JOURNAL OF SPACECRAFT AND ROCKETS
Vol. 45, No. 4, July–August 2008

Technology Readiness Level, Schedule Risk, and Slippage in Spacecraft Design

Gregory F. Dubos,* Joseph H. Saleh,† and Robert Braun‡
Georgia Institute of Technology, Atlanta, Georgia 30332

DOI: 10.2514/1.34947

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 U.S. Department of Defense in keeping its acquisition of space systems on schedule, and they identified the low Technology Readiness Level of the system/payload under development as a principal culprit driving schedule risk and slippage. In this paper, based on data from past space programs, we analyze 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 Technology Readiness Level schedule-risk curves to help program managers make risk-informed decisions regarding the appropriate schedule margins for a given program or the appropriate Technology Readiness Level to consider if the program's schedule were to be exogenously and rigidly constrained. Based on our findings, we recommend that the industry adopt and develop schedule-risk curves (instead of single-schedule point estimates), that these schedule-risk curves be made available to policy- and decision-makers in acquisition programs, and that adequate schedule margins be defined according to an agreed-upon and acceptable schedule-risk level.

Nomenclature

env(RSS)	=	envelope of relative schedule slippage
FTD	=	final total schedule duration
$f(\cdot)$	=	probability density function
IDE	=	initial schedule-duration estimate
LB	=	lower bound of the relative schedule slippage envelope
R^2	=	coefficient of determination in a regression analysis
RSS	=	relative schedule slippage
$\langle \mathrm{RSS} \rangle$	=	sample average of relative schedule slippage
$\langle \overline{\mathrm{RSS}} \rangle$	=	modeled average of relative schedule slippage
sm	=	schedule margin
TRL	=	Technology Readiness Level
UB	=	upper bound of the relative schedule slippage envelope
$\langle \overline{\mathrm{UB}} \rangle$	=	modeled upper bound of the relative schedule slippage envelope
WTRL	=	weighted average Technology Readiness Level
Z	=	standard normal random variable
α	=	multiplicative constant of the mean relative schedule slippage exponential model
α'	=	multiplicative constant of the upper-bound relative schedule slippage exponential model
λ	=	rate of the mean relative schedule slippage exponential model
λ'	=	rate of the upper-bound relative schedule slippage exponential model
μ	=	mean of the random-variable relative schedule slippage
σ	=	standard deviation of the random-variable relative schedule slippage

I. Introduction

THE Government Accountability Office (GAO) has repeatedly noted [1–5] the difficulties encountered by the U.S. Department of Defense (DOD) in keeping its acquisition of space systems on schedule (and within budget) [1]: "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."

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, can 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 within a very short time frame) is fundamentally at odds with schedule slippages in acquisition programs.

Identifying drivers 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 the 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; the more complex the system, the less likely that its development can remain on schedule. Requirements can also change while the system is still under development or the customer can modify and expand the scope of the system's performance and capabilities while the latter is still under development (e.g., the customer adds new requirements after

Presented as Paper 6020 at Space 2007, Long Beach, CA, 18–20 September 2007; received 2 October 2007; accepted for publication 26 February 2008.

*Graduate Research Assistant, School of Aerospace Engineering; greg.dubos@gatech.edu. Student Member AIAA (Corresponding Author).

†Assistant Professor, School of Aerospace Engineering. Senior Member AIAA.

‡David and Andrew Lewis Professor of Space Technology, School of Aerospace Engineering. Fellow AIAA.

the program requirements have been established, reviewed, and approved); this situation is not uncommon in product development and system design and is referred to as requirements creep. Requirements creep can significantly perturb the development schedule of a system and can result in major delays. Other factors that can (and often do) cause schedule slippage are funding delays or breaks in funding during the development of a system. When a customer or a funding agency delays or suspends funding after a project has been approved or initiated, several time-consuming complications result from this funding discontinuity, ranging from the suspension then reinitialization of new analyses when the program is restarted to the deployment and training of a new workforce. These factors imply that the resultant schedule slippage is often greater than the original funding delay that triggered the schedule slippage.

In summary, there are multiple causes of schedule slippage in space programs and they are not necessarily independent (in a statistical sense). Some of these causes, as discussed in this section, are beyond the control of the satellite manufacturers (e.g., requirement creeps and funding delays are primarily driven by the customer or the funding agency).

This paper deals with another factor that has been identified in several GAO reports as a principal culprit associated with schedule slippage: namely, the low Technology Readiness Level (TRL) of the system/payload under development [3,4,7–9]. In the following, the TRL is used as a proxy for technology maturity (or lack of) and is considered to be the independent variable in our analyses.

The purpose of this paper is to empirically explore, in the case of space systems, the correlation, if any, between the TRL and schedule slippage and how much schedule risk/slippage is associated with various levels of technology maturity or the TRL and to identify appropriate schedule margins to mitigate this risk.

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 the TRL as the index), and we propose to characterize the central tendency and dispersion of this random variable as a function of the TRL (the independent variable in this study) through data analysis and modeling. To this effect, in Sec. II, we briefly describe the TRL concept and highlight the relationship between technology uncertainty (i.e., low TRL) and schedule risk. In Sec. III, we discuss and characterize the data used in our study. In Sec. IV, we analyze the data and provide analytical models for schedule risk and slippage as a function of the TRL (and schedule margins). In Sec. V, we conclude with some implications and recommendations from our work.

II. Technology Readiness Level and Schedule Risk

The TRL is a widely adopted metric by NASA and the DOD; it was introduced by NASA in the 1980s to assess the maturity of a particular technology before its implementation in a system and to allow the "consistent comparison of maturity between different types of technology" [10]. 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 [10] for a more detailed description of these levels.

The lack of technology maturity (or a low TRL, sometimes described in the literature as technology uncertainty) is often associated with schedule risk, albeit qualitatively. Browning [11] defined 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 [12] 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, the more knowledge 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.

Table 1 Summary of different Technology Readiness Levels

TRL	Summary description
TRL 1	Basic principle observation and reporting
TRL 2	Technology concept and/or application formulation
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 relevant environment (ground or space)
TRL 7	System prototype demonstration in 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

As often noted by GAO, schedules overruns are most likely to occur when the burden of technology maturation is assumed within an acquisition program [7]. Indeed, because the low-TRL world (research environment, or science and technology 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 important, how much time it will take to bring a low-TRL technology (e.g., $\mathrm{TRL} = 4$) to a comfortable level of maturity (e.g., $\mathrm{TRL} = 8$). This issue is sometimes referred as the TRL gap and is described by George and Powers [13] 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 the 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 the Technology Readiness Level is a central theme in feasibility studies of system design (spacecraft and other), limited TRL data are 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 the 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. Most of the programs considered here are unmanned and include Earth science missions and interplanetary probes. Lee and Thomas [14] used these data to construct probability-based models for the cost growth of NASA's programs. Details about these data can be found in [14]. In this study, we focus instead on schedule slippage and concern ourselves with three parameters from the data set: 1) TRL at the start of the program, 2) initial schedule-duration estimate (IDE), and 3) final total schedule duration (FTD).

We define the relative schedule slippage (RSS) as the percentage of schedule growth, given the initial schedule estimate

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

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 the TRL and RSS. We analyze both the central tendencies and the dispersion of RSS as a function of the TRL and relate our results to schedule risk and slippage. The details are further discussed in Sec. IV.

Before we proceed, however, a subtlety concerning the TRL data should be addressed. TRLs usually define the maturity of a given technology and, by extension, a TRL value is commonly assigned to a component characterized by one single technology. However, to extend the notion of technology maturity to an entire program, an average TRL value for a complex system must be defined. Lee and Thomas [14] calculated a weighted average of the TRL (WRTL) for each program by taking the TRL of each component multiplied by their corresponding percent of the allocated cost against the entire program's cost:

$$\mathrm{WTRL}_{\mathrm{program}} = \sum_{\mathrm{components}\ c_i} w_i \cdot \mathrm{TRL}_{c_i} \tag{2}$$

where

$$w_i = \frac{\mathrm{cost}_i}{\mathrm{cost}_{\mathrm{program}}}$$

For example, a complex system such as the Hubble Space Telescope is first broken down into subsystems (e.g., attitude control), which are then decomposed into components (e.g., control-moment gyros).[§] The TRL of each component is then considered to regressively define the WRTL. In our study, we used the WRTL as a preliminary basis for the average system TRL, for which the influence on schedule slippage was investigated. The WRTL 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 WRTL for our schedule analysis. However, this WRTL 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 of taking the integer part of the WRTL. Here again, when considering components requiring a large resource investment, we contend that those with the lowest TRLs drive the schedule delays, because they represent the slowest links of the maturation chain. For example, consider a program for which $\mathrm{WRTL} = 4.62$. If it involves components with $\mathrm{TRL} = 5$ or 6, it also involves components with integer values of $\mathrm{TRL} \leq 4$. First, the WRTL of 4.62 gives a good indicator of the average TRL of the entire system. Then, considering that components with a low TRL (e.g., $\mathrm{TRL} = 4$) have a bigger impact on schedule slippage than do components with $\mathrm{TRL} = 5$, we retained the integer value: that is, $\mathrm{TRL} = 4$. (Following this logic, one could argue that the minimum of all the components' TRLs could be directly used in place of the WRTL. However, we think it is important to first capture the relative importance of every component in terms of the amount of resources spent. The WRTL provides this function.)

IV. Modeling Schedule Slippage and Risk

For each of the 28 NASA programs in our data set, we plot and analyze the doublet (TRL and RSS) in which the TRL consists of the integer values discussed in the previous section. The TRLs in our data set range from four to eight. We consider the relative schedule slippage to be a random variable: more precisely, a random vector or an indexed family of random variables with the 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, the TRL.

[§]Personal communication with D. Thomas, Aug. 2007.

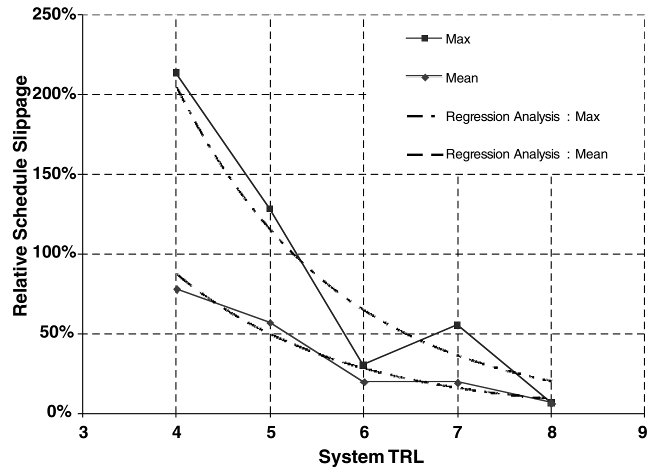

Fig. 1 RSS for 28 NASA programs (mean, maximum, and regression analysis) as a function of TRL; tabular data for this figure can be found in Table A1.

A. Mean Relative Schedule Slippage

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

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

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

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

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

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

We chose this model structure for both its simplicity and its 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 important, 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 the TRL increases. Condition 2 can be stated mathematically as follows: the absolute value of the derivative of the $\langle \overline{\mathrm{RSS}} \rangle$ with respect to the TRL should be a decreasing function, hence our choice of a decreasing exponential function (more details in the following sections).

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

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

Model parameter	Value
α	8.29
λ	0.56
R^2	0.94

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

TRL	Observed mean relative schedule slippage $\langle \overline{\mathrm{RSS}} \rangle_j$, %	$\langle \overline{\mathrm{RSS}} \rangle_j$, %	Error
4	78	88	10
5	57	50	7
6	20	29	9
7	19	16	3
8	7	9	2

the coefficient of determination R^2 (94%) and by the error between the model output and the observed data (less than 10%).

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 six points for each TRL), the R^2 value of our model (94%) should be considered with caution and not interpreted beyond the fact that it indicates an accurate model.

If y_i are the values of the dependent variable considered, $\hat{y}_i$ are the fitted values, and $\bar{y}$ is the sample mean, then 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.

B. Dispersion of the Relative Schedule Slippage

In addition to the mean relative schedule slippage, the data allow 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:

$$\begin{cases} \mathrm{UB}_j = \max_i (\mathrm{RSS}_i)|_{\mathrm{TRL}=j} \\ \mathrm{LB}_j = \min_i (\mathrm{RSS}_i)|_{\mathrm{TRL}=j} \end{cases} \tag{5}$$

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

$$\begin{cases} \mathrm{env}(\mathrm{RSS})_j = \{\mathrm{UB}_j; \mathrm{LB}_j\} \\ \mathrm{dispersion}_j = \mathrm{UB}_j - \mathrm{LB}_j \end{cases} \quad \text{for } j = 4, 5, 6, 7, 8 \tag{6}$$

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 IDE almost matched the 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):

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

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

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

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

Figure 1 shows that the dispersion of RSS narrows down as the TRL increases. This dispersion can be considered as a proxy for the time uncertainty in the technology-maturation process (the lower the TRL, the bigger the schedule uncertainty; that is, the larger variability will be around FTD). The 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 stage 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):

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

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.

Beyond the schedule estimation errors reflected by the mean RSS model [Eq. (4) and Table 2] (which may be due to a variety of factors, including intrinsically flawed schedule estimation methods in use by the industry), 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., the complexity of the system under development, experience of the program manager, funding delays, requirements creep, etc.). Further research is needed to identify the other parameters and correlate them with schedule slippage. Such work would be particularly relevant to the space industry, because it will 1) help develop credible schedule estimates, 2) limit programs' schedule risk, and 3) identify and disseminate the best practices related to maintaining acquisition programs on schedule. The GAO reports mentioned previously (see Sec. 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 nonzero mean relative schedule slippage strongly suggests the need to add a schedule margin to the 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 the start of the program, we calculate the probability of overshooting the initial schedule estimate

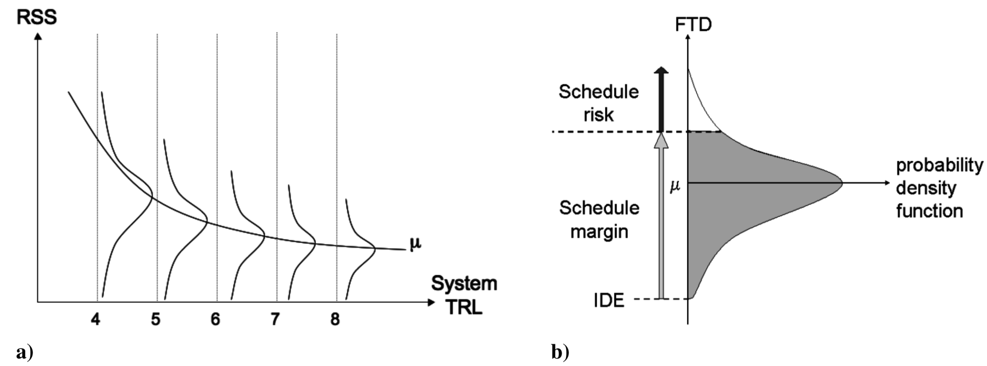

Fig. 2 Characterization of schedule slippage as a random variable: a) RSS probability density functions for each TRL value and b) FTD probability density function and schedule risk.

plus a given schedule margin. This will in turn help us determine appropriate schedule margins, as discussed subsequently.

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 lowercase to denote the relative value and uppercase 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 Fig. 2b. Mathematically, we express the schedule risk for a given TRL at the start of the program as follows:

$$\text{schedule risk}\bigg|_{\text{TRL}=j} = P\{\text{FTD}_j > \text{IDE}_j \cdot (1 + \text{sm}_j)\} \tag{9}$$

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. Note that Eq. (9) is equivalent to the following:

$$\text{schedule risk}\bigg|_{\text{TRL}=j} = P\{\text{RSS}_j > \text{sm}_j\} \tag{10}$$

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 are 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 a TRL schedule-risk curve, let us assume that for a given TRL, RSS (or FTD) is normally distributed. That is, we formulate the hypothesis that the RSS (or FTD) has a normal probability density function, which can be written as follows:

$$f(\text{RSS}_j)\bigg|_{\text{TRL}=j} = \frac{1}{\sigma_j\sqrt{2\pi}} e^{-\frac{(\text{RSS}_j-\mu_j)^2}{2\sigma_j^2}} \tag{11}$$

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

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:

$$\sigma \approx \frac{\text{dispersion}}{4} \tag{12}$$

The reader is referred to Lyman-Ott and Longnecker [15] 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:

$$\begin{cases} \mu_j \approx \langle\overline{\text{RSS}}\rangle|_{\text{TRL}=j} \\ \sigma_j \approx \frac{\text{dispersion}_j}{4} \end{cases} \quad \text{for } j = 4, 5, 6, 7, 8 \tag{13}$$

The numerical values for μ_j and σ_j used in the following are provided in Table A2. Given a system TRL at the start of a program and an at 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 Eqs. (9) and (10). Equation (10) can be rewritten in terms of the standard normal distribution as follows:

$$\text{schedule risk}\bigg|_{\text{TRL}=j} = 1 - P\{\text{RSS}_j \le \text{sm}_j\}$$
$$= 1 - P\{Z \le z\} = 1 - \phi(z) \tag{14}$$

with

$$Z = \frac{\text{RSS}_j - \mu_j}{\sigma_j} \tag{15}$$

and

$$z = \frac{\text{sm}_j - \mu_j}{\sigma_j} \tag{16}$$

where Z is the standardized random variable derived from RSS, and ϕ is the cumulative density function of the standard normal distribution function, for which the values can be found in tabulated form in the literature (see, for example, [16]).

1. Numerical Example

What should the schedule margin be if a schedule risk of 20% is called for, and the TRL = 6 at the start of the program? 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 42%. In other words, for TRL = 6 at the start of a program, a schedule margin of 42% will ensure a 20% schedule risk (or conversely, that there is an 80% likelihood that the FTD will fall below the IDE plus the 42% schedule margin). When done for all TRLs, not just for 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 Fig. 3.

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

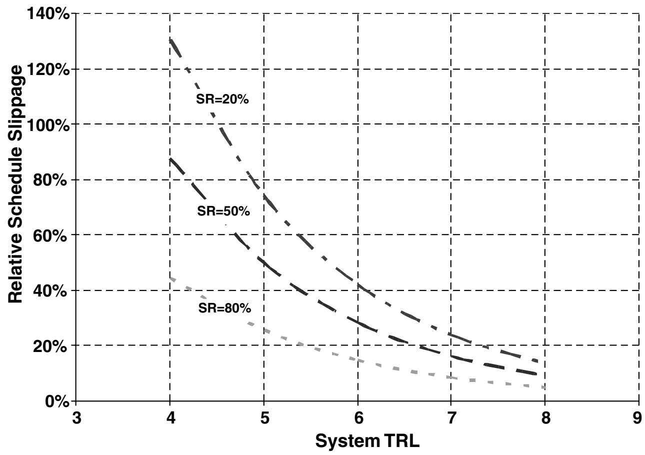

Fig. 3 TRL schedule-risk (SR) curves for a normally distributed relative schedule slippage.

1) When looked at through a vertical cut (i.e., for a given TRL at the 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, in Fig. 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 42% is required.

2) When looked at through a horizontal cut, the schedule-risk curves allow program managers to visualize the change (reduction) in schedule risk as the TRL changes (increases) for a given schedule margin. For example, a 40% schedule margin for TRL = 4 will result in a very high 80% schedule risk; however, the same 40% margin would ensure approximately a 20% schedule risk if the technologies considered for the program were more mature, with 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 for the appropriate TRL to consider in an acquisition program if the program's schedule were to be exogenously constrained. These points are further discussed in Sec. V, along with our recommendations.

2. Caveat

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

V. Conclusions

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, the TRL. Our empirical results, based on a data set of 28 NASA programs, support GAO's recommendation that to reduce schedule risk, no technologies below TRL = 6 or 7 be included in an acquisition program [5]: "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 studied strive for a TRL = 7, where the technology has been tested in an operational environment, that is, space."

Furthermore, the existence of a nonzero mean schedule slippage in our study strongly suggests the need to include adequate schedule margins to the 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 to (or greater than) the mean schedule slippage for a given TRL. For example, if the TRL = 6 criterion is adopted, we recommend that at a minimum, a schedule margin of 30% be adopted, because

$$\mathrm{sm}_6 \geq \langle \overline{\mathrm{RSS}} \rangle_{\mathrm{TRL}=6} \approx 30\%$$

But beyond a schedule point estimate, we recommend the following:

1) The industry should adopt and develop schedule-risk curves in space acquisition programs.

2) The schedule-risk curves should be made available to policy- and decision-makers.

3) Adequate schedule margins should be defined according to an agreed-upon acceptable schedule-risk level.

For example, under the assumption of Sec. IV.C, at TRL = 6, we would still incur a schedule risk of 50% with a schedule margin of 30%. To reach the 20% schedule-risk curve, we would need to include a schedule margin of 42%. Said differently, with a 42% 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.

Throughout this paper, we have avoided assigning blame for schedule slippage on any particular stakeholder in the space industry. As noted in the Introduction, several causes of schedule slippage, such as requirements creep and funding delays, are beyond the control of satellite manufacturers and are primarily driven by the customer or the funding agency. Furthermore, the choice of a low-TRL payload can be made independently by the manufacturer or it can be dictated by customer preference (or pressure); in both cases, this choice is likely to result in schedule slippage. Schedule slippage is therefore a delicate joint responsibility of the customer and the manufacturer. The present work can help in the following manner: should low-TRL technology be used in a system, acknowledging the risk of schedule slippage (and quantifying it) is essential for all stakeholders. The TRL schedule-risk curves allow all stakeholders to recognize and agree on the amount of schedule risk associated with the use of a given level of technology maturity.

It is important to recognize that although 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.

Finally, we hope that more TRL and schedule data will be made available (to academics) in the future, especially for U.S. Air Force and DOD programs. This would allow updating and expanding the results of the current work and including (if possible) a comparative analysis of schedule risk in various acquisition programs (space and other weapon systems, which would allow us to benchmark different acquisition policies and to identify best practices).

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

Agreed-upon schedule risk, %	Corresponding schedule margin, %
50	30
20	42
5	56

Appendix: Tabular Data for Observations and Models

Table A1 Tabular data for Fig. 1

TRL	Observed mean relative schedule slippage $\langle RSS \rangle|_{TRL=j}$, %	Modeled mean relative schedule slippage $\langle \overline{RSS} \rangle|_{TRL=j}$, %	Observed maximum relative schedule slippage $\langle UB \rangle|_{TRL=j}$, %	Modeled maximum relative schedule slippage $\langle \overline{UB} \rangle|_{TRL=j}$, %
4	78	88	214	205
5	57	50	128	115
6	20	29	30	65
7	19	16	55	37
8	7	9	7	21

Table A2 Normal distribution parameters of RSS for each TRL value

TRL	Mean $\mu_j \approx \langle \overline{RSS} \rangle|_{TRL=j}$, %	Standard deviation $\sigma_j \approx (\langle \overline{UB} \rangle|_{TRL=j})/4$, %
4	87.7	51.3
5	50.0	28.9
6	28.5	16.2
7	16.3	9.1
8	9.3	5.1

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. We also wish to thank the Editor-in-Chief, Associate Editor, and three anonymous reviewers for their comments on an earlier version of this paper.

References

[1] "Defense Acquisitions: Space System Acquisition Risks and Keys to Addressing Them," U.S. Government Accountability Office, Rept. GAO-06-776R, June 2006; also available online at http://www.gao.gov/new.items/d06776r.pdf [retrieved 15 Apr. 2008].

[2] "Technology Development: New DOD Space Science and Technology Strategy Provides Basis for Optimizing Investments, but Future Versions Need to be More Robust," U.S. Government Accountability Office, Rept. GAO-05-155, Jan. 2005; also available online at http://www.gao.gov/new.items/d05155.pdf [retrieved 15 Apr. 2008].

[3] "Defense Acquisitions: Improvements Needed in Space Systems Acquisition Management Policy," U.S. Government Accountability Office, Rept. GAO-03-1073, Sept. 2003; also available online at http://www.gao.gov/new.items/d031073.pdf [retrieved 15 Apr. 2008].

[4] "Defense Acquisitions: Assessments of Selected Weapon Programs," U.S. Government Accountability Office, Rept. GAO-07-406SP, Mar. 2007; also available online at http://www.gao.gov/new.items/d07406sp.pdf [retrieved 15 Apr. 2008].

[5] "Space Acquisitions: Actions Needed to Expand and Sustain Use of Best Practices," U.S. Government Accountability Office, Rept. GAO-07-730T, Apr. 2007; also available online at http://www.gao.gov/new.items/d07730t.pdf [retrieved 15 Apr. 2008].

[6] Chittenden, J. L., "Faster is Better...Can the USAF Acquisition Process Be SAIV'D?," *Streamlining DOD Acquisition: Balancing Schedule with Complexity*, edited by J. R. Rothenflue, and M. J. Kwolek, Air Univ., Maxwell-Gunter AFB, AL, Sept. 2006, pp. 81, 82; also available online at www.au.af.mil/au/awc/awcgate/cst/streamlining_acq.pdf [retrieved 15 Apr. 2008].

[7] "Best Practices: Better Management of Technology Development Can Improve Weapon Systems Outcome," U.S. Government Accountability Office, Rept. GAO/NSIAD-99-162, July 1999, pp. 14–15–16; also available online at http://www.gao.gov/archive/1999/ns991620.pdf [retrieved 15 Apr. 2008].

[8] "NASA: Implementing a Knowledge-Based Acquisition Framework Could Lead to Better Investment Decisions and Projects Outcomes," U.S. Government Accountability Office, Rept. GAO-06-218, Dec. 2005, pp. 15–16; also available online at http://www.gao.gov/new.items/d06218.pdf [retrieved 15 Apr. 2008].

[9] "Defense Acquisitions: Assessments of Selected Major Weapon Programs," Rept. GAO-05-301, Mar. 2005, pp. 10–11; also available online at http://www.gao.gov/new.items/d05301.pdf [retrieved 15 Apr. 2008].

[10] Mankins, J. C., "Technology Readiness Levels," NASA Office of Space Access and Technology, 1995; also available online at www.hq.nasa.gov/office/codeq/trl/trl.pdf [retrieved 15 Apr. 2008].

[11] Browning, T. R., "Sources of Schedule Risk in Complex System Development," *Systems Engineering*, Vol. 2, No. 3, 1999, pp. 129–142. doi:10.1002/(SICI)1520-6858(1999)2:3<129::AID-SYS1>3.0.CO;2-H

[12] Gupta, A., and Wilemon, D., "Accelerating the Development of Technology-Based New Products," *California Management Review*, Vol. 32, No. 2, 1990, pp. 24–44.

[13] George, T., and Powers, R., "Closing the TRL Gap," *Aerospace America*, Vol. 41, No. 8, Aug. 2003, pp. 24–26.

[14] Lee, T., and Thomas, L. D., "Cost Growth Models for NASA's Programs," *Journal of Probability and Statistical Science*, Vol. 1, No. 2, Aug. 2003, pp. 265–279.

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

[16] Yates, R. D., and Goodman, D. J., *Probability and Stochastic Processes*, 2nd ed., Wiley, New York, 2005, pp. 123–124.

O. de Weck
Associate Editor