Agile Software Development Cost Modeling for the US DoD

Wilson Rosa, Naval Center for Cost Analysis

Ray Madachy, Naval Postgraduate School

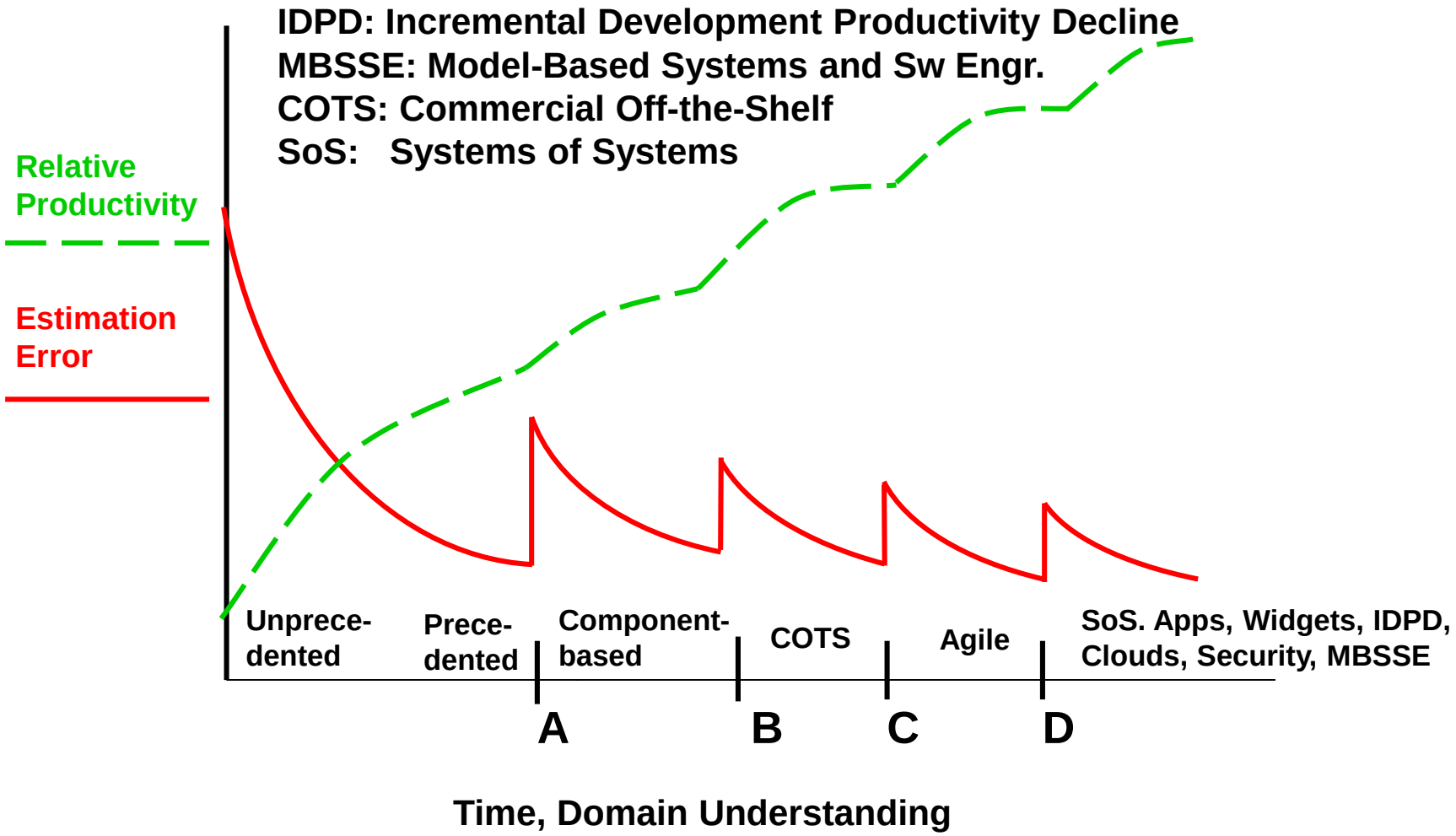
Bradford Clark, Software Metrics, Inc.

Barry Boehm, University of Southern California

SEI Software and Cyber Solutions Symposium

March 27, 2018

A Short History of Software Estimation Accuracy



Problem Statement

- In DoD, Popular Size Measures are often not available for Agile Effort Estimation at early phase
  - Function Points (FP)
  - COSMIC FP
  - Story Points
  - Source Lines of Code

- No Publicized/Empirical Agile Effort Estimation Models

Purpose

- Publish Agile Effort Estimation Models for
  - Crosschecking Contractor Cost Proposals
  - Validating Independent Government Cost Estimates
- Examine the validity of using Initial Software Requirements as proxy size measure
- Develop useful cost models using early phase information
- Model calibration comparison:

	Size		Cost Factors		Effort
Model Type	Initial Estimate	Final Actual	Initial Estimates	Final Actuals	Final Actual
Early Phase	x		x		x
Traditional		x		x	x

Outline

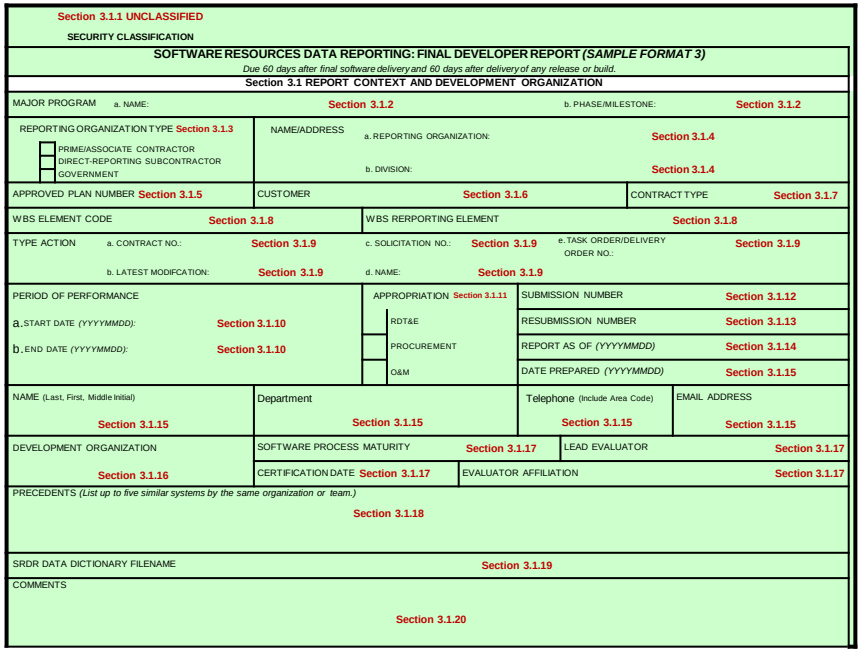
- Experimental Design
- Dataset Demographics
- Productivity Benchmarks
- Agile Effort Estimation Models
- Conclusion

Experimental Design

Primary Data Collection Form

- 2011 Software Resource Data Report (SRDR) (DD Form 2630)

SRDR Final Developer Report



Section 3.1.1 UNCLASSIFIED

SECURITY CLASSIFICATION

SOFTWARE RESOURCES DATA REPORTING: FINAL DEVELOPER REPORT *(SAMPLE FORMAT 3)*

Due 60 days after final software delivery and 60 days after delivery of any release or build.

Section 3.1 REPORT CONTEXT AND DEVELOPMENT ORGANIZATION

MAJOR PROGRAM a. NAME: Section 3.1.2 b. PHASE/MILESTONE: Section 3.1.2

REPORTING ORGANIZATION TYPE Section 3.1.3
- PRIME/ASSOCIATE CONTRACTOR
- DIRECT-REPORTING SUBCONTRACTOR
- GOVERNMENT

NAME/ADDRESS a. REPORTING ORGANIZATION: Section 3.1.4 b. DIVISION: Section 3.1.4

APPROVED PLAN NUMBER Section 3.1.5 | CUSTOMER Section 3.1.6 | CONTRACT TYPE Section 3.1.7

WBS ELEMENT CODE Section 3.1.8 | WBS REPORTING ELEMENT Section 3.1.8

TYPE ACTION a. CONTRACT NO.: Section 3.1.9 c. SOLICITATION NO.: Section 3.1.9 e. TASK ORDER/DELIVERY ORDER NO.: Section 3.1.9
b. LATEST MODIFCATION: Section 3.1.9 d. NAME: Section 3.1.9

PERIOD OF PERFORMANCE
a. START DATE *(YYYYMMDD)*: Section 3.1.10
b. END DATE *(YYYYMMDD)*: Section 3.1.10

APPROPRIATION Section 3.1.11
- RDT&E
- PROCUREMENT
- O&M

SUBMISSION NUMBER Section 3.1.12
RESUBMISSION NUMBER Section 3.1.13
REPORT AS OF *(YYYYMMDD)* Section 3.1.14
DATE PREPARED *(YYYYMMDD)* Section 3.1.15

NAME (Last, First, Middle Initial) Section 3.1.15 | Department Section 3.1.15 | Telephone (Include Area Code) Section 3.1.15 | EMAIL ADDRESS Section 3.1.15

DEVELOPMENT ORGANIZATION Section 3.1.16 | SOFTWARE PROCESS MATURITY Section 3.1.17 | LEAD EVALUATOR Section 3.1.17
CERTIFICATION DATE Section 3.1.17 | EVALUATOR AFFILIATION Section 3.1.17

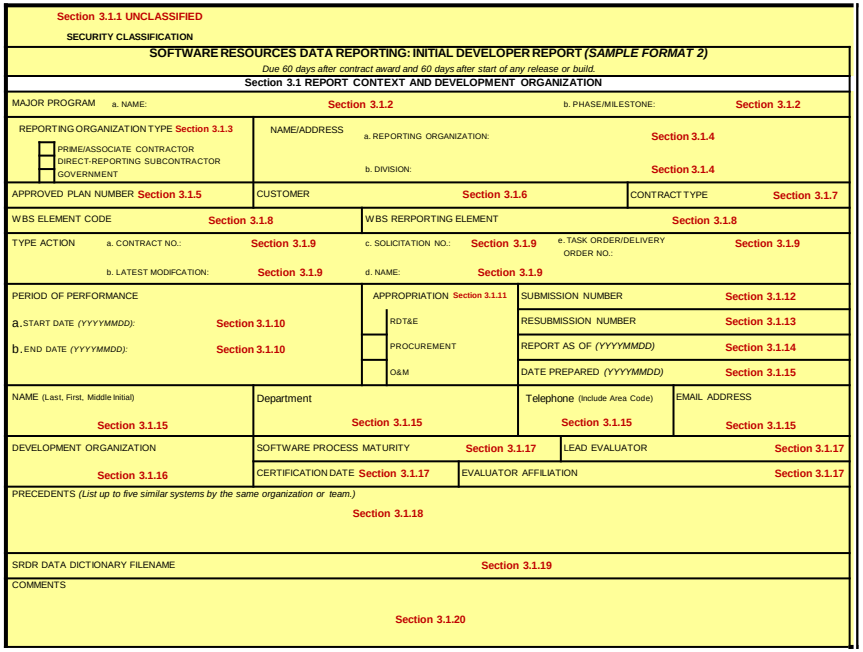
PRECEDENTS *(List up to five similar systems by the same organization or team.)* Section 3.1.18

SRDR DATA DICTIONARY FILENAME Section 3.1.19

COMMENTS Section 3.1.20

Actual Development Effort
Actual Development Process

SRDR Initial Developer Report



Section 3.1.1 UNCLASSIFIED

SECURITY CLASSIFICATION

SOFTWARE RESOURCES DATA REPORTING: INITIAL DEVELOPER REPORT *(SAMPLE FORMAT 2)*

Due 60 days after contract award and 60 days after start of any release or build.

Section 3.1 REPORT CONTEXT AND DEVELOPMENT ORGANIZATION

MAJOR PROGRAM a. NAME: Section 3.1.2 b. PHASE/MILESTONE: Section 3.1.2

REPORTING ORGANIZATION TYPE Section 3.1.3
- PRIME/ASSOCIATE CONTRACTOR
- DIRECT-REPORTING SUBCONTRACTOR
- GOVERNMENT

NAME/ADDRESS a. REPORTING ORGANIZATION: Section 3.1.4 b. DIVISION: Section 3.1.4

APPROVED PLAN NUMBER Section 3.1.5 | CUSTOMER Section 3.1.6 | CONTRACT TYPE Section 3.1.7

WBS ELEMENT CODE Section 3.1.8 | WBS REPORTING ELEMENT Section 3.1.8

TYPE ACTION a. CONTRACT NO.: Section 3.1.9 c. SOLICITATION NO.: Section 3.1.9 e. TASK ORDER/DELIVERY ORDER NO.: Section 3.1.9
b. LATEST MODIFCATION: Section 3.1.9 d. NAME: Section 3.1.9

PERIOD OF PERFORMANCE
a. START DATE *(YYYYMMDD)*: Section 3.1.10
b. END DATE *(YYYYMMDD)*: Section 3.1.10

APPROPRIATION Section 3.1.11
- RDT&E
- PROCUREMENT
- O&M

SUBMISSION NUMBER Section 3.1.12
RESUBMISSION NUMBER Section 3.1.13
REPORT AS OF *(YYYYMMDD)* Section 3.1.14
DATE PREPARED *(YYYYMMDD)* Section 3.1.15

NAME (Last, First, Middle Initial) Section 3.1.15 | Department Section 3.1.15 | Telephone (Include Area Code) Section 3.1.15 | EMAIL ADDRESS Section 3.1.15

DEVELOPMENT ORGANIZATION Section 3.1.16 | SOFTWARE PROCESS MATURITY Section 3.1.17 | LEAD EVALUATOR Section 3.1.17
CERTIFICATION DATE Section 3.1.17 | EVALUATOR AFFILIATION Section 3.1.17

PRECEDENTS *(List up to five similar systems by the same organization or team.)* Section 3.1.18

SRDR DATA DICTIONARY FILENAME Section 3.1.19

COMMENTS Section 3.1.20

Estimated Functional Requirements
Estimated External Interfaces
Estimated Peak Staff
Application Domain

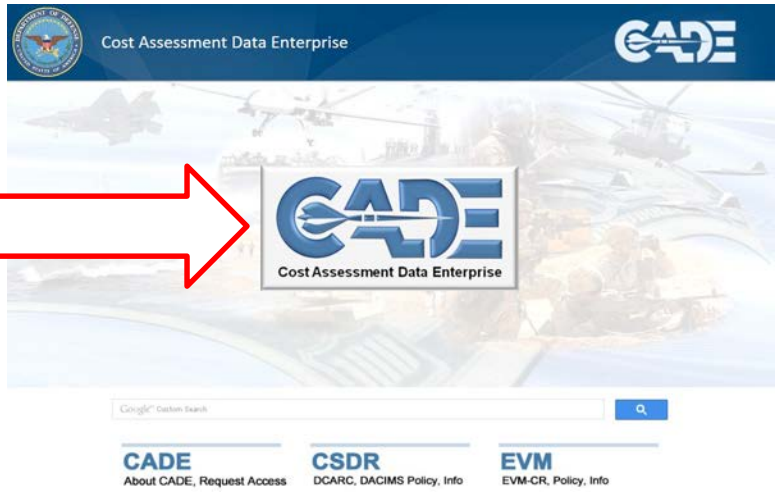
Population and Sample Size

Empirical data from 20 recent US DoD Agile programs:

12 Paired SRDRs from the Cost Assessment Data Enterprise (CADE)

Each paired SRDR includes:

SRDR Initial Developer Report (Estimates)
&
SRDR Final Developer Report (Actuals)



http://dcarc.cape.osd.mil/Default.aspx

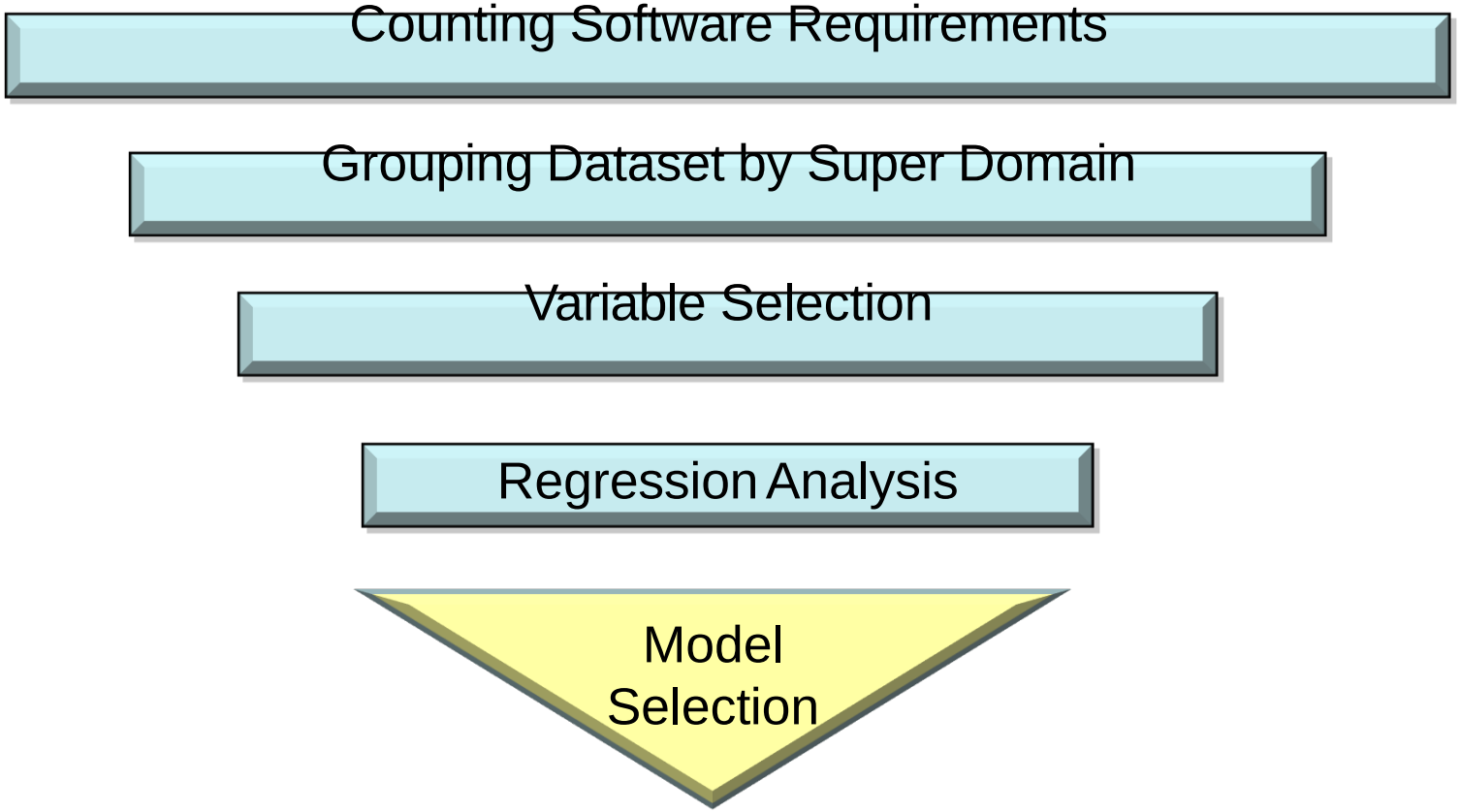
\+ **4** additional SRDRs from CADE (SRDR Final only)

\+ **4** Agile projects from proprietary source

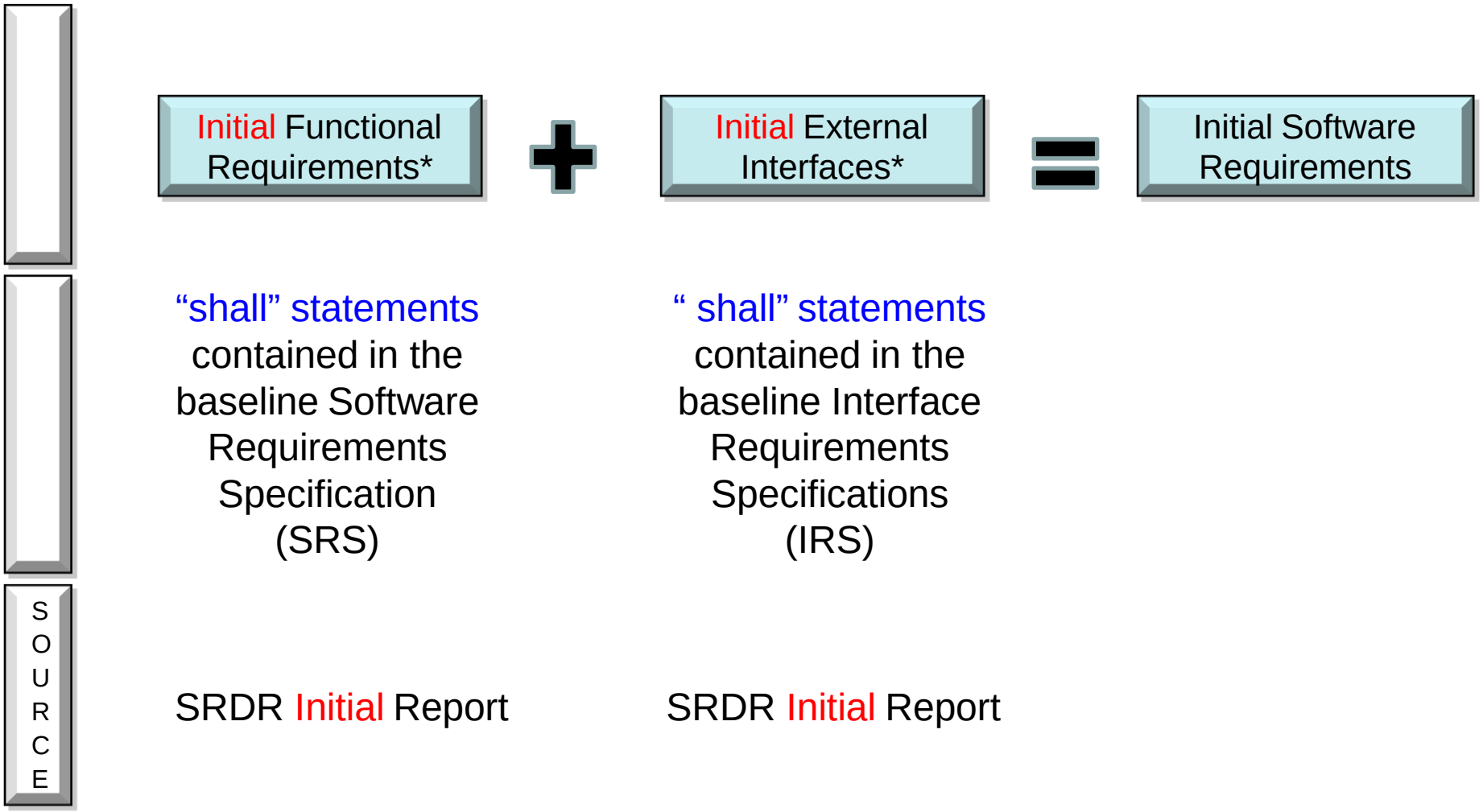
= **20** Agile projects analyzed in this study

Data Normalization and Analysis Workflow

- Dataset normalized to “account for sizing units, application complexity, and content so they are consistent for comparisons” (source: GAO)



Counting Software Requirements

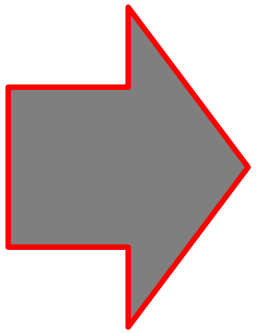


*Typically available before contract award
*Definitions align with IEEE std. 830-1998

Grouping Dataset by Super Domain

1) Dataset initially mapped into 17 Application Domains*
2) Then into 4 complexity groups called Super Domains

Application Domain
Software Tools
Training
Enterprise Information System
Enterprise Services
Custom AIS Software
Mission Planning
Test, Measurement, and Diagnostic Equipment
Scientific & Simulation
Process Control
System Software
Command & Control, Communications
Real Time Embedded
Vehicle Control/Payload
Signal Processing, Microcode & Firmware



Super Domain
Mission Support (SUPP)
Automated Information System (AIS)
Engineering (ENG)
Real Time (RTE)

*New DOD policy (http://cade.osd.mil/policy/srdr) requires that Application Domains are identified for reported software activities.

Grouping Dataset by Super Domain

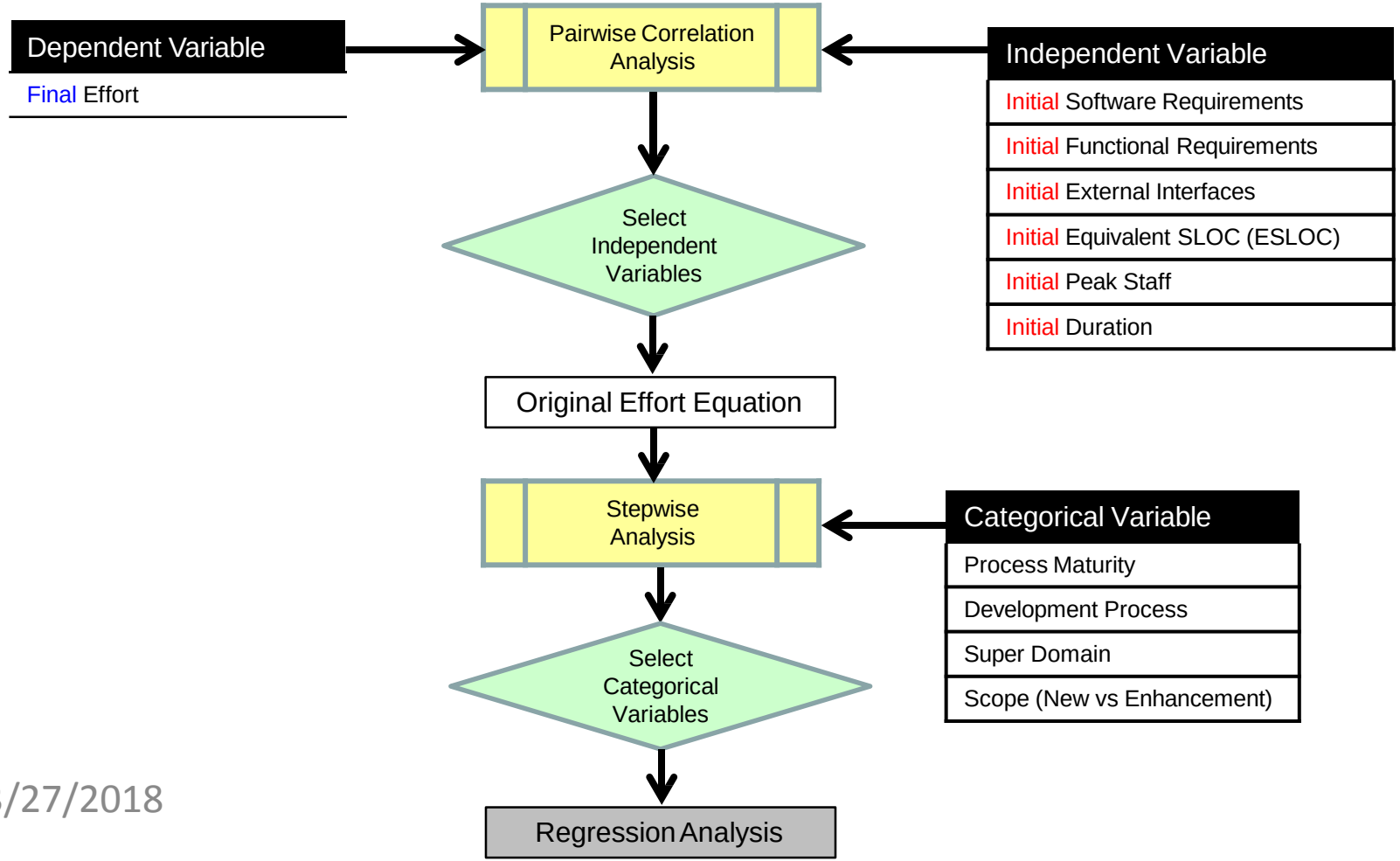
Super Domains

	Support	AIS	Engineering	Real Time	TOTAL
Aircraft	2	0	4	0	6
Business	1	3	0	0	4
C4I	0	1	3	5	9
Missile	0	0	0	1	1
	3	4	7	6	20

Top 2 Operating Environments à C4I and Aircraft

Variable Selection

1) Pairwise Correlation to select Independent Variables
2) Stepwise Analysis to select Categorical Variables



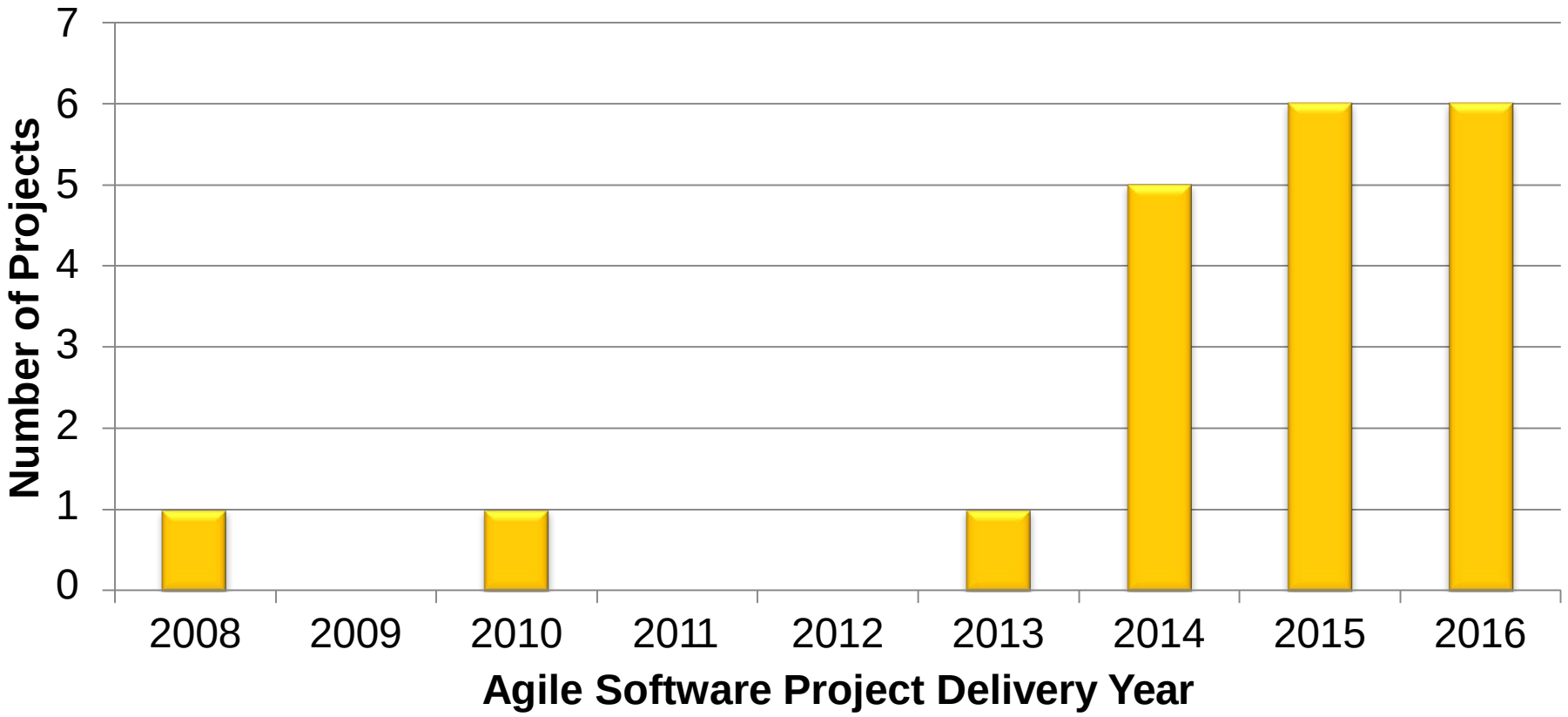
Model Selection

§ Model Selection Based on P-Value, lowest MMRE and CV

Measure	Symbol	Description
Coefficient of Variation	CV	Percentage expression of the standard error compared to the mean of dependent variable. A relative measure allowing direct comparison among models.
P-value	α	Level of statistical significance established through the coefficient alpha ($p \leq \alpha$).
Variance Inflation Factor	VIF	Indicates whether multi-collinearity (correlation among predictors) is present in multiple regression analysis.
Coefficient of Determination	R^2	The Coefficient of Determination shows how much variation in dependent variable is explained by the regression equation.
Mean Magnitude of Relative Error	MMRE	Low MMRE is an indication of high accuracy. MMRE is defined as the sample mean (M) of the magnitude relative error (MME). MME is the absolute value of the difference between Actual and Estimated effort divided by the Actual effort, $(A - E) / A$

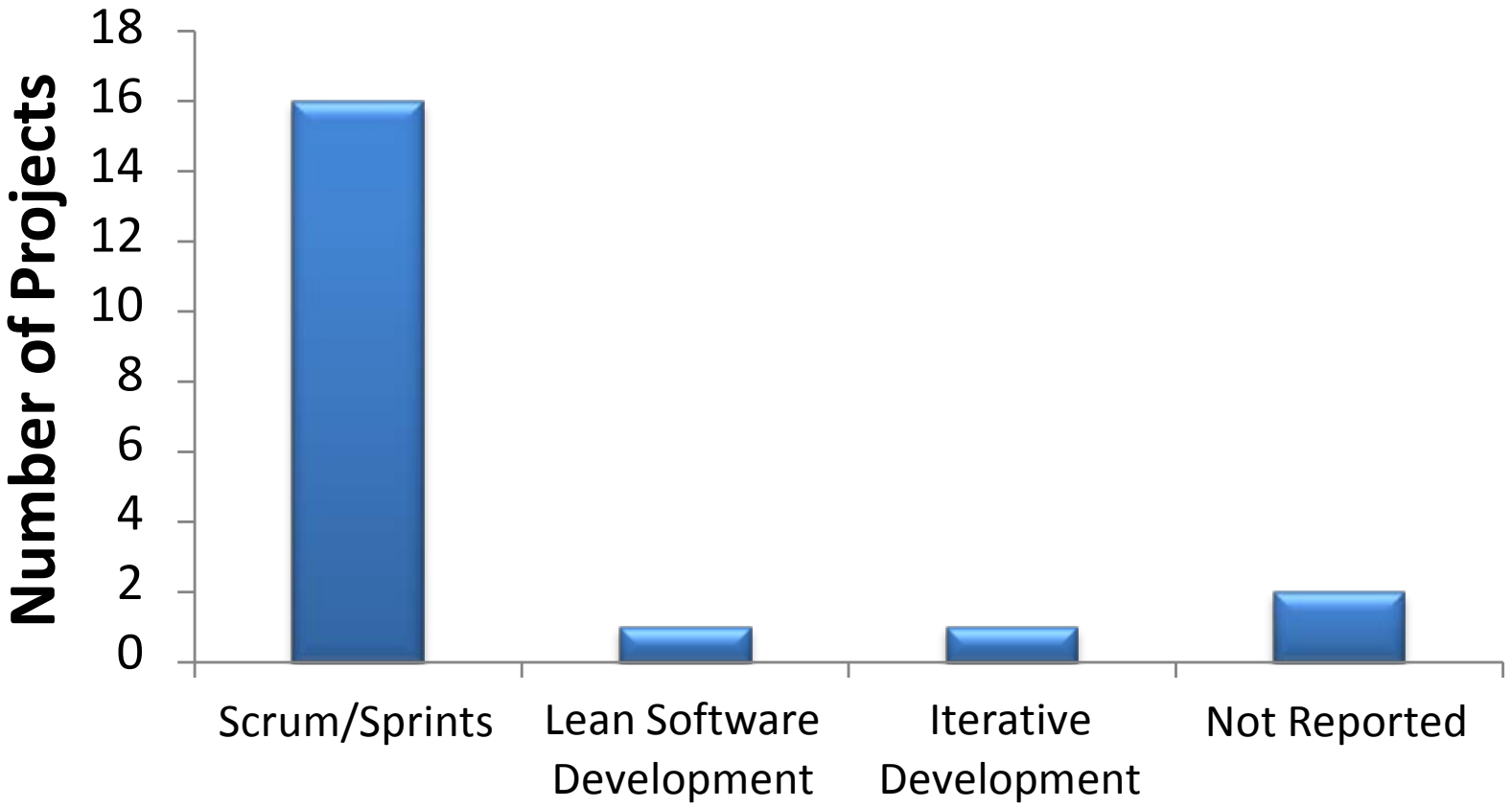
Dataset Demographics

Dataset by Delivery Year



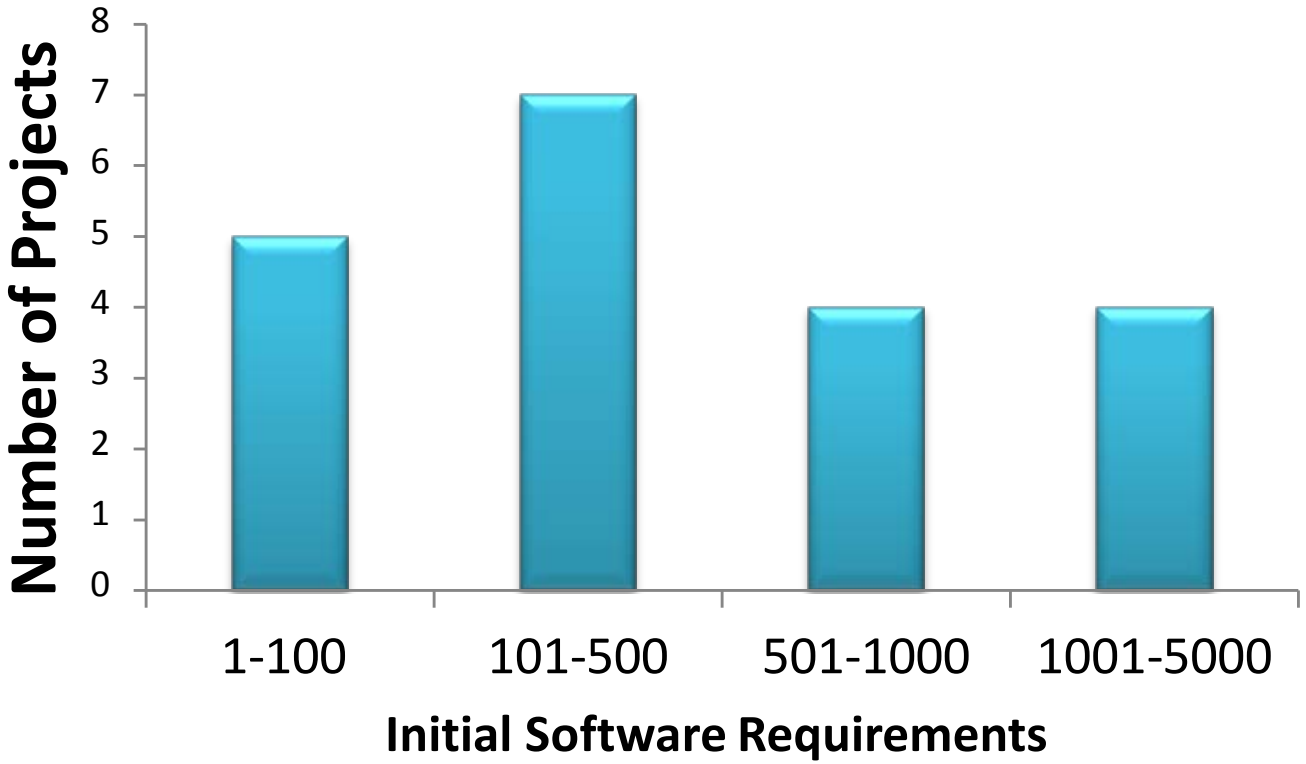
# of completed Agile Projects (reported in CADE) have increased since 2014

Dataset by Agile Framework



SRDR submissions provided limited information about Agile Framework
Future SRDR submissions will require developers to describe their Agile process

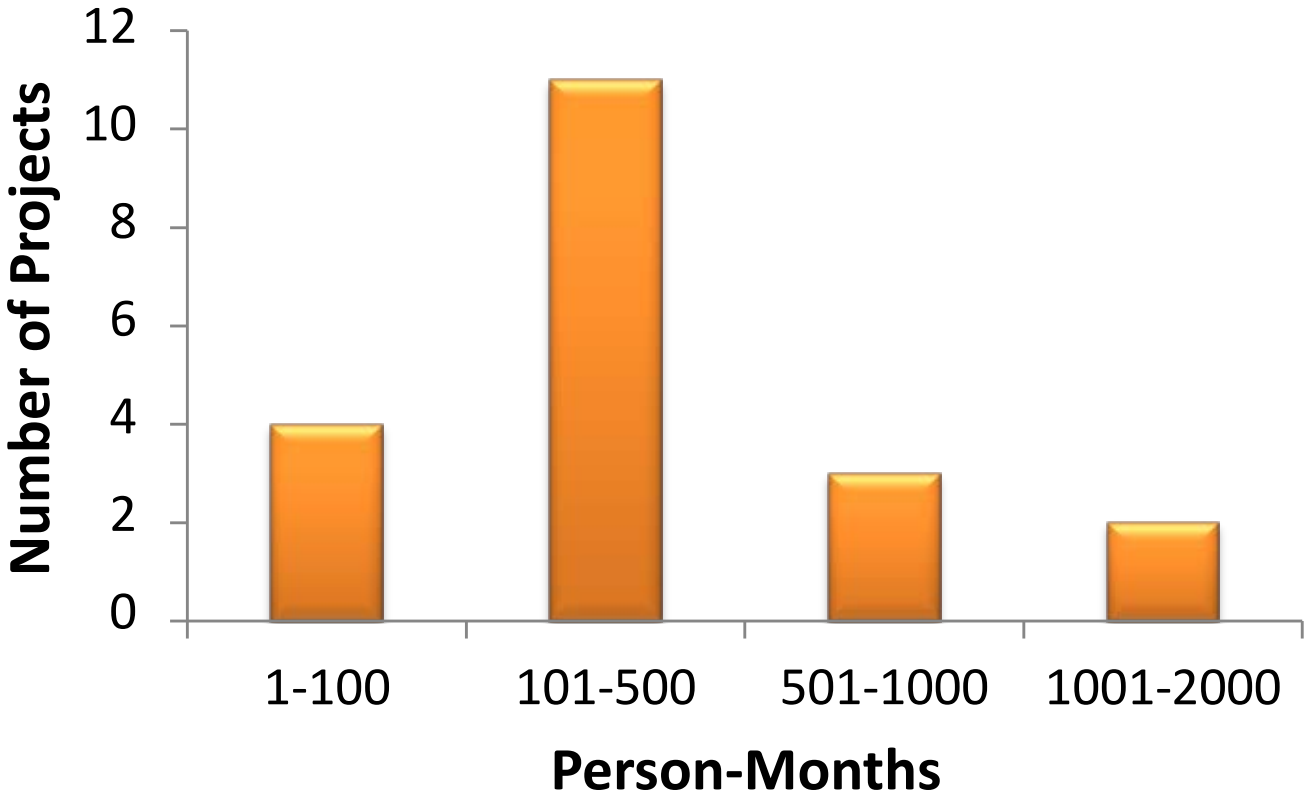
Dataset by Software Size* Range



Average software size is 704 Software Requirements

*Software Size refers to the Initial Software Requirements

Dataset by Expended Effort (in Person-Months)

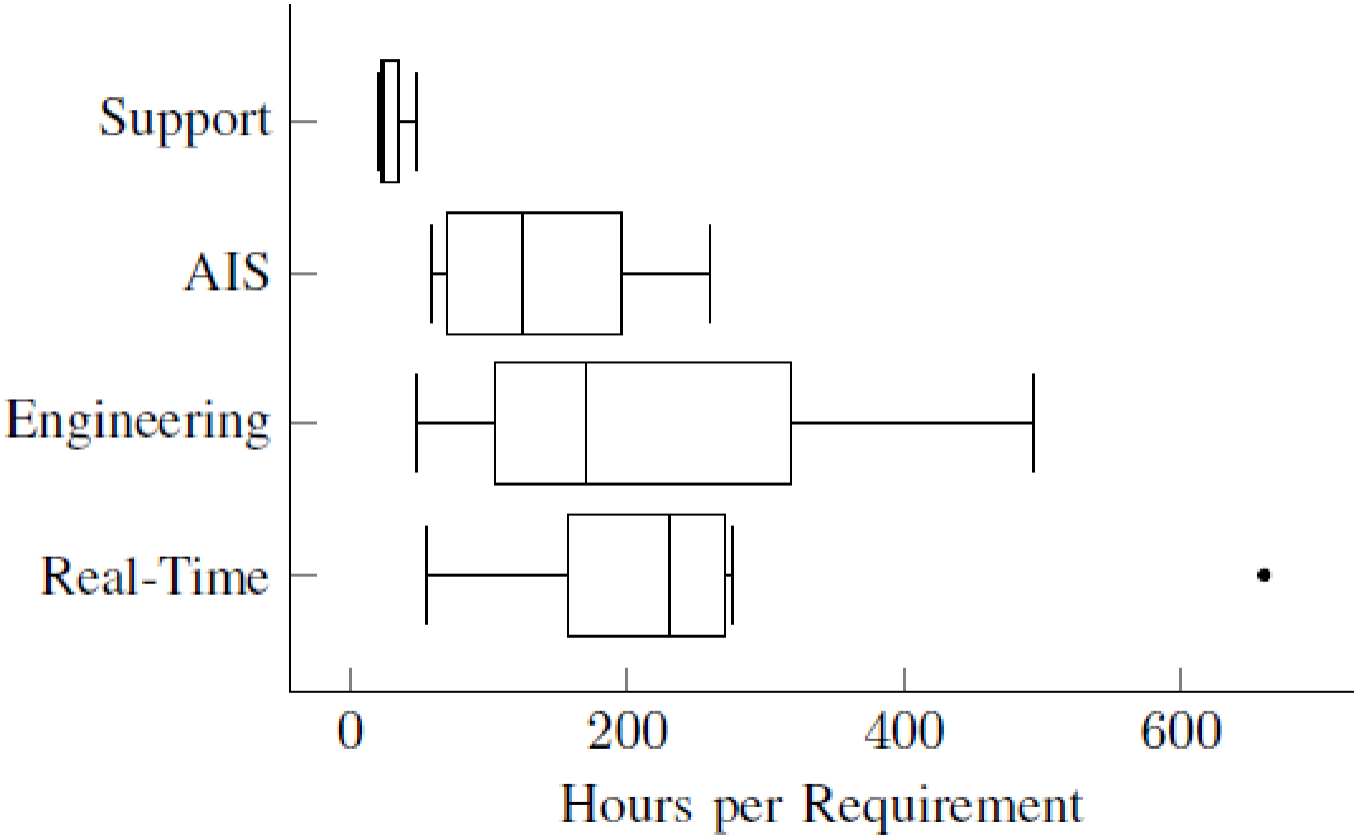


Average expended effort is 409 Person-Months

*Actual Effort Hours converted into Person Months using 152 hours/month

Productivity Benchmarks

Productivity by Super Domain



Grouping by Software Domain shows significant effect on Agile Software Productivity

Productivity Comparison
Agile vs Non-Agile

	Average Productivity*	
Size Range	Agile	Non-Agile
1-100	0.37	0.33
101-500	0.96	0.80
501-5000	1.97	1.16
Composite Average	0.8	0.66

* Initial Software Requirements per Person-Months

When grouped by Size, Agile Software Projects appear to be more productive

Agile Effort Estimation Models

Agile Effort Model Variables

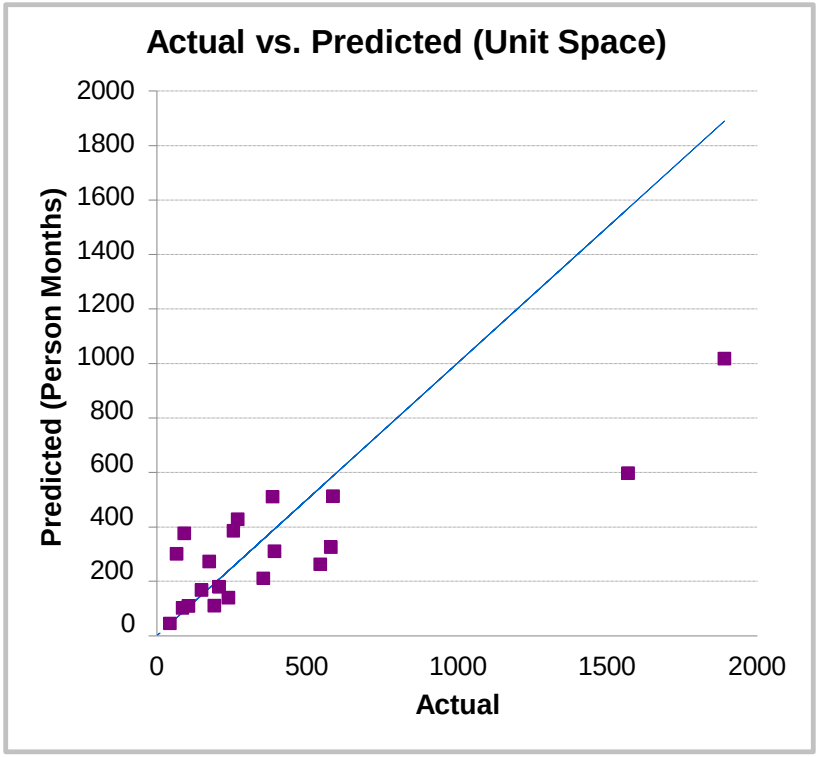
Name	Acronym	Type	Definition
Final Effort	EFFORT	Dependent	Actual software engineering effort (in Person-Months) at contract completion
Initial Software Requirements	REQ	Independent	Sum of Initial Functional Requirements and Initial External Interface Requirements collected at contract award. Counting convention based on “shall statements”
Initial Peak Staff	STAFF	Independent	Estimated peak team size at contract award, measured in full-time equivalent staff
Super Domain	SD	Categorical	Software primary application. Four Types: Mission Support, Automated Information System (AIS), Engineering, or Real Time

Agile Effort Estimation Model
(Single Variable)

Model	**Equation Form**	N	R^2%	CV%	Mean	MMRE%	REQ Min	REQ Max
1	Effort = 14.5 **x** REQ$^{0.5009}$	20	53	48	409	64	10	4,867

Effort = Final Effort (in Person Months) at contract completion

REQ = Initial Software Requirements at contract start



Coefficient Statistics:

Variable	P-value	VIF
Intercept	0.0000	
REQ	0.0002	
STAFF		
SD		

Agile Estimation Model not accurate when simply using REQ as input

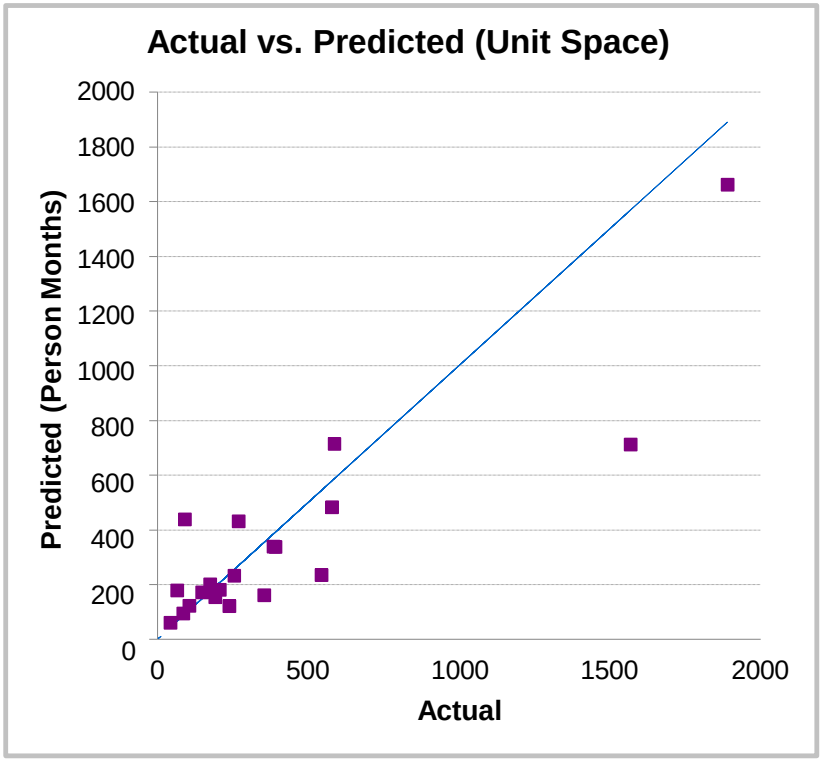
Agile Effort Estimation Model
(Two Variables)

Model	**Equation Form**	N	R^2%	CV%	Mean	MMRE %	REQ Min	REQ Max
2	Effort = 6.8 **x** REQ$^{0.4071}$ **x** STAFF$^{0.4404}$	20	60	36	409	52	10	4,867

Effort = Final Effort (in Person Months) at contract completion

REQ = Initial Software Requirements at contract start

STAFF = Initial (or Estimated) Peak Staff at contract start



Coefficient Statistics:

Variable	P-value	VIF
Intercept	0.0000	
REQ	0.0015	1.22
STAFF	0.0559	1.22
SD		

Agile Estimation Model improves when Peak Staff is treated with REQ

Agile Effort Estimation Model
(Three Variables)

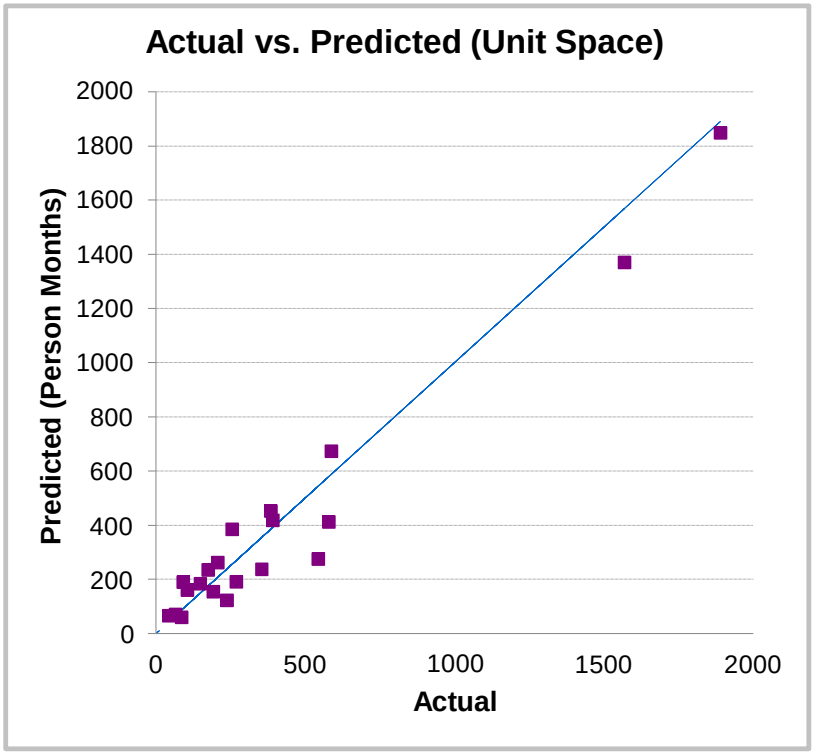
Mod	Equation Form	N	R^2 %	CV %	Mean	MMRE %	REQ Min	REQ Max
3	Effort = 1.3 **x** REQ$^{0.5126}$ **x** STAFF$^{0.4782}$ **x** SD$^{1.001}$	20	81	22	409	32	10	4,867

Effort = Final Effort (in Person Months) at contract completion

REQ = Initial Software Requirements at contract start

STAFF = Initial (or Estimated) Peak Staff at contract start

SD = 1 for Mission Support Super Domain (SD)
2 for Automated Information System SD
3 for Engineering SD
4 for Real Time SD



Coefficient Statistics:

Variable	P-value	VIF
Intercept	0.0000	
REQ	0.0000	1.45
STAFF	0.0045	1.37
SD	0.0003	1.07

Agile Estimation Model more accurate when all 3 variables are added

Conclusion

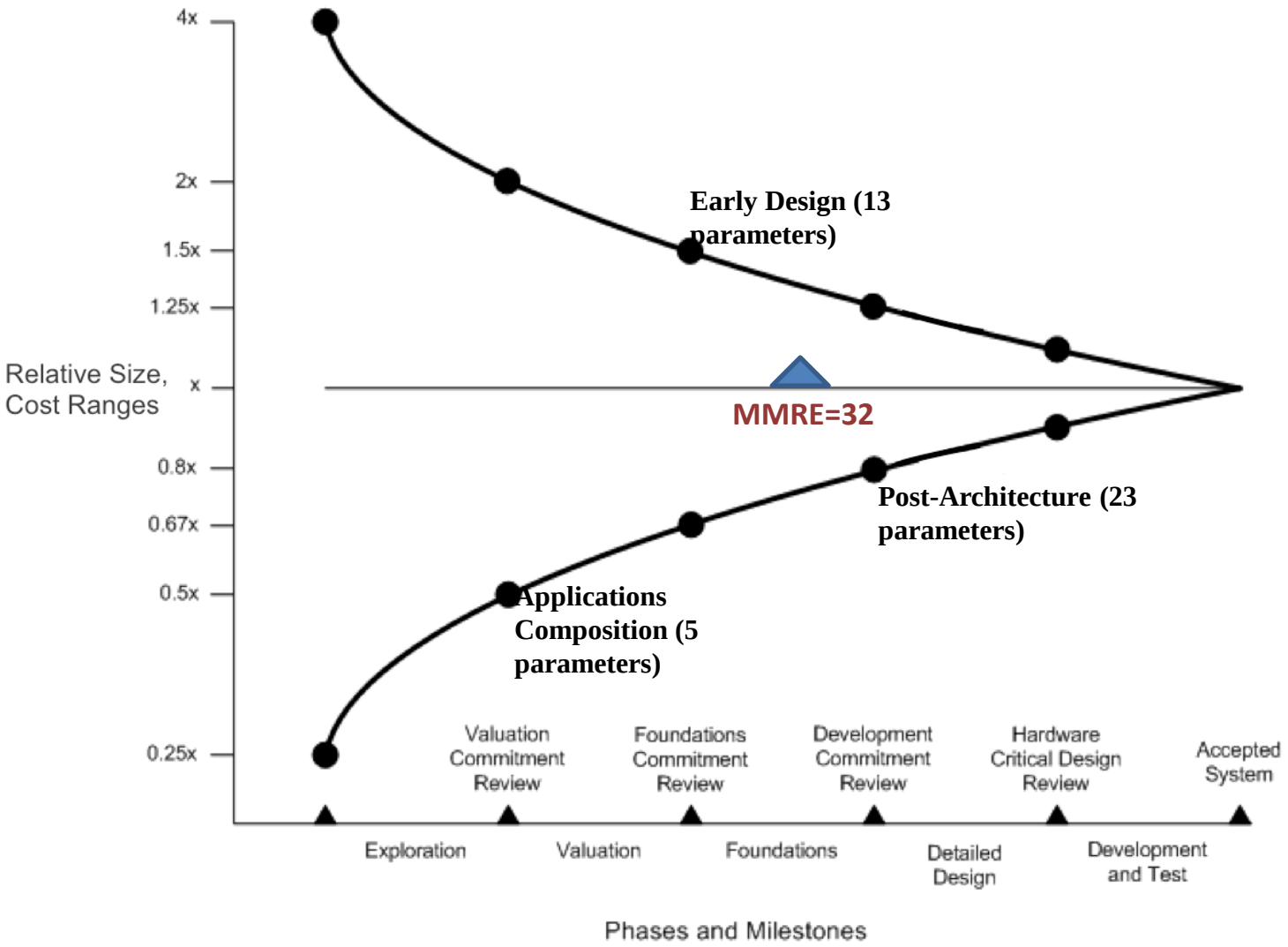
Primary Findings

- ü Initial Software Requirements* is a valid size proxy for Software Effort Estimation Models
- ü Models' accuracy improves when Peak Staff and Super Domain, are treated along with Initial Software Requirements*

Model	**Equation Form**	N	R^2%	CV%	MMRE%
1	Effort = 14.5 **x** $REQ^{0.5009}$	20	53	48	64
2	Effort = 6.8 **x** $REQ^{0.4071}$ **x** $STAFF^{0.4404}$	20	60	36	52
3	Effort = 1.3 **x** $REQ^{0.5126}$ **x** $STAFF^{0.4782}$ **x** $SD^{1.001}$	20	81	22	32

*Initial Software Requirements = Initial Functional Requirements + Initial External Interfaces

The Cone of Uncertainty



Model Limitations and Usefulness

- Since data was analyzed at the CSCI level, effort models may not be appropriate for projects reported at the Roll-Up Level.
- Do not use Effort Estimation Models if your input parameters are outside of the model's dataset range.
- **P**roposed Effort Models may be used to either crosscheck or validate contract proposals as input parameters used in the study are typically available during proposal evaluation phase
- **A**pplicable for both, Defense and Business Systems
- **A**pplicable for Agile Software Projects