SYM-AM-19-064

PROCEEDINGS
OF THE
SIXTEENTH ANNUAL ACQUISITION RESEARCH SYMPOSIUM

THURSDAY SESSIONS
VOLUME II

Acquisition Research:
Creating Synergy for Informed Change

May 8–9, 2019

Published: April 30, 2019

Prepared for the Naval Postgraduate School, Monterey, CA 93943.

System Maturity Estimation During Program Execution

John C. Kamp—is a retired naval officer and served as a program manager. He has over 30 years of direct experience in defense operations, research, engineering, program management, resource sponsorship, requirements generation, and acquisition spanning multiple domains and services. He received his Doctor of Engineering degree in Engineering Management from The George Washington University in 2019. Dr. Kamp is a Fellow in the Royal Institution of Naval Architects and has been a member of the ACM and IEEE since 2005. [jckamp2@gmail.com].

Ebrahim Malalla—is a Visiting Associate Professor of Engineering and Applied Science at The George Washington University. He worked previously as an Assistant Professor at Ahlia University, Bahrain, and at King Fahd University for Petroleum and Minerals, Saudi Arabia. While working at Ahlia University, he held the positions of the President Assistant for University Compliance (2015–2017), the Associate Dean of Graduate Studies and Research (2012–2014), the Chairperson of the IT Department (2008–2011), and the Chairperson of the M.Sc. Program in Engineering Management offered by The George Washington University at Ahlia University (2012–2017). Dr. Malalla received his PhD in Probabilistic Analysis of Algorithms from McGill University, Montreal, Canada, in 2004. His research interests include multidisciplinary problems in computer science, mathematics and engineering management and have led to many papers published in refereed international journals. [emalalla@gwu.edu].

Abstract

Defense acquisition programs integrate mature and immature new technologies into developing and in-service systems to offset future threats and needs. Mature technologies may be nearly ready-for-use; less mature technologies may mitigate anticipated threats or create new capabilities but may also take more time to develop and integrate into a system leading to schedule growth. The Department of Defense uses Technology Readiness Assessments to assess system technology maturity and to satisfy statutory requirements to evaluate system technical readiness prior to starting system development. The Government Accountability Office independently conducts annual assessments of selected weapon system programs. These are useful but require program offices to expend significant time and effort as part of program execution. This research examines different measures of technology and system maturity and identifies maturity-related factors. Regression analysis is used to identify statistically significant predictors of program technology and system maturity and schedule growth. The research results provide program offices insight into technology and system maturity and the sources of schedule growth based on resource, programmatic, operational testing, and schedule-related factors, allowing them to monitor and adjust acquisition program planning and execution.

Introduction

A recent unclassified summary of the National Defense Strategy described a changing acquisition strategy emphasizing speed of delivery as part of a response to capable, innovative adversaries (Mattis, 2018). One way that Department of Defense (DoD) Major Defense Acquisition Programs[1] (MDAPs) respond to an adversarial threat is by

[1] MDAPs are weapon system programs with research and development expenditures greater than $300 million or procurement expenditures greater than $1.80 billion indexed to fiscal year 1990 constant dollars (MDAP Defined, 2007).

integrating proven (*mature*) and emerging (*immature*) technologies into new and in-service systems. Katz et al. (2015) noted that DoD programs may select less mature technologies to hedge system performance against future threats or to create new capabilities but use more mature technologies to reduce the likelihood of schedule growth. The problem is that MDAP schedules can grow by over 25% when integrating immature systems. In 2013, the Government Accountability Office (GAO) reported an average schedule delay of 27 months, or 37%, for MDAPs to deliver an initial capability (Dodaro, 2013).

Maturity-Related Terms and Measures

System maturity is different than technology maturity. *System maturity* means the system satisfies the *design requirements*. The literature describes system maturity in terms of requirements validation (Tetlay & John, 2009) and includes validating functional requirements (Gove & Uzdzinski, 2013). *Technology maturity* describes how well a technology is understood. During system development, the DoD focuses on technology maturity instead of system maturity (Ramirez-Marquez & Sauser, 2009).

The GAO assesses MDAP *technology maturity, design maturity,* and *production maturity* as part of its annual independent assessments of selected DoD weapon system programs (Dodaro, 2007). Katz et al. (2015) showed that GAO technology maturity occurs for a system when the TRLs of all critical technologies were greater than or equal to 7. For the MDAPs considered in this research, most (about 54%) achieved GAO technology maturity, fewer (about 41%) achieved design maturity, and few (about 7%) achieved production maturity.

Product maturity reflects a product's market position. Day (1981) described product maturity in terms of customer understanding (learning), market share (potential) and competition (turbulence). Mature products respond more to customer and competitive demands (orientations) than to innovation (Wang, Wang, & Zhao, 2015). Nolte (2008) states that technology maturity is related to how *well* something is understood, while product maturity includes concepts of product obsolescence and competitive market share.

Product quality is a measure of how well a product meets customer requirements (Kandt et al., 2016). Azizian et al. (2011) identified the relationship between product (system) quality as measured by existing international standards and technology maturity, and found that system *development* and *operational tests*, system *prototyping,* and actual system *demonstrations* were statistically significant predictors of product quality.

Readiness describes context-specific system suitability for use (Tetlay & John, 2009), which is similar to maturity. Technology Readiness Assessments (TRAs) are used by the DoD to assess system technology maturity and to satisfy statutory requirements to evaluate system technical readiness prior to starting system development (Weapon System Acquisition Reform Act [WSARA], 2009). Bailey et al. (2014) noted that the TRA process is qualitative and subjective, and found the underlying system engineering activities, *not the TRA itself*, were statistically significant predictors of quality and program performance.

Technology Readiness Levels (TRLs) are used by the DoD to indicate where the maturity of either a technology or system lies within a qualitative nine-level ordinal scale (Mankins, 2009). Note that TRLs do not by themselves characterize risk or difficulty of progressing between levels (Conrow, 2011), nor do they describe integration readiness or risk (Ramirez-Marquez & Sauser, 2009). TRLs are characterized by completion of discrete events and activities, but are typically not reported in the public literature. There are tools such as TRL calculators (Nolte, 2008) to help consistently assign TRLs.

This paper summarizes recent research that explored how MDAP system maturity and performance are reflected in data routinely collected by program offices. Regression analysis was used to create *system maturity* models for selected MDAPs between Milestone B (approval to start Engineering and Manufacturing Development) and declaration of Initial Operational Capability (IOC). These system maturity models were used to test the hypothesis that *system maturity is correlated to program schedule growth*.

Data Sources and Dataset Creation

The original research dataset[2] (Kamp, 2019) was created from publicly-released annual reports to Congress issued by the GAO, the Director, Operational Test and Evaluation (DOT&E), and the DoD for MDAPs between 2007 through 2017. Program data (observations) were included in the dataset if a program was assessed in both the GAO and DOT&E annual reports in a given report year.[3] The dataset was filtered to eliminate cancelled programs and programs without published schedule estimates or with missing data elements, and recorded in comma-separated variable files. This resulted in 154 observations of 48 programs from the 2007 through 2017 reports. Three observations[4] were outside the research program window from Milestone B IOC and were eliminated, leaving 151 observations in the database. No programs had observations in all years, but three programs had more than six observations.[5] Tests for observation independence were performed on these observations and on the complete dataset, and no additional observations were deleted.

Response and Predictor Variables

There are two explicit technology maturity response variables in the dataset, GAO technology maturity and estimated TRL.[6] The GAO's technology maturity assessment is an independent check of technology maturity. By definition, a *technology may be mature* when demonstrated within a system in a relevant (TRL 6) or operational (TRL 7) environment, but a *system is mature* when tested in its production version under operational conditions (TRL 8) or when used in an actual mission (TRL 9; Assistant Secretary of Defense for Research and Engineering [ASD(R&E)], 2011). The predictor variables used in the research are clustered into four groups: resource-, programmatic-, operational testing- and schedule-related predictors. The following tables will summarize the predictors by group and summarize their significance to technology maturity response variables.

1. <u>Resource-Related Predictors</u>: Resources are planned or budgeted quantities. All resource-related variables in the dataset were continuous and were derived from GAO annual reports or Selected Acquisition Report (SAR) Summary reports. These are summarized in Table 1.

---

[2] Available upon request.
[3] The GAO and DOT&E do not issue publicly available reports on all MDAPs each year, resulting in relatively few programs reported by both agencies. This criterion eliminated about 90% of MDAPs in any given year, but ensured concurrent programmatic and operational testing information.
[4] The three program observations outside the research window were C-130J in 2008, JLTV in 2011, and MQ-9 in 2014.
[5] The Joint Strike Fighter (F-35), CVN-78 and WIN-T programs all had more than six observations.
[6] An estimated TRL was created as TRLs are not typically reported in public documents.

Table 1. Resource-Related Predictors

Descriptive Name	Symbol	Explanation/ Description	Source	Maturity significance
Cost change assigned to Engineering - current year	Eng	PM reported cost changes allocated to Engineering in the GAO year, $ millions	SAR Summary	
Cost change assigned to Engineering - one year prior	Eng.1	PM reported cost changes allocated to Engineering ONE YEAR PRIOR to the GAO year, $ millions	SAR Summary	
Cost change assigned to Engineering - two years prior	Eng.2	PM reported cost changes allocated to Engineering TWO YEARS PRIOR to the GAO year, $ millions	SAR Summary	Sys maturity Binary, p=0.071
procurement funding	P.M	GAO or Program Office reported procurement funding, $ Millions (GAO value) (natural log of P.M is LN.P.M)	GAO	
research and development funding	RD.M	GAO or PM reported total research and development funding, $ Millions (GAO value) (natural log of RD.M is LN.RD.M)	GAO	for LN.RD.M, Binary, p=0.000
procurement quantities	P_no	Planned procurement quantities	GAO	
External Program cycle, months	Cycle.Mo	GAO or Program Office reported Program Office cycle time estimate, months	GAO	TRL Ordinal, p=0.000

2. *Schedule Predictors*: Schedule-related predictors may be interpreted as mimicking the program office view of progressing between events. These are continuous valued predictors calculated as differences between key milestone dates (program start, Milestone B, OT events, Milestone C, and IOC) within the GAO reporting. The schedule-related predictors are summarized in Table 2.

Table 2. Schedule-Related Predictors

Descriptive Name	Symbol	Explanation/ Description	Source	Maturity significance
Time from program start to Milestone B	B.st	time between MILESTONE B and Program start date, years	Calculated	TRL Ordinal, p=0.000
Time from Milestone B to Milestone C	C.B	time between MILESTONE B and MILESTONE C, years	Calculated	TRL Ordinal, p=0.010
Time from Milestone C to EVENT	C.ev	time between MILESTONE C and EVENT, years	Calculated	System maturity, Binary, p=0.000
Time from program start to EVENT	ev.st	time between EVENT and program start date, years	Calculated	TRL Ordinal, p=0.000
Relative Schedule Change	RSC	Relative Schedule Change (RSC) - *H2 DEPENDENT VARIABLE*	Calculated	TRL Ordinal, p=0.001

3. *Programmatic-Related Predictors:* Programmatic predictors reflect both programmatic strategies and external factors. There were 11 categorical (TRUE or FALSE) predictors derived from GAO annual reports and cross checked against DOT&E reports and any available SAR reports. These predictors are summarized in Table 3.

Table 3. Programmatic-Related Categorical Predictors

Descriptive Name	Symbol	Explanation/ Description	Source	Maturity significance
Commercial basis	COML	Program procures Commercially available system design/product (i.e. a helicopter or ship)	GAO	Both models Binary, p=0.031, Ordinal p=0.038
Complex system	complex_sys	System is complex	GAO	
External program dependencies	DEPEND	SYSTEM function depends on other programs not controlled by Program Office	GAO	Sys maturity Binary, p=0.079
Unstable funding	Fin_Uns	Indications that program is financially unstable (i.e. funding change > 10% in a year)	GAO	
Joint program	Joint	Joint Program indicator	GAO	Sys maturity Binary, p=0.008
Nunn-McCurdy Breach	NM	Nunn-McCurdy Breach occurred	GAO	
External program issues	PM.oth	other Program management issues (political direction or sponsorship)	GAO	
System prototypes	Prototype	Program uses prototypes representative of objective system, capable of operating in realistic environments	GAO	Sys maturity Binary, p=0.006
Unstable requirements	Req_Uns	Program requirements are unstable (i.e. >10% change in procurement; identified requirement changes)	GAO	Sys maturity Binary, p=0.006
Restructured program	Restr	Program is restructured	GAO	
Part of a System of Systems	SoS_Part	System is required to operate as part of a system-of-systems	GAO	Sys maturity Binary, p=0.079

The Complex system predictor is identified by the program office controlling subsystem and system selection and integration, while the Part of a System of Systems predictor identifies a system requiring other systems to accomplish its design mission (such as an aircraft carrier needing aircraft; Stuckey, Sarkani, & Mazzuchi, 2017). Unstable requirements are primarily indicated by year-to-year procurement quantity changes of more than 10%. A nominal variable identifies the system type (symbol "Type")—such as an aircraft, ship, missile, or ground vehicle system. These have been used by other researchers (Tate, 2016). Some useful predictor variables such as system mission, prototyping, program funding, technology maturity (Monaco & White III, 2005), and Drezner and Smith's (1990) programmatic structural and external factors may be found in the GAO Annual Assessments. Other predictors found in the literature, such as Low Rate Initial Production quantities and contract type (Monaco & White III, 2005), are not within the research data sources.

4. *Operational Testing (OT) Predictors:* The OT categorical (binary) predictors represent system issues found during OT events and are described in Table 4.

Table 4. Operational Testing-Related Predictors

Descriptive Name	Symbol	Explanation/ Description	Source	Maturity significance
System command and control issues	C3I	Testing issues with command, control, communications, intelligence (i.e. communications range and data rate) (0=FALSE, 1=TRUE)	DOTE	System maturity, Binary, p=0.002
System control and controllability issues	CONTROL	Testing issues with system controllability (i.e. precision, maneuverability) (0=FALSE, 1=TRUE)	DOTE	
System integration issues	INTEG	Testing system integration issues (i.e. fit, quality, non-compliance with requirements) (0=FALSE, 1=TRUE)	DOTE	Both models Binary, p=0.011, Ordinal p=0.005
Interoperability issues with other systems	INTEROP	Testing issues with system interoperability (i.e. can't exchange data with other systems, crypto) (0=FALSE, 1=TRUE)	DOTE	
System reliablity, maintainability, availability issues	RMA	Testing issues with system reliability, maintainability, availability (i.e. mean time between failures) (0=FALSE, 1=TRUE)	DOTE	
System operator usability issues	OPER	Testing issues with operator usability (i.e. safety, tactics, doctrine, procedures, training, cybersecurity) (0=FALSE, 1=TRUE)	DOTE	
System propulsion issues	PROP.PW.EN	Testing issues with propulsion power or energy (i.e. underpowered, prime mover issues) (0=FALSE, 1=TRUE)	DOTE	
System payloads issues	SEN.W	Testing issues with system payloads (i.e. sensors or weapons) (0=FALSE, 1=TRUE)	DOTE	TRL Ordinal, p=0.028
System structural issues	STRUCT	Testing physical structural issues (i.e. cracking, vibration, loading) (0=FALSE, 1=TRUE)	DOTE	
System size, weight, or power issues	SWAP	Testing issues with size, weight, or power (i.e. overweight) (0=FALSE, 1=TRUE)	DOTE	System maturity, Binary, p=0.046
System software performance issues	SW	Testing issues with system software (i.e. logic errors, production, vulnerabilities) (0=FALSE, 1=TRUE)	DOTE	System maturity, Binary, p=0.000

Azizian et al. (2011) identified critical technologies, analyses of alternatives, operational tests, certification and accreditation, and system engineering plans (all processes supporting technology readiness assessments) as important predictors affecting technology maturity and program performance. The OT predictors were iteratively derived using word frequency counting software DOT&E annual reports, and by reading the reports to derive context and relevance to system effectiveness and suitability.

5. *Testing Events and Predictors*: In general, common developmental and operational testing events or milestones were used. Flight Test was generalized to include first operational test (First Flight or first underway sea test) to represent operation in a realistic environment. Additionally, this research mapped DOT&E reported test event completion to an estimated TRL between 5 and 9 inclusive.[7] The TRL mappings to events were based upon DoD TRL definitions (Mankins, 2009) and were reviewed by independent experts. These events and their TRL mappings are summarized in Table 5.

[7] In lieu of using Nolte's (2008) TRL calculator to estimate TRLs

Table 5. Testing Events and TRL Mapping

Descriptive Name	Symbol	Explanation/ Description	Source	Maturity significance
Crtical Design Review	CDR	Critical Design Review (TRL=6 after completion)	GAO/DOTE	Event dependent
Design Review	DR	Design review (unspecified) (TRL at least 5, value dependent on description)	GAO/DOTE	Event dependent
Development testing	DT	Development testing (unspecified) (TRL test dependent)	GAO/DOTE	Event dependent
Early Fielding	EFR	Early Fielding Report – following directed deployment (TRL =9)	GAO/DOTE	Mature system
Force Deployment Evaluation	FDE	Force Deployment Evaluation	GAO/DOTE	Event dependent
Follow-on Test	FOTE	Follow On Test and Evaluation – testing following IOC (TRL=9)	GAO/DOTE	Mature system
Flight Test	FT	Flight Test (or first Sea Test) (unspecified) (TRL at least 6, value dependent on test description)	GAO/DOTE	Event dependent
Initial Operational Test	IOTE	Initial Operational Test and Evaluation – uses a production representative system (TRL>7, value dependent on test description)	GAO/DOTE	Event dependent
Land Based Test	LBT	Land Based Test (of any type) (TRL at least 5, value dependent on test description)	GAO/DOTE	Event dependent
Live Fire Test	LFTE	Live Fire Test and Evaluation – survivability testing of components or system (TRL >6, value dependent on test description)	GAO/DOTE	Event dependent
Limited User Test	LUT	Limited User Test – a subset of OT – for example a subset of effectiveness testing (TRL> 6, value dependent on test description)	GAO/DOTE	Event dependent
Milestone B	MSB	Milestone B (not a test event) (statutory TRL=6 after 2008)	GAO/DOTE	Event dependent
Milestone C	MSC	Milestone C (not a test event) (TRL=8)	GAO/DOTE	Mature system
Operational Assessment	OA	Operational Assessment – a subset of operational test (specific objective) (TRL at least 6, value dependent on test description)	GAO/DOTE	Event dependent
Operational Test	OT	Operational Test (unspecified) (TRL at least 7, value dependent on test description)	GAO/DOTE	Event dependent
Quick Reaction Assessment	QRA	Quick Reaction Assessment – for a specific end use (TRL =8)	GAO/DOTE	Mature system

Methodology

We performed binary and ordinal logistic regression analyses on the dataset using Minitab 18 and SPSS. The response variables are GAO technology maturity for the binary logistic regression and estimated TRL for the ordinal logistic regression. The regressions were reduced using backwards elimination or manually (one variable at a time) until only significant predictors remained. A random 10% subset of the dataset was withheld for model validation. Residuals were inspected to assess if regression assumptions were satisfied; then model goodness of fit and accuracy were evaluated.

Results

GAO Technology Maturity Regression Model

Table 6 and Figure 1 summarize the GAO technology maturity binary logistic regression model. Significant terms are identified using the predictor symbols from Tables 1 through 5. The model is significant at the α = 0.05 level.

Table 6. Results of GAO Technology Maturity Binary Logistic Regression

Term	Coef	Contribution	P-Value	VIF	Odds Ratio
Regression		39.85%			
Constant	7.500		0.000		
LN.RD.M	-0.944	7.89%	0.000	1.60	0.389
C.ev	-0.381	8.12%	0.000	1.87	0.683
[Req_Uns=1]	-1.461	1.46%	0.009	1.57	0.232
[COML=1]	-1.303	1.29%	0.024	1.50	0.272
[Prototype=1]	1.404	2.46%	0.015	1.67	4.070
[SW=1]	2.536	7.11%	0.000	1.56	12.632
[C3I=1]	1.404	0.99%	0.010	1.29	4.071
[INTEG=1]	-1.217	3.15%	0.013	1.21	0.296
[DEPEND=1]	-1.271	3.71%	0.014	1.26	0.281
[Joint=1]	-1.814	3.67%	0.009	1.38	0.163

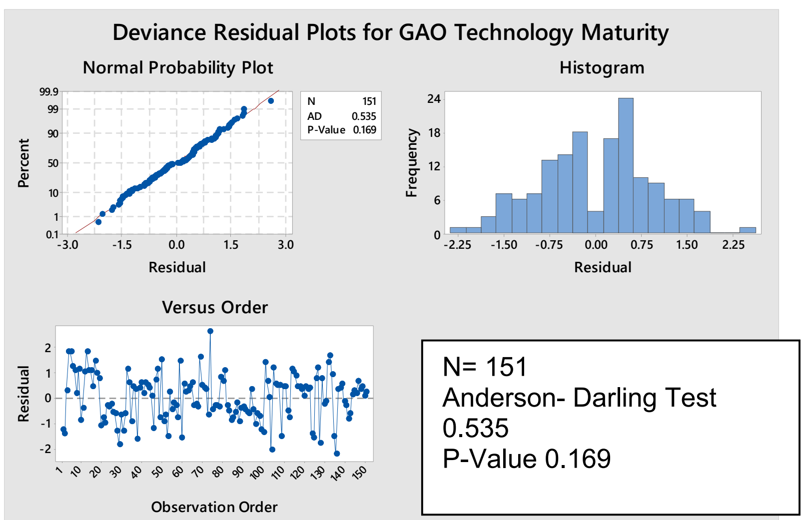


Figure 1. Technology Maturity Regression Deviance Residual Plots

The model satisfies all binary logistic regression assumptions. Model accuracy was assessed by withholding 14 random observations and re-performing the regression on the reduced dataset. This *accuracy test* model was used to predict the GAO technology maturity of these 14 withheld observations, and predictions and observations were compared. The regression predictors changed slightly between the two models. The accuracy test model correctly predicted GAO Technology Maturity 11 of 14 times (78.6%). Table 7 summarizes the goodness-of-fit differences between these two models.

Table 7. Binary Logistic Regression Goodness-of-Fit and Association Measures

	Goodness-of-Fit Tests					Model Summary			Measures of Association
						Deviance	Deviance		
Hosmer-Lemeshow	Observations	Model α	DF	Chi-Square	P-Value	R-Sq	R-Sq(adj)	AIC	Kendall's Tau-a
Table 6 model	151	0.05	8	6.85	0.552	39.85%	35.06%	147.42	0.39
Accuracy test model	137	0.05	8	5.4	0.714	40.25%	34.98%	135.26	0.40

Association is between the response variable and predicted probabilities

Summary of GAO Technology Maturity Regression Results

The model correctly predicted GAO technology maturity over 75% of the time. Most of the 10 predictors contributed less than 4% of the variance. The following were the top three predictors in contribution order: time between Milestone C and defined events, the natural log of research and development funds, and software issues during Events. The model was most affected by programmatic (five predictors), then resource, operational testing (three predictors), and finally schedule predictors. Issues found during Testing Events indicate system immaturity, or a system without issues during an event is considered mature for that predictor. In particular, software issues discovered during operational testing had the largest odds ratio, as such issues were commonly discovered later in system development and testing relative to other issues.

TRL Ordinal Logistic Regression Model

An ordinal logistic regression of estimated technology readiness levels (TRLe) was performed in Minitab. Predictors were removed until only those with p-values less than or equal to 0.05 remained, and the regression is significant at α = 0.05. The model was re-run in SPSS using the Minitab predictors to test the proportional odds assumption using the test of parallel lines. The Minitab and SPSS logistic regression results are summarized in Table 8.

Table 8. Ordinal Logistic Regression of Technology Maturity (TRLs)

TRL Ordinal Logistic Regression Table

Minitab results					Odds	95% CI		SPSS results						95% CI	
Predictor	Coef	SE Coef	Z	P	Ratio	Lower	Upper		Estimate	Std. Error	Wald	df	Sig.	Lower	Upper
[TRL=5]	-6.609	0.955	-6.92	0.000				[TRLe = 5]	-7.219	1.002	51.883	1	0.000	-9.184	-5.255
[TRL=6]	-2.499	0.675	-3.70	0.000				[TRLe = 6]	-3.109	0.728	18.246	1	0.000	-4.536	-1.683
[TRL=7]	-1.341	0.656	-2.04	0.041				[TRLe = 7]	-1.951	0.701	7.745	1	0.005	-3.325	-0.577
[TRL=8]	1.475	0.662	2.23	0.026				[TRLe = 8]	0.864	0.692	1.559	1	0.212	-0.492	2.220
Cycle.Mo	0.030	0.007	4.47	0.000	1.03	1.02	1.04	Cycle.Mo	-0.030	0.007	19.950	1	0.000	-0.043	-0.017
ev.st	-0.561	0.075	-7.50	0.000	0.57	0.49	0.66	ev.st	0.561	0.075	56.224	1	0.000	0.414	0.708
B.st	0.523	0.131	4.00	0.000	1.69	1.31	2.18	B.st	-0.523	0.131	15.999	1	0.000	-0.779	-0.267
C.B	0.180	0.070	2.59	0.010	1.20	1.04	1.37	C.B	-0.180	0.070	6.683	1	0.010	-0.316	-0.044
RSC	2.194	0.673	3.26	0.001	8.97	2.40	33.55	RSC	-2.194	0.673	10.631	1	0.001	-3.513	-0.875
[COML=1]	-0.819	0.396	-2.07	0.038	0.44	0.20	0.96	[COML=0]	-0.819	0.396	4.286	1	0.038	-1.594	-0.044
[SEN.W=1]	-0.809	0.368	-2.20	0.028	0.45	0.22	0.92	[SEN.W=0]	-0.809	0.368	4.846	1	0.028	-1.530	-0.089
[INTEG=1]	1.018	0.363	2.80	0.005	2.77	1.36	5.63	[INTEG=0]	1.018	0.363	7.863	1	0.005	0.306	1.729

The model differences are due to software implementation differences. All ordinal logistic regression model assumptions were satisfied. The pseudo-R2 is 0.577. SPSS provides a simpler prediction performance summary, and full dataset results are shown in Table 9.

Table 9. SPSS Ordinal Logistic Regression TRL Prediction Results

SPSS model performance		Actual TRL estimate 5 Count	6 Count	7 Count	8 Count	9 Count
Predicted Response Category	5	0	0	0	0	0
	6	4	31	6	13	0
	7	0	0	0	0	0
	8	0	10	18	38	11
	9	0	0	0	4	16

		Actual TRL estimate [5,6,7]	[8,9]		
Predicted TRL estimate	[8,9]	28	69	exact	56.3%
	[5,6,7]	41	13	+/- 1	84.8%
		[5,6,7]	[8,9]	correct	72.8%
	[8,9]	18.5%	45.7%		
	[5,6,7]	27.2%	8.6%		

The predicted responses represent the highest probability for each observation; the actual TRL is the estimated TRL in the data set. The model predicts the exact TRL (matches the estimated TRL) 56.3% of the time and is within + 1 TRL level nearly 85% of the time. The response gap seen in Table 9 at a predicted TRL of 7 mimics the TRL and system maturity relationship (system mature if TRL is 8 or 9, system immature if TRL is 5, 6, or 7).

Summary of TRL Ordinal Regression Results

The TRL ordinal logistic regression model is dominated by schedule-related predictors. It predicted an *exact* TRL slightly over half the time; this is not useful if a program needs high confidence in the prediction precision. The SPSS model correctly predicted *system maturity* over 70% of the time and shows a significant relationship between TRL (maturity) and relative schedule change. *Program offices may use the inability to hold to schedule as an indicator of technology immaturity, and during Engineering and Manufacturing Development (EMD), as an indicator of system immaturity.*

System Maturity and Schedule Results

Finally, the dataset was partitioned into immature (estimated TRL = 5, 6, 7) and mature (estimated TRL = 8, 9) subsets. A linear regression of Relative Schedule Change (RSC) using the research predictors was performed on each partition. The immature model (estimated TRL = 5, 6, 7) did not satisfy residual normal distribution assumptions at α = 0.05, so comparisons are made when both models are significant at α = 0.1. These models are presented in Table 10 and Figure 2.

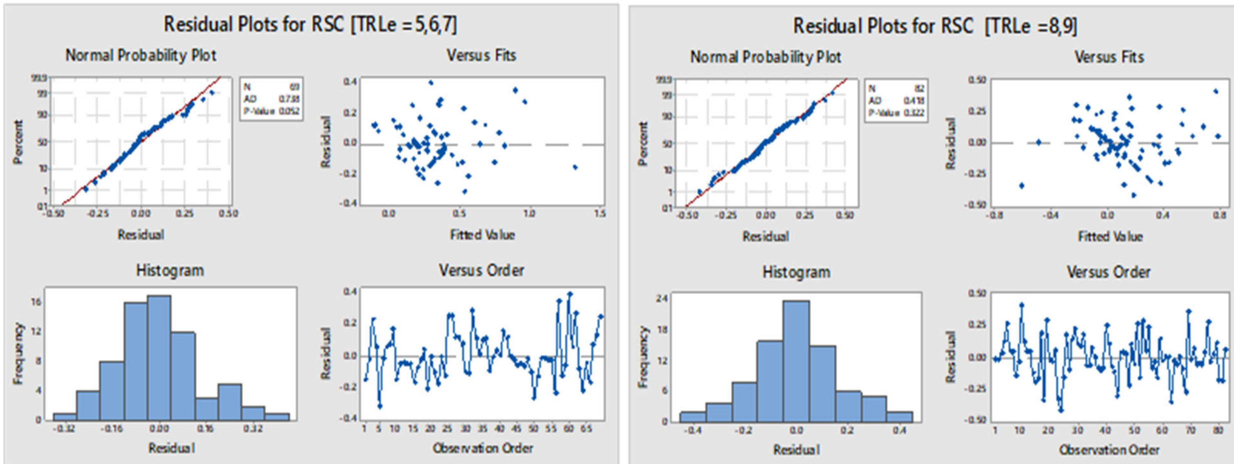
Table 10. RSC Regression Summaries by System Maturity

Systems with TRLe = [5,6,7]

Source	Coef	Contribution	p	VIF
Regression	0.552	74.27%	0	
LN.RD.M	0.0635	0.25%	0.009	2.82
LN.P.M	-0.0807	16.51%	0	2.02
Cycle.Mo	-0.004734	27.51%	0	2.74
B.st	0.0633	9.63%	0	2.81
C.B	0.0207	0.48%	0.009	1.5
Restr	0.152	2.14%	0.002	1.49
PM.oth	-0.1146	1.49%	0.012	1.32
complex_sys	0.3115	0.31%	0	2.68
SoS_part	0.3782	14.22%	0	2.85
DEPEND	-0.1301	0.80%	0.014	1.58
OPER	-0.1003	0.92%	0.058	1.73
S	R-sq	R-sq(adj)	PRESS	R-sq(pred)
0.160606	74.27%	69.30%	2.3534	58.81%

Systems with TRLe = [8,9]

Source	Coef	Contribution	p	VIF
Regression	0.693	68.58%	0	
LN.P.M	-0.1058	10.59%	0	1.65
Eng.2	-0.000414	0.23%	0.008	3.26
P_no	-0.000001	3.43%	0.013	1.87
C.B	0.0278	0.01%	0.013	2.92
Restr	-0.3253	2.78%	0	3.05
PM.oth	0.3648	1.64%	0	2.03
COML	0.1854	1.75%	0.006	2.44
Prototype	-0.278	2.51%	0	2.23
NM	0.1334	0.54%	0.039	1.86
Type	* by factor	37.30%	0	* by factor
complex_sys	0.2167	1.73%	0.013	2.00
SoS_part	0.1271	1.68%	0.072	1.86
SWAP	0.2466	0.10%	0.003	*
RMA	0.1334	4.28%	0	1.96
S	R-sq	R-sq(adj)	PRESS	R-sq(pred)
0.188684	68.58%	59.60%	3.852	46.03%



Estimated TRL = [5,6,7]
N= 69
Anderson- Darling Test
0.738
P-Value 0.052

Estimated TRL = [8,9]
N= 82
Anderson- Darling Test
0.418
P-Value 0.322

Figure 2. Residual Plots for System Maturity-Related Schedule Regressions

Clustering the linear regression predictors into four broad groups—Resources, Programmatic, OT and Schedule—provides an efficient representation of the changing relative importance of factor groups as a system proceeds from immature to mature per Table 11.

Table 11. Schedule Variance Percentages by Factor Groups

	[TRLe=5,6,7]	[TRLe = 8,9]
Resources	44.27%	14.26%
Program	18.16%	49.93%
OT	1.72%	4.38%
Schedule	10.11%	0.01%
error	25.74%	31.42%

These schedule growth models showed that immature systems need adequate resources (including both time and money as resources), a good initial schedule plan (getting the schedule right), and a plan to manage system complexity driven by interactions with the larger system of systems. Mature system schedule growth is driven by commodity type (for example, aircraft or ship) integration issues.

Conclusions

This research examined different measures of technology and system maturity and identified maturity-related factors. We used logistic regression analysis to show relationships between system maturity and program schedule growth. This research is valid for MDAPs with reports issued by both the GAO and DOT&E in the same year from 2007 through 2017. Research findings may not be valid for MDAPs not in these reports, highly classified programs, defense business systems and smaller expenditure acquisition programs. The research provides program offices insight into technology and system maturity and the sources of schedule growth based on resource, programmatic, operational testing, and schedule-related factors, allowing them to monitor and adjust acquisition program planning and execution. Examples of useful results for program managers include the following:

- This research developed operational testing performance factors. These factors were shown to be related to system maturity and program schedule growth. Program managers may use such factors to help develop quantitative measures of system maturity related to factors seen during development and testing.
- Program managers may use the combination of a reported or estimated GAO technology maturity and a predicted or estimated TRL of 8 or 9 as an indicator of *system* maturity during system Engineering and Manufacturing Development.
- The research showed that resources (having enough money and time to develop a system) matter most before a system is mature and that program structure and execution matter later in program execution. However, program managers need both program resources and structure from the start to deliver and support their products.
- The research showed that an *inability* to adhere to planned schedule indicates system immaturity.

References

Assistant Secretary of Defense for Research and Engineering (ASD[R&E]). (2011, April). *Technology readiness assessment (TRA) guidance*. Washington, DC: DoD. Retrieved from https://www.acq.osd.mil/ecp/DOCS/DoDGuidance/TRA2011.pdf

Azizian, N., Mazzuchi, T., Sarkani, S., & Rico, D. F. (2011). A framework for evaluating technology readiness, system quality, and program performance of U.S. DoD acquisitions. *Systems Engineering*, *14*(4), 410–426.

Bailey, R. U., Mazzuchi, T. A., Sarkani, S., & Rico, D. F. (2014). A comparative analysis of the value of technology readiness assessments. *Defense Acquisition Research Journal*, *21*(4), 825–850.

Conrow, E. H. (2011). Estimating technology readiness level coefficients. *Journal of Spacecraft and Rockets*, *48*(1).

Day, G. (1981). The product life cycle: Analysis and applications issues. *Journal of Marketing*, *45*(4), 60. https://doi.org/10.2307/1251472

Dodaro, G. L. (2007). *Defense acquisitions: Assessments of selected weapon programs* (GAO-07-406SP; p. 178). Washington, DC: GAO.

Dodaro, G. L. (2013). *Defense acquisitions: Assessments of selected weapon programs* (GAO-13-294SP; p. 190). Washington, DC: GAO.

Drezner, J. A., & Smith, G. K. (1990). *An analysis of weapon system acquisition schedules* (Acquisition No. R3937; p. 212). Santa Monica, CA: RAND.

Gove, R., & Uzdzinski, J. (2013). A performance-based system maturity assessment framework. *2013 Conference on Systems Engineering Research*, *16*, 688–697. https://doi.org/10.1016/j.procs.2013.01.072

Kamp, J. (2019). *Integrating immature systems and program schedule growth*. Washington, DC: The George Washington University.

Kandt, A., Pickshaus, T., Fleischer, K., & Schmitt, R. (2016). A new model to ascertain product maturity in product development processes. *26th CIRP Design Conference*, *50*, 173–178. https://doi.org/10.1016/j.procir.2016.05.006

Katz, D. R., Sarkani, S., Mazzuchi, T., & Conrow, E. H. (2015). The relationship of technology and design maturity to DoD weapon system cost change and schedule change during engineering and manufacturing development: Relationship of maturity to DoD weapon system effectiveness. *Systems Engineering*, *18*(1), 1–15. https://doi.org/10.1111/sys.21281

Major Defense Acquisition Program Defined, 10 U.S.C. § 2430 (2007). Retrieved from https://www.gpo.gov/fdsys/pkg/USCODE-2006-title10/html/USCODE-2006-title10-subtitleA-partIV-chap144-sec2430.htm

Mankins, J. C. (2009). Technology readiness assessments: A retrospective. *Acta Astronautica*, *65*(9), 1216–1223. https://doi.org/10.1016/j.actaastro.2009.03.058

Mattis, J. (2018). *Summary of the 2018 national defense strategy: Sharpening the American military's competitive edge*. Washington, DC: DoD. Retrieved from https://dod.defense.gov/Portals/1/Documents/pubs/2018-National-Defense-Strategy-Summary.pdf

Monaco, J. V., & White III, E. D. (2005). Investigating schedule slippage. Defense Technical Information Center.

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

Ramirez-Marquez, J. E., & Sauser, B. J. (2009). System development planning via system maturity optimization. *IEEE Transactions on Engineering Management*, *56*(3), 533–548. https://doi.org/10.1109/TEM.2009.2013830

Stuckey, R. M., Sarkani, S., & Mazzuchi, T. A. (2017). Complex acquisition requirements analysis: Using a systems engineering approach. *Defense Acquisition Research Journal: A Publication of the Defense Acquisition University*, *24*(2), 266–301.

Tate, D. M. (2016). *Acquisition cycle time: Defining the problem* (Revised).

Tetlay, A., & John, P. (2009). Determining the lines of system maturity, system readiness and capability. In *Readiness in the system development lifecycle*. England: Loughborough University. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.614.9724

Wang, Q., Wang, Z., & Zhao, X. (2015). Strategic orientations and mass customisation capability: The moderating effect of product life cycle. *International Journal of Production Research*, *53*(17), 1–18. https://doi.org/10.1080/00207543.2015.1027012

Weapon Systems Acquisition Reform Act of 2009, Pub. L. No. 111-23 (2009). Retrieved from https://www.gpo.gov/fdsys/pkg/PLAW-111publ23/content-detail.html

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

www.acquisitionresearch.net