSS

* Graduate Research Assistant. School of Aerospace Engineering, Georgia Institute of Technology. AIAA student member.
† Assistant Professor. School of Aerospace Engineering, Georgia Institute of Technology. AIAA Senior Member. Corresponding author, jsaleh@gatech.edu.
‡ David and Andrew Lewis Professor of Space Technology. School of Aerospace Engineering, Georgia Institute of Technology. AIAA Fellow.

$\left\langle \overline{UB} \right\rangle$	=	modeled upper bound of the RSS envelope
α	=	multiplicative constant of the mean RSS exponential model
α'	=	multiplicative constant of the upper bound RSS exponential model
λ	=	rate of the mean RSS exponential model
λ'	=	rate of the upper bound RSS exponential model
μ	=	mean of the random variable RSS
σ	=	standard deviation of the random variable RSS

I. Introduction

The Government Accountability Office (GAO) has repeatedly noted[1,2,3,4,5] the difficulties encountered by the Department of Defense (DOD) in keeping its acquisition of space systems on schedule (and within budget):

> "*DOD's space system acquisitions have experienced problems over the past several decades that have driven up costs by hundreds of millions, even billions, of dollars; stretched schedules by years; and increased performance risks. In some cases, capabilities have not been delivered to the warfighter after decades of development.*" [1]

Several reasons testify to the criticality of maintaining space programs on schedule. First, the cost penalties associated with schedule slippage can significantly strain a program's budget, and in some cases jeopardize the (financial) survival and continued (political) support for the program. Second, the recent emphasis by the DOD on accelerated acquisition cycles in general, and on a more **responsive space** in particular—i.e., the ability to develop and launch a space asset on orbit within a very short timeframe—is fundamentally at odds with schedule slippages in acquisition programs.

Identifying the root causes of schedule slippage, and understanding the mechanics of their subversive influence, is critical for the DOD and other Federal agencies' acquisition functions. Furthermore, eliminating these causes and better controlling program schedule is increasingly emphasized as a priority. Chittenden[6] for example advised that "in order to resolve these enduring [schedule] concerns, the United States Air Force should pursue using schedule adherence or "SAIV" (schedule as an independent variable) as the first order programmatic driver for the majority of its acquisition programs."

What drives schedule slippage? Several factors come to mind as possible culprits. For example, it is easy to conceive of a system's **complexity** as having an important influence on the ability of a manufacturer or system integrator to stay on schedule or not—the more complex the system, the less likely its development can remain on schedule. Similarly, "**requirements creep**" can significantly perturb the development schedule of a system and result in major delays. This paper deals with a third factor that has been identified in several GAO reports as a principal culprit driving schedule slippage, namely the low **Technology Readiness Level** (TRL) of the system/payload under development.

The purpose of this paper is to empirically explore, in the case of space systems, **how much schedule risk/slippage is associated with various levels of technology maturity** or TRL, and to identify appropriate schedule margins to mitigate the risk of schedule slippage. In the following, we assume that schedule slippage is a random variable—more precisely, a random vector or an indexed family of random variables with TRL as the index—and we propose to characterize through data analysis and modeling the central tendency and dispersion of this random variable as a function of TRL (the independent variable in this study). To this effect, we briefly describe in Section 2 the TRL concept and highlight the relationship between technology uncertainty (i.e., low TRL) and schedule risk. In Section 3, we discuss and characterize the data used in our study. In Section 4, we analyze the data and provide analytical models for schedule risk and slippage as a function of TRL (and schedule margins). In Section 5, we conclude with some implications and recommendations from our work.

II. Technology Readiness Level and Schedule Risk

The Technology Readiness Level (TRL) is a widely adopted metric by NASA and the DOD; it was introduced by NASA in the 1980s, first to assess the maturity of a particular technology before its implementation in a system, and second, to allow the "consistent comparison of maturity between different types of technology."[7] This metric is organized on a scale of nine levels corresponding to key stages of development of a given technology. A brief description of these levels is provided in Table 1; the reader is referred to Mankins[7] for a more detailed description of these levels.

Table 1. Summary of different Technology Readiness Levels

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

The lack of technology maturity or low TRL, sometimes described in the literature as technology uncertainty, is often associated with schedule risk, albeit qualitatively. Browning[8] defines schedule risk as the "*uncertainty* in the ability of a project to develop an acceptable design […] within a span of time, *and* the *consequences* thereof." The author also defines technology risk as the "*uncertainty* in capability of technology to provide performance benefits (within cost and/or schedule expectations), *and* the *consequences* thereof." By their definitions alone, these concepts suggest a close relationship between technology uncertainty and schedule risk. In fact, in a study conducted by Gupta and Wilemon[9] of large technology-based firms, "about 58% of the interviewees cited technological uncertainties as a major reason for delays." The link between technology uncertainty and technology maturity is intuitive: the more mature a technology is, the more knowledge is available concerning its development, manufacturing, and mode(s) of operation. This, in turn, provides a higher confidence level that the mission requirements will be met. As a result, technology uncertainty in the project is reduced. Therefore, maturing technology is critical to completing a program on schedule and within budget.

As often noted by GAO, schedules overruns are most likely to occur when the burden of technology maturation is assumed within an acquisition program.[10] Indeed, as the low-TRL world (research environment, or S&T in government parlance) and the high-TRL world (e.g, development and production) are significantly different and do not always interact seamlessly, it is hard to predict how smooth this maturation process will be, and more importantly, how much time it will take to bring a low TRL technology (e.g., TRL = 4) to a comfortable level of maturity (e.g., TRL = 8). This issue is sometimes referred as the **TRL gap** and is described by George and Powers[11] as "the problem of efficiently transitioning a new technology from concept to viable product in the **shortest possible time** and at the least cost."

In the following, we propose to quantify the relationship between technology uncertainty and schedule risk using TRL as a proxy for the former. We start in the following section by describing the data collected for our analysis.

III. Data Description

Paradoxically, despite the fact that technology readiness level is a central theme in feasibility studies of system design (spacecraft and other), limited TRL data is available to the technical community for analysis—unlike other parameters such as system cost for example for which quantitative data and a number of (cost) models exist and are widely available. In some cases, when TRL is discussed in the technical literature, qualitative maturity levels ("Low/Medium/High") are employed.

For the purpose of this study, we were provided with programmatic data from 28 NASA programs. Lee and Thomas[12] used this data to construct probability-based models for the cost growth of NASA's programs. Details about this data can be found in Ref. 12. In this study, we focus instead on schedule slippage and concern ourselves with three parameters from the data set:

1. TRL at start of program
2. Initial schedule Duration Estimate (IDE)
3. Final Total schedule Duration (FTD)

We define the Relative Schedule Slippage (RSS) as the percentage schedule growth given the initial schedule estimate:

$$RSS = \frac{(FTD - IDE)}{IDE} \cdot 100 \tag{1}$$

Recall that the objective of this paper is to quantify how much schedule risk/slippage is associated with different levels of technology maturity or TRL. Given this objective, we perform a regression analysis on the data and investigate the relationship between TRL and RSS. We analyze both the central tendencies and the dispersion of RSS as a function of TRL and relate our results to schedule risk and slippage. The details are further discussed in Section IV.

Before we proceed however, a subtlety concerning the TRL data should be addressed:

Lee and Thomas[12] calculated a weighted average of TRL for each program (WTRL), by taking the "TRL of each component multiplied by their corresponding percent of the allocated cost against the entire program's cost".

$$WTRL_{program} = \sum_{components\, c_i} w_i \cdot TRL_{c_i} \qquad \text{where } w_i = \frac{cost_i}{cost_{program}} \tag{2}$$

In our study, we used the WTRL as a preliminary basis for the "average system-TRL" whose influence on schedule slippage we investigated. The WTRL is proportional to the amount of resources spent for each component. Components with a small w_i are either of minor importance in the design, or their TRL is already sufficiently high to limit the allocated cost for their development and implementation. In both cases, it is reasonable to assume that such components will not critically impact the advancement of the schedule, which justifies the use of the WTRL for our schedule analysis. However, this WTRL calculation results in a value with decimal digits. Such a degree of precision was not relevant for our study. To obtain the average system-TRL, the final step consisted in taking the integer part of the WTRL. Here again, when considering components requiring a large resource investment, we contend that those with the lowest TRLs drive the schedule delays, as they represent the "slowest links" of the maturation chain. For example, consider a program whose WTRL is 4.62. If it involves components with TRL 5 or 6, it also involves components with integer values of TRL less or equal than 4. First, the WTRL of 4.62 gives a good indicator of the "average TRL" of the entire system. Then, considering that components with low TRL (e.g., TRL = 4) have a bigger impact on schedule slippage than components with TRL 5, we retained the integer value, that is TRL = 4.[§]

IV. Modeling Schedule Slippage and Risk

For each of the 28 NASA programs in our data set, we plot and analyze the doublet (TRL; RSS) where the TRL consists of the integer values discussed in the previous section. The TRLs in our data set range from 4 to 8. We consider the relative schedule slippage a random variable—more precisely, a random vector or an indexed family of random variables with TRL as the index. In the following, we analyze and model both the central tendency and the dispersion of this random variable as a function of the independent variable in this study, namely TRL.

[§] Following these logics, one could argue that the minimum of all the components' TRLs could be directly used in place of the WTRL. However, we think it is important to capture first the relative importance of every component in terms of the amount of resources spent. The WTRL provides this function.

A. Mean relative schedule slippage

We capture the central tendency of RSS by its mean or average value, which for a given TRL is defined as follows:

$$\langle RSS \rangle_j = \sum_{i=1}^{n} \frac{RSS_i}{n} \bigg|_{TRL=j} \tag{3}$$

Figure 1 shows the mean RSS for each TRL. For example, for a TRL = 4 at start of the program, Figure 1 shows that an average 78% schedule slippage has been observed in all 28 programs considered—in other words, programs' schedules have been consistently underestimated by 78% when the TRL at start of the program was 4 (this is low maturity technology in the context of a space acquisition program). Similarly, when TRL at start of the program was 7, Figure 1 shows a mean RSS of 19%.

More generally, Figure 1 shows a monotonically decreasing average RSS as a function of TRL. This result can be interpreted as follows: the quality of the initial schedule estimate (IDE) at start of the program improves (i.e., is more accurate) as the technologies considered for the program are more mature. Conversely, the lower the maturity of the technology considered, the less we can predict with accuracy the actual schedule or FTD (i.e., the bigger the error in the program's initial schedule estimate). While this result may be considered intuitive, Figure 1 provides an empirical confirmation of this intuition.

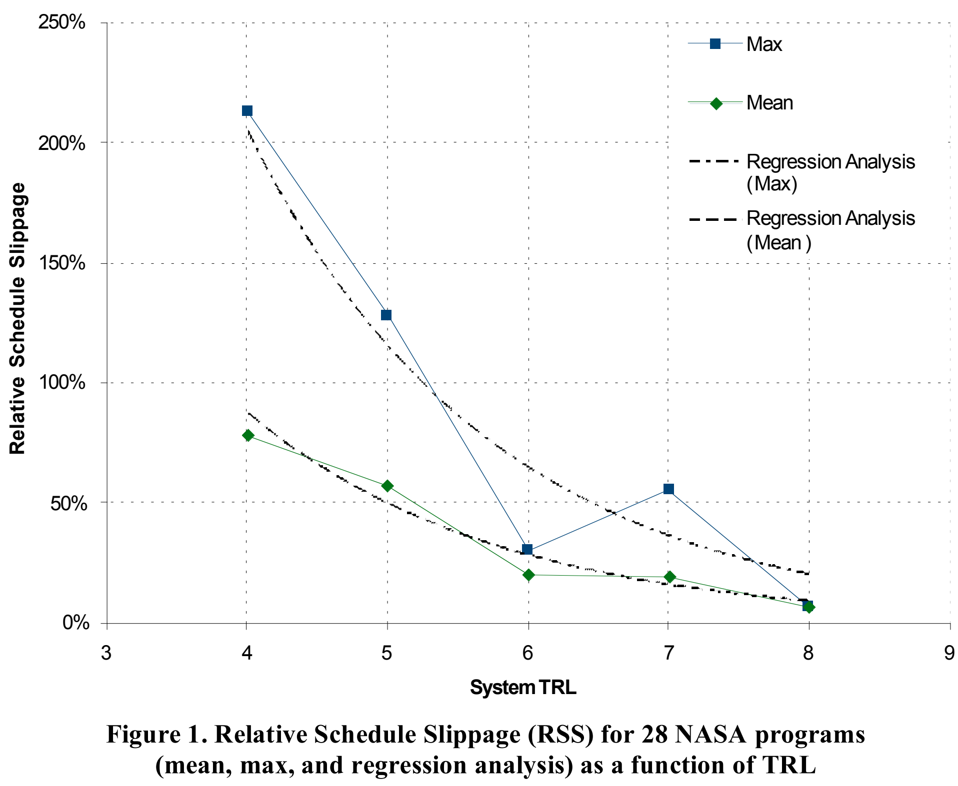

Figure 1. Relative Schedule Slippage (RSS) for 28 NASA programs (mean, max, and regression analysis) as a function of TRL

To analytically reflect this trend, we propose to model the mean relative schedule slippage with a decreasing exponential function of TRL, and perform a regression analysis on our data set to fit the model parameters. Equation (4) represents the model structure:

$$\left\langle \overline{\mathrm{RSS}} \right\rangle = \alpha \cdot e^{-\lambda \cdot TRL} \tag{4}$$

We chose this model structure both for its simplicity and conceptual relevance. A polynomial fit of order $n > 1$ for example would be meaningless considering the small size of the sample, and the absence of a conceptual interpretation of the coefficients needed to ensure goodness-of-fit. More importantly, we needed a function that 1) accounts for the reduction of the schedule slippage with higher TRLs, and 2) provides increasingly smaller increments in schedule slippage as TRL increases. Condition 2 can be stated mathematically as follows: the absolute value of the derivative of the $\left\langle \overline{\mathrm{RSS}} \right\rangle$ with respect to TRL should be a decreasing function. Hence, our choice of a decreasing exponential function (more details in the following paragraphs).

Table 2 shows the results of our regression analysis using this model structure (Eq. 4). A comparison of the observed and modeled mean relative schedule slippage is provided in Table 3. Our model of the mean relative schedule slippage, which consists of Eq. 4 and the value of its parameters in Table 2, is fairly accurate, as reflected by the coefficient of determination R^2, 94%, and by the error between the model output and the observed data (less than 10 percent).

The R^2 parameter[**] indicates that the variability in the mean relative schedule slippage is primarily accounted for by the TRL. However, due to the limited size of our sample (28 data points with an average of 6 points for each TRL), the R^2 value of our model, 94%, should be considered with caution and not interpreted beyond the fact it indicates an accurate model.

Table 2. Model parameters for the average schedule slippage in our data set

Model parameter	Value
α	8.29
λ	0.56
R^2	0.94

[**] If y_i are the values of the dependant variable considered, $\hat{y}_i$ the fitted values, and $\bar{y}$ the sample mean, the coefficient of determination is defined by $R^2 = \frac{\sum_i (\hat{y}_i - \bar{y})^2}{\sum_i (y_i - \bar{y})^2}$, and takes a value between 0 and 1.

Table 3. Model's accuracy: mean relative schedule slippage and TRL

TRL	Observed mean relative schedule slippage (data)	Modeled mean relative schedule slippage	Error
4	78%	88%	10
5	57%	50%	7
6	20%	29%	9
7	19%	16%	3
8	7%	9%	2

B. Dispersion of the relative schedule slippage

In addition to the mean relative schedule slippage, the data allows us to model the envelope or range within which the relative schedule slippage falls for each TRL. We refer to the range of the relative schedule slippage as its dispersion. In the following, we propose to model the range or envelope of RSS by the upper- and lower bound (UB, and LB respectively) values of RSS for each TRL level:

$$\begin{cases} UB_j = \max\left(RSS_i\right)\Big|_{TRL=j} \\ \\ LB_j = \min\left(RSS_i\right)\Big|_{TRL=j} \end{cases} \tag{5}$$

The envelope and dispersion of our data set are defined by Eq. 6:

$$\begin{cases} env\left(RSS\right) = \left\{UB_j;\ LB_j\right\} \\ \\ Dispersion_j = UB_j - LB_j \end{cases} \quad \text{for } j = 4,\ 5,\ 6,\ 7,\ 8 \tag{6}$$

The lower-bound model (LB_j) is trivial and equal to zero for all TRLs. In other words, for each TRL, at least one data point was found in our sample for which the initial estimated schedule (IDE) almost matched the actual schedule (FTD), thus resulting in an RSS almost equal to zero.[††] Consequently, the upper-bound model is also a model of our data dispersion.

We model the upper-bound with a decreasing exponential as defined in Eq. 7:

$$\left\langle \overline{\mathrm{UB}} \right\rangle = \alpha' \cdot e^{-\lambda' \cdot TRL} \tag{7}$$

[††] This was a surprising result for the low TRL (4 and 5). We can assume that for these exceptional cases a significant schedule margins was probably factored into the initial schedule estimate, although unfortunately the data we have does not allow us to verify this assumption. More details on schedule margins are discussed in the next section.

Figure 1 shows that the dispersion of RSS narrows down as TRL increases. This dispersion can be considered a proxy for the time uncertainty in the technology maturation process: the lower the TRL, the bigger the schedule uncertainty, that is, the less we can predict with accuracy the time it will take to complete a project. GAO[5] put it more forcefully:

> *"There is no way to estimate how long it would take to design, develop, and build a satellite system when critical technologies planned for that system are still in a relatively early stages of discovery and invention."*

Our results provide additional nuance to, and quantification of, this statement by GAO. Table 4 shows the results of our regression analysis using this model structure (Eq. 7). Notice that the upper-bound is a scaled up version of the mean relative schedule slippage with $\lambda' \approx \lambda$, and $\alpha' \approx 2.5\alpha$. Mathematically, the quasi-equality of the exponential coefficients implies that the changes in the mean and maximum RSS values for a given TRL jump are identical (Eq. 8). In practice, this quasi-equality indicates that the same phenomena resulting from the technology uncertainty affect the mean and the worst-case (maximum) schedule slippage for the range of TRL values in our data set, as captured in Eq. 8:

$$\frac{\left\langle \overline{\mathrm{RSS}} \right\rangle_{TRL=j}}{\left\langle \overline{\mathrm{RSS}} \right\rangle_{TRL=k}} \approx \frac{\left\langle \overline{\mathrm{UB}} \right\rangle_{TRL=j}}{\left\langle \overline{\mathrm{UB}} \right\rangle_{TRL=k}} \tag{8}$$

Our model of the dispersion of the relative schedule slippage is fairly accurate, as reflected by the coefficient of determination R^2 (83%). However, the same caveat regarding the R^2 parameter discussed previously (IV-A) also applies in this case of the dispersion model.

Table 4. Model parameters for the maximum schedule slippage in our data set

Model parameter	Value
α'	20.47
λ'	0.57
R^2	0.83

Beyond the schedule estimation errors reflected by the mean RSS model (Equation 4 and Table 2)—these may be due to a variety of factors including intrinsically flawed schedule estimation methods in use by the industry, or misleading industry practices—the dispersion of the RSS data suggests the existence of other sources of discrepancies between FTD and IDE (i.e., other than TRL), specific to each space program (e.g., complexity of the system under development, experience of the program manager, etc.). Further research is needed to identify these other parameters and correlate them with schedule slippage. Such work would be particularly relevant to the space industry, as it will 1) help develop credible schedule estimates, 2) limit programs' schedule risk, and 3) identify and disseminate best practices related to maintaining acquisition programs on schedule. The GAO reports mentioned previously (see Section I) constitute an important step in this direction. Academic analyses of these issues are required and can usefully complement the GAO reports and help disseminate the results beyond the reach of the traditional GAO readership.

C. Schedule risk curves and schedule margins

The existence of a non-zero mean relative schedule slippage strongly suggests the need to add a schedule margin “on top of” the initial schedule estimate or IDE (said differently, traditional schedule estimate methods are biased and consistently underestimate actual programs’ schedules). In this subsection, we discuss schedule margins and introduce the concept of schedule risk curves, or more precisely, TRL-schedule risk curves. In particular, for a given TRL at start of the program, we calculate the probability of overshooting the initial schedule estimate plus a given schedule margin. This in turn will help us determine appropriate schedule margins, as discussed below.

A schedule margin (*SM* or *sm*) can be expressed in the same units as IDE (e.g., months), or it can be expressed as a percentage of IDE; we use lower-case to denote the relative value, and upper-case to denote the absolute value of the schedule margin. Recall that RSS, and consequently FTD, which results from a linear transformation of RSS (see Eq. 1), are random variables. Consequently, we define schedule risk as the probability that the actual schedule (FTD) overshoots the initial schedule estimate plus a schedule margin, as shown in Figure 2b. Mathematically, we express the schedule risk for a given TRL at start of the program as follows:

$$Schedule\ \ Risk\big|_{TRL=j} = P\left\{FTD_j > IDE_j \cdot \left(1 + sm_j\right)\right\} \tag{9}$$

Assuming we can evaluate Eq. 9, the result would be interpreted as follows: a 20% schedule risk for example corresponds to a 20% chance that the initial schedule estimate plus the schedule margin underestimate the actual program’s schedule. Conversely, a 20% schedule risk corresponds to a confidence level of 80% that the actual schedule will fall within the initial schedule estimate plus the schedule margin. The higher the schedule margin, the lower the schedule risk. It should be noted that Eq. 9 is equivalent to the following:

$$Schedule\ \ Risk\big|_{TRL=j} = P\left\{RSS_j > sm_j\right\} \tag{10}$$

In order to evaluate Eq. 9, we need to find or assume a probability distribution function for FTD or RSS for each TRL. The data in our sample is not rich enough to allow us to infer a probability distribution function for FTD or RSS. However, for the purpose of introducing the concept of TRL-schedule risk curve, let us assume that for a given TRL, RSS (FTD) is normally distributed. That is, we formulate the *hypothesis* that the RSS (FTD) has a normal probability density function, which can be written as follows:

$$f\left(RSS_j\right)\big|_{TRL=j} = \frac{1}{\sigma_j\sqrt{2\pi}}e^{-\frac{(RSS_j-\mu_j)^2}{2\sigma_j^{\,2}}} \tag{11}$$

For a given value j of the TRL, μ_j is approximated by our sample average RSS calculated in Eq. (4) and illustrated in Figure 1, and σ_j represents the standard deviation, which is related to the dispersion of the data. Figure 2 provides a graphic illustration of Eq. 10 and 11.

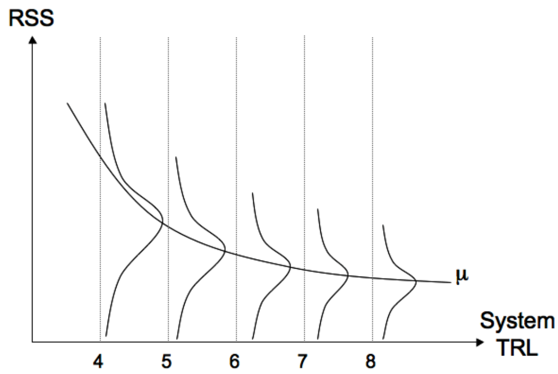

Figure 2a. RSS probability density functions for each TRL value

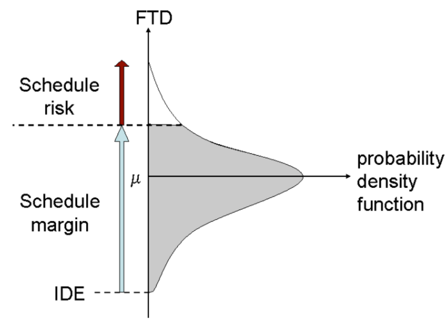

Figure 2b. FTD probability density function and schedule risk

In statistical data analysis, the standard deviation of a normally distributed random variable is approximated by the dispersion of a set of measurements (i.e., the sample) as follows:

$$\sigma \approx \frac{Dispersion}{4} \tag{12}$$

The reader is referred to Lyman-Ott and Longnecker[13] for a discussion of this approximation. Finally, the probability distribution function of the RSS as defined in Eq. 11 for each TRL value is fully determined by the following:

$$\begin{cases} \mu_j \approx \left\langle \overline{\mathrm{RSS}} \right\rangle \Big|_{TRL=j} \\ \\ \sigma_{\mathrm{j}} \approx \dfrac{Dispersion_j}{4} \end{cases} \qquad \text{for } j = 4, 5, 6, 7, 8 \tag{13}$$

Given a TRL at start of a program, and an agreed upon schedule risk, we propose to find the corresponding schedule margin that should to be added to IDE to ensure this level of schedule risk. Mathematically, *sm* now becomes the unknown in Eq. (9) and (10). Equation 10 can be rewritten in terms of the standard normal distribution as follows:

$$Schedule\ \ Risk\Big|_{TRL=j} = 1 - \Pr\left\{RSS_j \le sm_j\right\}$$

$$= 1 - \Pr\left\{Z \le z\right\} = 1 - \phi(z) \tag{14}$$

$$\text{with } Z = \frac{RSS_j - \mu_j}{\sigma_j} \tag{15}$$

$$\text{and } z = \frac{sm_j - \mu_j}{\sigma_j} \tag{16}$$

where Z is the standardized random variable derived from RSS, and ϕ is the cumulative density function of the standard normal distribution function, whose values can be found in tabulated form in the literature (see for example Ref. 14).

Numerical example: What should the schedule margin be if a schedule risk of 20% is called for, and the TRL at start of the program is 6? In this case, $\phi(z) = 1 - 0.2 = 0.8$, and we find the corresponding tabulated value of z = 0.84. Using Eq. 16, we find a schedule margin of 43%. In other words, for TRL = 6 at start of a program, a schedule margin of 43% will ensure a 20% schedule risk (or conversely, that there is 80% likelihood that the FTD will fall below the initial schedule estimate IDE plus the 43% schedule margin).

When done for all TRLs, not just one TRL value as in the previous numerical example, this analysis results in the construction of a family of **TRL-schedule risk curves,** as shown in Figure 3.

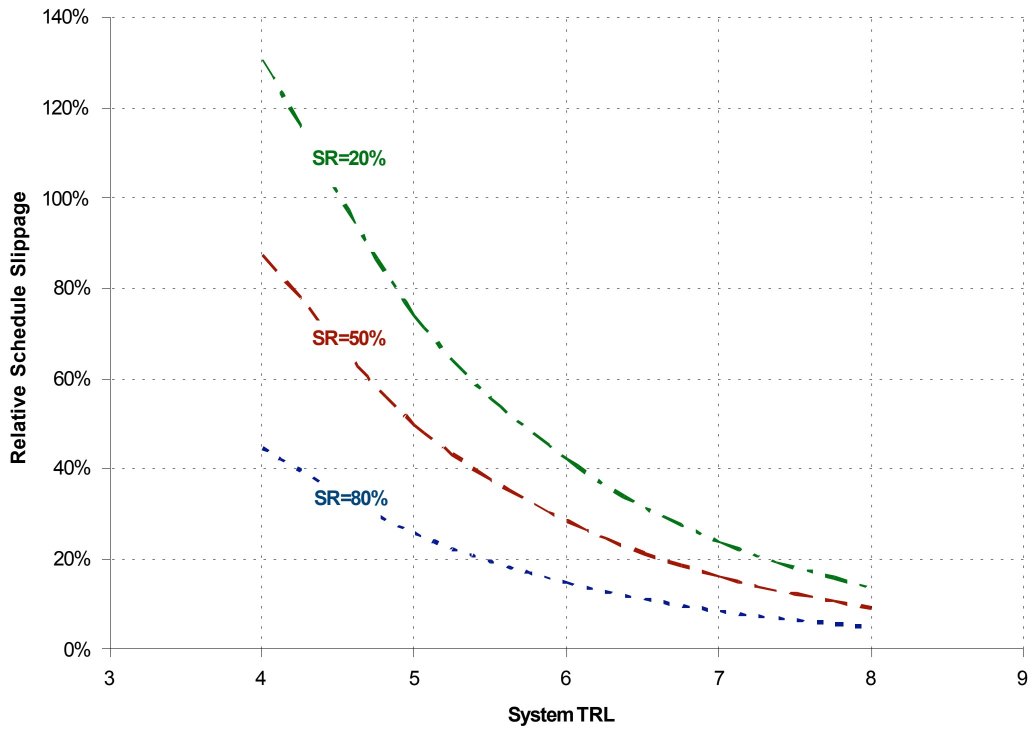

Figure 3. TRL-schedule risk curves (SR) for a normally distributed Relative Schedule Slippage

Schedule risk curves, as shown in Figure 3, are useful tools for program managers and decision-makers. These curves are can help in the following way:

- When looked at through a "vertical" cut (i.e., for a given TRL at start of the program), schedule risk curves allow program managers to visualize the incremental schedule margin needed to achieve a **reduction in schedule risk**. For example, on Figure 3, we see that for TRL = 6, a schedule margin of 30% is required to achieve a schedule risk of 50%; however, if this schedule risk is unacceptable, and, say a 20% schedule risk is required, then a schedule margin of 43% is required.

- When looked at through a "horizontal" cut, the schedule risk curves allow program managers to visualize for a given schedule margin, the change (reduction) in schedule risk as the TRL changes (increases). For example, a 40% schedule margin for TRL = 4 will result in a very high 80% schedule risk; however this same 40% margin will ensure approximately a 20% schedule risk if the technologies considered for the program were more mature with a TRL = 6.

In short, schedule risk curves allow program managers to make risk-informed decisions regarding the appropriate schedule margins for a given program, or the appropriate TRL to consider in an acquisition program should the program's schedule be exogenously constrained. These points are further discussed in the Section V along with our recommendations.

Caveat: Figure 3 was created based on the model developed in section IV-A and IV-B, and given the assumption that RSS (and consequently FTD) is normally distributed. Should more RSS empirical data be available and warrant a different probability distribution function, schedule risk curves can still be developed by replacing Eq. 11–16 with the new probability distribution function (pdf) and its parameters; schedule risk however, as defined in Eq. 9 and 10, remains the same irrespective of the pdf considered. The concept and usefulness of the schedule risk curves to program managers and decision-makers are independent on the particular distribution function for RSS (or FTD).

V. Implications and Recommendations

Schedule slippage plagues the space industry, and is antinomic with the recent emphasis on space responsiveness. This work focused on one key culprit driving schedule slippage, namely TRL. Our empirical results, based on a data set of 28 NASA programs, support GAO's recommendation that, in order to reduce schedule risk, no technologies below TRL = 6 or 7 be included in an acquisition program:

> *"If programs adhere to the TRL = 6 criteria, they will greatly reduce the risk of encountering costly technical delays, though not completely […]. Moreover, the best practice programs we have studied strive for a TRL = 7, where the technology has been tested in an operational environment, that is, space."*[5]

Furthermore, the existence of a non-zero mean schedule slippage in our study strongly suggests the need to include adequate schedule margins to the initial schedule estimate (IDE). The schedule margin can be adapted to the schedule risk the program is willing to accept; however, at a minimum, we recommend that programs adopt a schedule margin that is equal (or greater) to the mean schedule slippage for a given TRL. For example, if the TRL = 6 criteria is adopted, we recommend that at a minimum, a schedule margin of 30% be adopted, since:

$$sm_6 \geq \left\langle \overline{\mathrm{RSS}} \right\rangle_{TRL=6} \approx 30\%$$

But beyond a schedule point estimate, we recommend, 1) that the industry adopts and develops schedule risk curves in space acquisition programs; 2) that these schedule risk curves be made available to policy- and decision-makers; and 3) that adequate schedule margins be defined according to an agreed upon acceptable schedule risk level. For example, under the assumption of IV-C, at TRL = 6, we would still incur a schedule risk of 50% with a schedule margin of 30%. In order to "reach" the 20% schedule risk curve, we would need to include a schedule margin of 43%. Said differently, with a 43% schedule margin, we are 80% confident that the actual program's schedule will be within our initial estimate plus margin. Table 5 shows three levels of schedule risk and the required or corresponding schedule margins.

Table 5. Schedule risk and the corresponding schedule margin for a TRL = 6 system

(agreed upon) Schedule risk	Corresponding schedule margin
50%	30%
20%	43%
5%	56%

Schedule margins and cost implications: It is important to recognize that while schedule margins decrease schedule risk, they do come at a cost. In other words, increasing a schedule margin will increase a program's budget, and many programs may not be able to afford large schedule margins (e.g., the 56% schedule margin in Table 5, which results in a trifle schedule risk of 5%). Program managers should therefore carefully balance the cost implications of schedule margins with the schedule risk that the program can or is willing to afford. In future work, we propose to investigate the (marginal) cost of schedule margins, and to develop an analytical tool for program managers to assess and trade cost, schedule risk, and schedule margins.

Finally, we hope that more TRL and schedule data will be made available (to academics) in the future, especially for Air Force and DOD programs. This would allow to update and expand the results of the current work, and to include if possible a comparative analysis of schedule risk in various acquisition programs (space and other weapon systems).

Acknowledgments

We gratefully acknowledge the support of Dale Thomas from NASA Marshall Space Flight Center, and Tzesan Lee who provided us with very helpful information about the data used in this study.

References

[1] "Defense Acquisitions: Space System Acquisition Risks and Keys to Addressing Them", GAO-06-776R, June 2006. URL: http://www.gao.gov/docsearch/repandtest.html [cited 26 July 2007].

[2] "Technology Development: New DOD Space Science and Technology Strategy Provides Basis for Optimizing Investments, but Future Versions Need to be More Robust," GAO-05-155, Jan. 2005. URL: http://www.gao.gov/docsearch/repandtest.html [cited 26 July 2007].

[3] "Defense Acquisitions: Improvements Needed in Space Systems Acquisition Management Policy", GAO-03-1073, Sep. 2003. URL: http://www.gao.gov/docsearch/repandtest.html [cited 26 July 2007].

[4] "Defense Acquisitions: Assessments of Selected Weapon Programs", GAO-07-406SP, Mar. 2007. URL: http://www.gao.gov/docsearch/repandtest.html [cited 26 July 2007].

[5] "Space Acquisitions: Actions Needed to Expand and Sustain Use of Best Practices", GAO-07-730T, Apr. 2007. URL: http://www.gao.gov/docsearch/repandtest.html [cited 26 July 2007].

[6] Rothenflue, J.R., Kwolek, M.J., "Streamlining DOD Acquisition: Balancing Schedule with Complexity", September 2006, Chapter 3 by Chittenden, J. L., "Faster is Better… Can the USAF Acquisition Process Be SAIV'D?" pp. 81,82. URL: www.au.af.mil/au/awc/awcgate/cst/streamlining_acq.pdf [cited 2 August 2007].

[7] Mankins, J. C., "Technology Readiness Levels," NASA Office of Space Access and Technology, White Paper, 1995. URL: www.hq.nasa.gov/office/codeq/trl/trl.pdf [cited 2 August 2007].

[8] Browning, T. R., "Sources of Schedule Risk in Complex System Development", *Systems Engineering*, Vol. 2, No. 3, 1998, pp. 129-142.

[9] Gupta, A. and Wilemon, D., "Accelerating the development of technology-based new products", *California Management Review*, Vol. 32, No.2, Winter 1990, pp. 24-44.

[10] "Best Practices: Better Management of Technology Development Can Improve Weapon Systems Outcome," GAO/NSIAD-99-162, pp.14-15-16. Jul. 1999. URL: http://www.gao.gov/docsearch/repandtest.html [cited 26 July 2007].

[11] George, T., Powers, R., “Closing the TRL gap”, *Aerospace America*, August 2003, pp 24-26.

[12] Lee, T., Thomas, L. D., “Cost Growth Models for NASA’s Programs”, *Journal of Probability and Statistical Science*, Vol. 1, 2003, pp. 265-279.

[13] Lyman-Ott, R., Longnecker, M., *Statistical Methods and Data Analysis*. 5th ed., Duxbury Thomson Learning, Pacific Grove, CA, 2001, pp. 91, 92.

[14] Yates, R. D., Goodman, D. J., *Probability and Stochastic Processes*. 2nd ed., Wiley, Hoboken, NJ, 2005, pp. 123–124.