Initial Results Building a Normalized Software Database using SRDRS

2009 ISPA/SCEA
Professional Development and Training Workshop
St. Louis, MO
2 – 5 June 2009

Michael Gallo
Paul Hardin
Elizabeth Koza
Robert Bailey

Sponsor: Deputy Assistant Secretary of the Army for Cost & Economics

Technomics
The Science of Informed Decision Making

Background

- The Army is collecting software data to build a database
- Two primary collection sources
  - SRDRs
  - Army internal collection
- Current research is focused on weapon system software, not AIS programs
- Research approach
  - Minimize contractor's effort to report the data (i.e. maximize use of artifacts used internally by the contractors)
  - Avoid the use of subjective data fields

Army's Desired Uses of the Data

- Productivity factors
- Parametric estimating equations
- Calibration of commercial software cost models
- Sizing estimates
- Visualizing trends
- Sanity checks

Selected SRDR Data Element Summary

Section I - General Context

- System/Element Name
- Report as-of date
- Authorizing Vehicle
- Development Organization
- Software Process Maturity
- Precedents
- SRDR Data Dictionary Filename
- Comments

Section II – Product Description

- Functional Description
- Software Development Characterization
- Application Type
  - Primary and Secondary Programming language
  - Percentage of Overall Product Size
  - Development Process
  - Upgrade or New Development?
  - SW Development Method
- Non-Developmental Software
  - COTS/GOTS Applications Used
  - Integration Effort (Optional)
- Staffing
  - Peak Staff
  - Peak Staff Date
  - Hours per Staff-Month
- Personnel Experience by Domain
- Comments

Selected SRDR Data Element Summary

Section III – Product Size

- Requirements Counts
  - Total Software Requirements
  - New Software Requirements
  - Total External Interface Requirements
  - New External Interface Requirements
  - Requirements Volatility
- Total Delivered Code Count
  - New Code
  - Reused With Modifications
  - Reused Without Modifications
  - Carryover Code
  - Auto-generated Code
  - Sub-contractor Code
  - Counting Convention
- Comments

- **Effort (staff-hours)**
- **Effort must be partitioned into a set of activities**
- **For <u>each</u> SW activity reported the contractor must provide:**
  - **WBS Element reference**
  - **Start Month**
  - **End Month**
  - **Prime Contractor hours**
  - **All Other Sub-ctr hours**

+ SRDR Data Dictionary

The Problem

- Data is collected from a variety of contractors and product mission areas and typically used and reflected as 'industry averages'
- Contractors use different definitions for reporting size, effort, and schedule data
  - There is no universal standard or mandate for software accounting and metric data
  - Data sources such as the SRDR permit contractors to tailor the report to the contractor's internal accounting and metrics systems
- Before the data can be used across projects/developers it must be made comparable

Approach For Normalizing SW Data

1. Review data quality
2. Identify both prevalent sizing metric in the data and prevalent sizing categories used
3. Identify prevalent software activities included in reported 'chunks' of effort, especially activities that are reported discretely.
4. Formulate a series of estimating equations that estimate constituent activities buried within reported 'chunks'; Review context for important variables that can be used to explain differences in the data; derive coefficients
5. Apply estimating equations *relatively* to estimate missing pieces and to break apart chunks into discrete activities

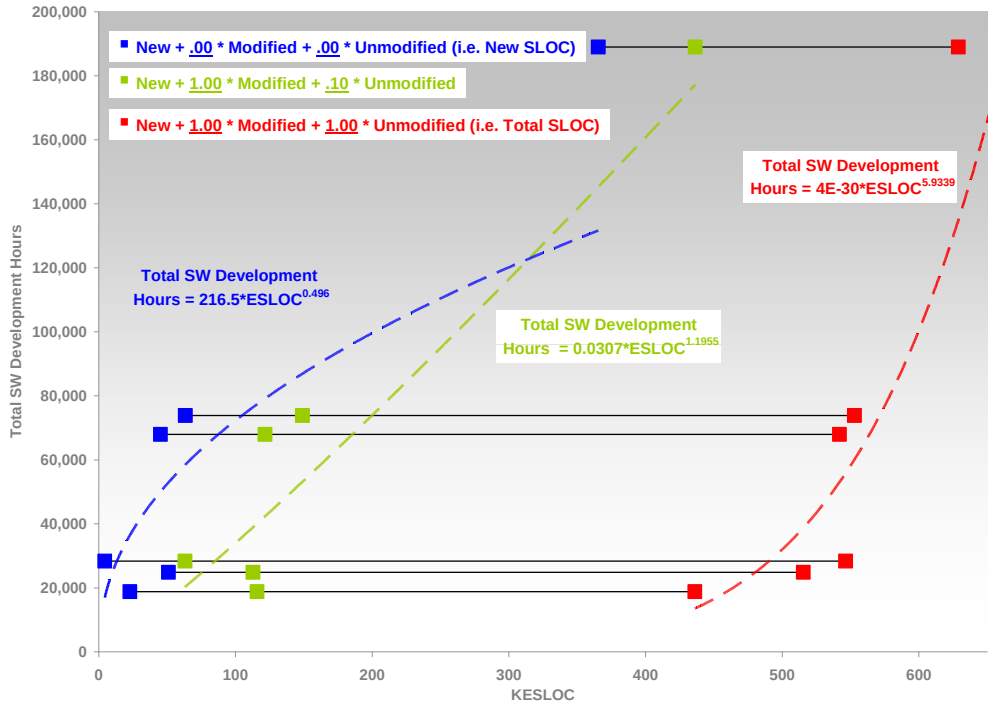
Review Data Quality

- Raw data points were filtered to remove data deemed inadequate
  - Missing or incomprehensible definitions
  - No size or effort reported
  - No language reported
  - No counting convention reported
  - Reported only Total SLOC
  - Foreign Military Sales (FMS)
- Many of the data points removed came from a 3rd party

- Understand the SW product
  - Mission, function and complexity of the software
  - Platform and operating environment of the software
  - Understand what programming languages were used
- Understand the development project
  - Characterization of the development work
  - Understand how the software product is put together
    - How it's integrated
    - How much was built with reused components
    - How much was auto-generated
- Understand who developed it
  - Primes
  - Subs
- Understand what's in the data reported
  - Scope of effort reported (what's included/excluded)
  - Understand the units of sizing and rules for sizing categories
- Understand other attributes that might drive cost/schedule/quality

Identify Prevalent Sizing Metric and Categories

- What are the basic sizing units of measure in the dataset?
  - What is the prevalent counting convention?
  - What are the programming languages used?
- What sizing categories are used?
  - Varying definitions are used
  - Additionally, Auto-Generated, Carryover, Deleted, COTS
- ESLOC weights have a significant effect on the normalized data and will ultimately influence
  - Computed productivity (ESLOC/hr)
  - Observed effort vs. Size trend



Each ■—■—■ represents one project, but with a different computation of that project's ESLOC.

Identify Prevalent SW Activities in the Data

Source

1	Detailed Design + Coding + Unit Testing		Sys Req + Req Def + Integration + FQT		SW PM + SQE + CM		Meetings				
2	DIT	CTS	CPTO	LIM	MIS	CFIN	REQ	SBVT	Demo	QA	CM
3	SW Requirements Analysis			SW Coding and Unit Testing			SW Developmental T&E		Other		
	Systems Engineering	Software Engineering		Integration & Test		Database	Mgmt & Direct Spt Functions		Subcontractor Effort		
	System Eng Requirements Analysis and Arch		SW Requirements Analysis and Arch	SW Design	SW Coding and Unit Testing	SW Integration and Test	System Integration and Test	Management, Support, and Labs			
4	SW Requirements Analysis	SW Architecture and Detailed Design	SW Coding and Unit Testing	Systems and System Test	Independent Test Group		Metrics, SCM, Documentation & Other Dept Support Efforts				
5	Requirements	Design, Code, Test		System Integration	SW CM		SW Qual	SW Mgmt			
6	Ada (Staff-months)		Ada83 (Staff-months)		Jovial (Staff-months)		Jovial/ASM (Staff-months)				
7	SW Requirements	Preliminary Design	Detailed Design		Code & Unit Test		Mega Level Test (MLT)		Element Level Test (ELT)		
	Detailed Design + Code & Unit Test + Mega Level Test + Element Level Test + PIV Defects										
8	Design and Document Review	Interface Design Documents	Coding		Code Reviews		Integration and Test		Other Documentation		
9	SW Requirements Analysis	SW Architecture and Detailed Design	SW Coding and Unit Testing	Software Qualification Testing	SW Integration and System/SW Integration		SW Developmental T&E		Other		
	Requirements Analysis	High Level Design, PDR	Detailed Design Iterations, CDR & OCSD	Code & Unit Test	SW/SW Integration Testing	SW/HW Integration Testing	Support to Systems, Test, ER & ILS	SDP, Management	SCM, Sys Admin, VDD		SQE
	System Requirements Definition	Implementation		SW/HW Test		System Integration		System Verification		CM, QA, PM	
10	SW Requirements Analysis	SW Coding and Unit Testing	SW Architecture and Detailed Design	SW Int. and System/SW Int.	SW Qualification Testing	SW Development Test and Evaluation		CM, SW Safety, SW Process Improvement			
11	SW Requirements Analysis	SW Coding and Unit Testing	SW Architecture and Detailed Design	SW Int. and System/SW Int.	SW Qualification Testing		Mgmt, SW CM, SW Process, System Admin Spt, Subcontract Mgmt Spt				
12	SW Requirements Analysis	SW Coding and Unit Testing	SW Architecture and Detailed Design		SW Int. and System/SW Int.		SW Qualification Testing				
13	Systems Engineering Activities			SW Engineering Activities			SW Integration Activities				
14	SW Requirements Analysis		SW Design/Code/Test & Integration		SW Quality Assurance		SW Configuration Mgmt		SW Environment Support		
	SW Requirements Analysis			SW Design/Code		Test and Integration			Other mgmt and spt (build related)		
	SW Requirements Analysis			SW Design/Code		SW Test and Integration	Help Desk and User Spt		Other mgmt and spt (build related)		
	SW Eng Requirements Analysis	SW Architecture and Detailed Design		SW Coding and Unit Testing	SW Integration and System/Software Integration		SW Development Test and Evaluation		CM, SW QA, and Dev Environment		
15	SW Requirements Analysis	SW Coding and Unit Testing	SW Architecture and Detailed Design	SW Integration and System/SW Integration		SW DT&E	SW Environmental Dev: SW Mgmt, CM, SW QM, Environmental Dev				
16	SW Requirements Analysis	SW Architecture and Detailed Design	SW Coding	SW Unit Testing	SW Integration and Test	SW Qualification Testing		Build Specific Systems Development			
17	SW Requirements Analysis	SW Design		SW Implementation and Unit Testing		SW Verification		QA, SCM, SW Environment			
18	SW Requirements Analysis		SW Architecture and Detailed Design		SW Coding and Unit Testing			SW Integration and System/SW Integration			
19	SW Requirements Analysis	SW Architecture & Detailed Design	SW Coding & Unit Testing	SW Integration & System/SW Integration		SW Developmental Test & Evaluation		Spiral Level Leadership		Other Direct Hours	
20	SW Requirements + Preliminary Design + Detailed Design + Code & Unit Test + SW I&T										
21	Requirements Analysis	Model Development	Design	C&UT	Integration & Test	FQT	PM	Data	QA	DevEnv	CM & QA

Each ▭ is a discretely reported 'chunk' of software activity.

Activities in blue font reflect data from an SRDR source

Identify SW Activities in the Data

Source											
1	Detailed Design + Coding + Unit Testing	Sys Req + Req Def + Integration + FQT	SW PM + SQE + CM	Meetings							
2	DIT	CTS	CPTO	LIM	MIS	CFIN	REQ	SBVT	Demo	QA	CM
3	SW Requirements Analysis	SW Coding and Unit Testing	SW Developmental T&E	Other							
3	Systems Engineering	Software Engineering	Integration & Test	Database	Mgmt & Direct Spt Functions	Subcontractor Effort					
3	System Eng Requirements Analysis and Arch	SW Requirements Analysis and Arch	SW Design	SW Coding and Unit Testing	SW Integration and Test	System Integration and Test	Management, Support, and Labs				
4	SW Requirements Analysis	SW Architecture and Detailed Design	SW Coding and Unit Testing	Systems and System Test	Independent Test Group	Metrics, SCM, Documentation & Other Dept Support Efforts					
5	Requirements	Design, Code, Test	System Integration	SW CM	SW Qual	SW Mgmt					
6	Ada (Staff-months)	Ada83 (Staff-months)	Jovial (Staff-months)	Jovial/ASM (Staff-months)							
7	SW Requirements	Preliminary Design	Detailed Design	Code & Unit Test	Mega Level Test (MLT)	Element Level Test (ELT)					
7	Detailed Design + Code & Unit Test + Mega Level Test + Element Level Test + PIV Defects										
8	Design and Document Review	Interface Design Documents	Coding	Code Reviews	Integration and Test	Other Documentation					
9	SW Requirements Analysis	SW Architecture and Detailed Design	SW Coding and Unit Testing	Software Qualification Testing	SW Integration and System/SW Integration	SW Developmental T&E	Other				
9	Requirements Analysis	High Level Design, PDR	Detailed Design Iterations, CDR & OCSD	Code & Unit Test	SW/SW Integration Testing	SW/HW Integration Testing	Support to Systems, Test, ER & ILS	SDP, Management	SCM, Sys Admin, VDD	SQE	
9	System Requirements Definition	Implementation	SW/HW Test	System Integration	System Verification	CM, QA, PM					
10	SW Requirements Analysis	SW Coding and Unit Testing	SW Architecture and Detailed Design	SW Int. and System/SW Int.	SW Qualification Testing	SW Development Test and Evaluation	CM, SW Safety, SW Process Improvement				
11	SW Requirements Analysis	SW Coding and Unit Testing	SW Architecture and Detailed Design	SW Int. and System/SW Int.	SW Qualification Testing	Mgmt, SW CM, SW Process, System Admin Spt, Subcontract Mgmt Spt					
12	SW Requirements Analysis	SW Coding and Unit Testing	SW Architecture and Detailed Design	SW Int. and System/SW Int.	SW Qualification Testing						
13	Systems Engineering Activities	SW Engineering Activities	SW Integration Activities								
14	SW Requirements Analysis	SW Design/Code/Test & Integration	SW Quality Assurance	SW Configuration Mgmt	SW Environment Support						
14	SW Requirements Analysis	SW Design/Code	Test and Integration	Other mgmt and spt (build related)							
14	SW Requirements Analysis	SW Design/Code	SW Test and Integration	Help Desk and User Spt	Other mgmt and spt (build related)						
14	SW Eng Requirements Analysis	SW Architecture and Detailed Design	SW Coding and Unit Testing	SW Integration and System/Software Integration	SW Development Test and Evaluation	CM, SW QA, and Dev Environment					
15	SW Requirements Analysis	SW Coding and Unit Testing	SW Architecture and Detailed Design	SW Integration and System/SW Integration	SW DT&E	SW Environmental Dev: SW Mgmt, CM, SW QM, Environmental Dev					
16	SW Requirements Analysis	SW Architecture and Detailed Design	SW Coding	SW Unit Testing	SW Integration and Test	SW Qualification Testing	Build Specific Systems Development				
17	SW Requirements Analysis	SW Design	SW Implementation and Unit Testing	SW Verification	QA, SCM, SW Environment						
18	SW Requirements Analysis	SW Architecture and Detailed Design	SW Coding and Unit Testing	SW Integration and System/SW Integration							
19	SW Requirements Analysis	SW Architecture & Detailed Design	SW Coding & Unit Testing	SW Integration & System/SW Integration	SW Developmental Test & Evaluation	Spiral Level Leadership	Other Direct Hours				
20	SW Requirements + Preliminary Design + Detailed Design + Code & Unit Test + SW I&T										
21	Requirements Analysis	Model Development	Design	C&UT	Integration & Test	FQT	PM	Data	QA	DevEnv	CM & QA

System Requirements	System Design	Requirements	Prototyping	Architecture	Initial Design	Detailed Design	Code & Unit Test	Formal Integration	Integration Testing	System Testing	Acceptance Testing	Field Testing	CM	Project Plans	PM	Quality Assurance	Process Improvement	User Documentation	Safety	IV&V	Life-Cycle Support	Independent Testing	Development Environment
1		1		1	1	1	1	1	1	1	1		1	1	1	1							
1		1		1	1	1	1	1	1	1	1		1	1	1	1							
1		1		1	1	1	1	1	1	1	1		1	1	1	1							
		1		1	1	1	1	1	1	1	1		1	1	1	1							
		1		1	1	1	1	1	1	1	1		1	1	1	1							
		1		1	1	1	1	1	1	1	1		1	1	1	1							
		1		1	1	1	1	1	1	1	1		1	1	1	1							
		1		1	1	1	1	1	1				1				1	1				1	
		1		1	1	1	1	1	1									1					
						1	1	1	1									1					
		1			1	1	1	1	1		1		1	1	1	1		1				1	1
						1	1	1	1									1					
						1	1	1	1									1					
						1	1	1	1									1					
						1	1	1	1									1					
		1		1	1	1	1	1	1	1	1	1	1	1	1	1							
		1		1	1	1	1	1	1		1		1	1	1	1	1		1	1			
		1		1	1	1	1	1	1	1	1		1	1	1	1	1		1				
		1		1	1	1	1	1	1	1	1		1	1	1	1	1		1				
		1		1	1	1	1	1	1		1												
		1		1	1	1	1	1	1					1				1	1				
1	1	1	1	1	1	1	1	1	1					1	1		1	1					
	1	1		1	1	1	1	1	1		1												
		1		1	1	1	1	1	1	1	1		1			1				1			1
1	1	1		1	1	1	1	1	1	1	1		1	1	1								
1	1	1		1	1	1	1	1	1	1	1		1	1	1						1		
		1		1	1	1	1	1	1	1													
1	1	1		1	1	1	1	1					1			1		1					
1	1	1		1	1	1	1	1	1	1	1			1	1			1					
1	1	1		1	1	1	1	1	1	1	1			1	1			1					

Blue font = SRDR Data Source

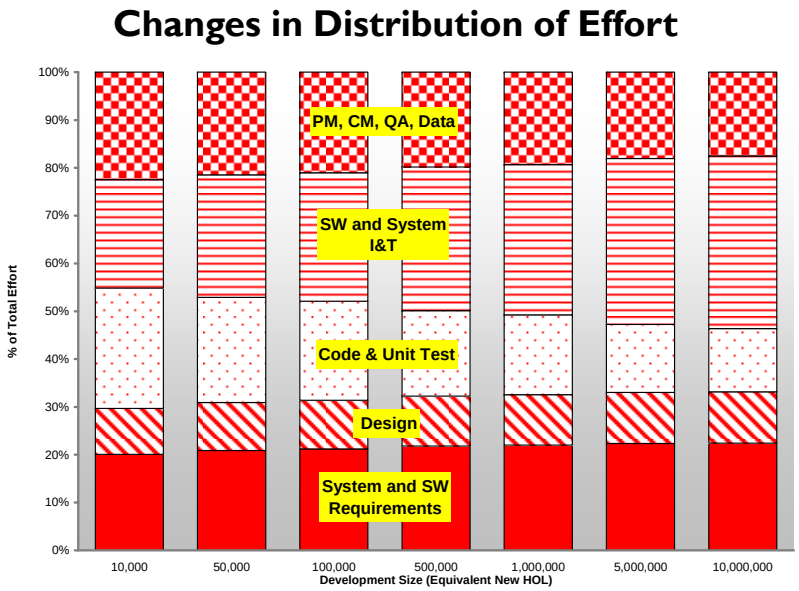
Graphical View of the 'Chunks'

System Requirements	System Design	Requirements	Architecture	Initial Design	Detailed Design	Code & Unit Test	Formal Integration	Integration Testing	System Testing	Acceptance Testing	CM	Project Plans	PM	Quality Assurance	Process Improvement	User Documentation	Safety	IV&V	Independent Testing	Development Environment
2		2	2	2	1	1	2	2	2	2	3	3	3,4	3						
2		2	2	2	1	1	2	2	2	2	3	3	3,4	3						
2		2	2	2	1	1	2	2	2	2	3	3	3,4	3						
		4	1	1	1	1	1	2	6	5	8	3	3	7						
		7	1	1	1	1,4	1	2	9	8	11	3,5,6	3,5,6	10						
		7	1	1	1	1,4	1	2	9	8	11	3,5,6	3,5,6	10						
		7	1	1	1	1,4	1	2	9	8	11	3,5,6	3,5,6	10						
		1	2	2	2	3	4	4			6				6	6			5	
		1	1	2	3	4	5	6								1,2,3, 4,5,6				
					1	1	1	1								1				
		1		2,3	3	4	5	5		6	11,13, 14,15	7,13, 14,15	7,13, 14,15	9,11, 12,14,1 5		8,13, 15			5	10,12
					1	1	1	1								1				
					1	1	1	1								1				
					1	1	1	1								1				
					1	1	1	1								1				
		1	2	2	2	3	4	4	4	5	6	6	6	6						
		1	2	2	2	3	4	4		5	7	7	7	7	7		7	6		
		1	2	2	2	3	4	4	5	5	6	6	6	6	6		6			
		1	2	2	2	3	4	4	5	5	6	6	6	6	6		6			
		1	2	2	2	3	4	4		5										
		1	1	1	2	2	2	3				1				2	1			
1	1	2	2	3	3	4	5	6				7	7		7	7				
	7	1	2	2	2	3,4	5	5		6										
		1	2	2	2	3	4	4	4	4	5			5				4		5
1	1	1	2	2	2	2	3	3	3	3	4	4	4							
1	1	1	2	2	2	2	3	3	3	3	4	4	4							
		1	2	2	2	3	4	4	4											
1	1	2	2	2	3	3	4				6			6		5				
1	1	1	2	2	2	3	3	4	5	5		6	6			6				
1	1	1	2	2	2	3	3	4	5	5		6,7	6,7			6				

= Discretely reported activity = Activity is included in a 'chunk' total = Activity was not reported

Some Approaches to Normalize Effort

- Filter to include projects with the same set of activities
  - Need many data points
  - Or filter to maximize the # of activities (results in a small # of projects)
  - Or Filter to maximize the # of projects (results in a small # of activities)
- Use a factor to fill in missing activities
  - Can't derive a factor in isolation to add/remove activities because:
  - Factor can change as size increases
  - Need unique factors for each unique 'chunk'
- Derive non-linear equation in isolation to add/remove activities
  - Basically, filtering on activities reported discretely (greens)
  - Multiple projects must have reported activity discretely
- Would result in dropping data that contain actuals that are included in chunks (yellows)



Source: VERA

Our Analytical Approach

- Derive an effort estimating relationship (EER) for each software development activity. Use the EERs to break apart the chunks (yellows) & estimate the missing activities (reds)
- Relies on both the discretely reported activities (greens) and information buried in the chunks.

System Requirements	System Design	Requirements	Architecture	Initial Design	Detailed Design	Code & Unit Test	Formal Integration	Integration Testing	System Testing	Acceptance Testing	CM	Project Plans	PM	Quality Assurance	Process Improvement	User Documentation	Safety	IV&V	Independent Testing	Development Environment
2		2	2	2	1	1	2	2	2	2	3	3	3,4	3						
2		2	2	2	1	1	2	2	2	2	3	3	3,4	3						
2		2	2	2	1	1	2	2	2	2	3	3	3,4	3						
		4	1	1	1	1	1	2	6	5	8	3	3	7						
		7	1	1	1	1,4	1	2	9	8	11	3,5,6	3,5,6	10						
		7	1	1	1	1,4	1	2	9	8	11	3,5,6	3,5,6	10						
		7	1	1	1	1,4	1	2	9	8	11	3,5,6	3,5,6	10						
		1	2	2	2	3	4	4			6				6	6			5	
		1	1	2	3	4	5	6								1,2,3, 4,5,6				
					1	1	1	1								1				
		1		2,3	3	4	5	5		6	11,13, 14,15	7,13, 14,15	7,13, 14,15	9,11, 12,14,1 5		8,13, 15			5	10,12
					1	1	1	1								1				
					1	1	1	1								1				
					1	1	1	1								1				
					1	1	1	1								1				
		1	2	2	2	3	4	4	4	5	6	6	6	6						
		1	2	2	2	3	4	4		5	7	7	7	7	7		7	6		
		1	2	2	2	3	4	4	5	5	6	6	6	6	6		6			
		1	2	2	2	3	4	4	5	5	6	6	6	6	6		6			
		1	2	2	2	3	4	4		5										
		1	1	1	2	2	2	3				1				2	1			
1	1	2	2	3	3	4	5	6				7	7		7	7				
	7	1	2	2	2	3,4	5	5		6										
		1	2	2	2	3	4	4	4	4	5			5				4		5
1	1	1	2	2	2	2	3	3	3	3	4	4	4							
1	1	1	2	2	2	2	3	3	3	3	4	4	4							
		1	2	2	2	3	4	4	4											
1	1	2	2	2	3	3	4				6			6		5				
1	1	1	2	2	2	3	3	4	5	5		6	6			6				
1	1	1	2	2	2	3	3	4	5	5		6,7	6,7			6				

General Formulation of Approach

- Each 'chunk' consists of a set of estimated activities

$$\hat{E}_{\mathrm{CH}} = \left(\sum \hat{E}_{\mathrm{i}}\right)$$

- Total estimated project effort consists of a set of estimated chunks

$$\hat{E}_{\mathrm{T}} = \left(\sum \hat{E}_{\mathrm{CH}}\right)$$

E_T = Total Reported Project Effort

$\hat{E}_{\mathrm{T}}$ = Total Estimated Project Effort

$\hat{E}_{\mathrm{CH}}$ = Estimated Chunk Effort

$\hat{E}_{\mathrm{i}}$ = Estimated effort of activity i

- Each effort estimating relationship (EER) provides an estimate for one activity

$$\hat{E}_{\mathrm{i}} = \mathrm{A}_{\mathrm{i}}[k_{\mathrm{i}}(\mathrm{ESLOC}_{\mathrm{i}})^{b_{\mathrm{i}}} * (EAF_1)^{\mathrm{Platform}} * (EAF_2)^{\mathrm{Dev\_Type}} * (EAF_3)^{\mathrm{CMM}} * (EAF_4)^{\mathrm{SystemEffect}}]$$

where

$\mathrm{A}_{\mathrm{i}} = 1$ if activity i is included in total effort (E_{T}); else $\mathrm{A}_{\mathrm{i}} = 0$

$\mathrm{ESLOC}_{\mathrm{i}}$ = Equivalent New Lines of Code for Activity i

Platform = 1 if platform is Air; else Platform = 0

Developer Type = 1 if single developer or subcontractor; else Developer Type = 0

CMM = CMM Level of Developer - Avg CMM Level of Database

SystemEffect = 1 if system is part of larger development project; else SystemEffect = 0

- ESLOC is specified uniquely for each activity

$$\mathrm{ESLOC}_{\mathrm{i}} = New * CF_i + \sum_{j=1}^{m} Not\_New_j * CF_i * w_{i_j}$$

where CF_{i} = Counting Convention Factor

w_{i} = Equivalent new fractional cost

j = SLOC categories that are considered "Not New"

Challenges Encountered

- Not enough data points. 211 unknowns, but only 168 d.o.f. in the data
- 'Partial' programs drove analysis to illogical results
- Required significant tailoring
  - Reduced SLOC categories
  - Common ESLOC weights
  - Common D.O.S. for groups of activities

- Use constrained optimization to solve for coefficients that minimize residuals of estimated effort at both the project level and chunk level

Initial Results

- $\text{Effort} = A * \text{ESLOC}^{B} * C^{\text{PLATFORM}} * D^{\text{DEV-TYPE}} * E^{(\text{CMM-4.23})}$

SW Activity[1]	A	B	C	D	E
Requirements	0.00470	1.1100	1.6	0.89	0.91
Architecture	0.00110	1.1733	1.6	0.89	0.91
Initial Design	0.00190	1.1733	1.6	0.89	0.91
Detailed Design	0.00450	1.1733	1.6	0.89	0.91
Code & Unit Test	0.03000	1.0893	1.6	0.89	0.91
Formal Integration	0.00400	1.2393	1.6	0.89	0.91
Integration Testing	0.01540	1.2393	1.6	0.89	0.91
System Testing	0.00450	1.2393	1.6	0.89	0.91
Acceptance Testing	0.00370	1.2393	1.6	0.89	0.91
Configuration Mgt	0.00540	1.1116	1.6	0.89	0.91
Project Plans	0.00100	1.1116	1.6	0.89	0.91
Program Mgt	0.00390	1.1116	1.6	0.89	0.91
Quality Assurance	0.00800	1.1116	1.6	0.89	0.91

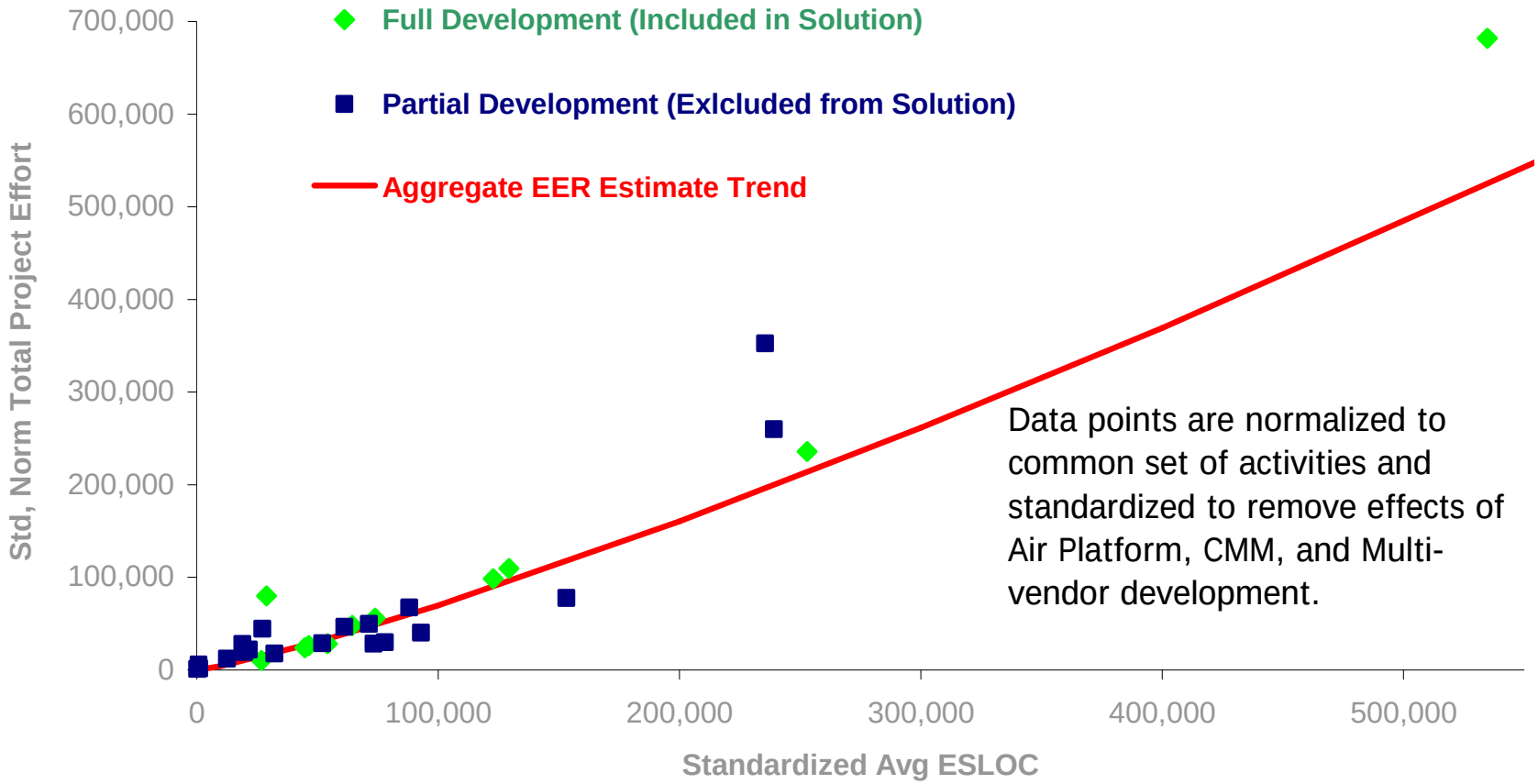
SW Activity	ESLOC				
	New	Mod	Reused	Carryover	Autogen
Requirements	1.0000	0.5624	0.2200	0.0420	0.0400
Architecture	1.0000	0.5624	0.2200	0.0420	0.0400
Initial Design	1.0000	0.5624	0.2200	0.0420	0.0400
Detailed Design	1.0000	0.5624	0.2200	0.0420	0.0400
Code & Unit Test	1.0000	0.5624	0.2200	0.0420	0.0400
Formal Integration	1.0000	0.5624	0.2200	0.0420	0.0400
Integration Testing	1.0000	0.5624	0.2200	0.0420	0.0400
System Testing	1.0000	0.5624	0.2200	0.0420	0.0400
Acceptance Testing	1.0000	0.5624	0.2200	0.0420	0.0400
Configuration Mgt	1.0000	0.5624	0.2200	0.0420	0.0400
Project Plans	1.0000	0.5624	0.2200	0.0420	0.0400
Program Mgt	1.0000	0.5624	0.2200	0.0420	0.0400
Quality Assurance	1.0000	0.5624	0.2200	0.0420	0.0400

[1]Included in Normalized Database

SLOC = Logical Lines of Code

	Average Percent Error		PRED(20)	
	Data Pts Included	Data Pts Excluded	Data Pts Included	Data Pts Excluded
Project	4%	200%	80%	20%
Chunk	169%	344%	26%	9%

Initial Results (Cont'd)



Application of EERs

Breaking Apart Chunks ("Yellows")

$$\text{Factor}_i = \frac{\text{EER}_{i_j}}{\sum \text{EER}_{i_j}}$$

where

EER_{i_j} = EER Estimate of activity i in chunk j

$\sum \text{EER}_{i_j}$ = Estimate of all activities that are included in chunk j

$$\hat{E}_i = \text{Factor}_i * E_{CH_j}$$

where E_{CH_j} = Actual effort of chunk j

Estimating Missing Activities ("Reds")

$$\text{Factor}_i = \frac{\text{EER}_i}{\sum \text{EER}_i}$$

where

EER_i = EER Estimate of activity i

$\sum \text{EER}_i$ = Estimate of all activities that are included in the actual project effort

$$\hat{E}_i = \text{Factor}_i * E_{\text{Project}}$$

where E_{Project} = Actual Project Effort

- Estimating equations are used *relatively* to add/remove activities from normalized effort

Reporting Normalized Data

Project	Platform	Developer Type	Completeness	Total Raw Effort (Hours)	Average ESLOC	Requirements Hours	Architecture Hours	Initial Design Hours	Detailed Design Hours	Code & Unit Test Hours	Formal Integration Hours	Integration Testing Hours	System Testing Hours	Acceptance Testing Hours	CM Hours	Project Plans Hours	PM Hours	Quality Assurance Hours	Total Normalized Effort (Hours)
1	ground	multi	partial	1,565	120	393	27	46	109	304	7	28	8	7	197	14	53	292	1,484
2	ground	multi	partial	5,447	783	288	102	176	106	402	146	3,552	665	543	340	61	246	504	7,131
3	ground	multi	partial	1,646	1,076	343	68	118	281	784	3	13	184	36	90	17	65	134	2,136
4	air	multi	partial	32,366	12,523	299	151	260	619	8,608	998	3,843	3,221	2,632	1,845	343	1,338	2,737	26,893
5	ground	multi	partial	32,080	18,849	3,013	1,303	328	779	3,186	2,256	9,786	3,774	3,083	1,284	292	1,137	1,904	32,124
6	ground	multi	partial	21,601	19,833	11,546	1,237	2,136	5,084	1,356	41	156	46	2,210	914	170	663	1,356	26,914
7	ground	multi	partial	23,976	21,606	6,386	751	1,298	3,088	6,481	172	661	3,039	4,598	1,016	189	737	1,507	29,923
8	air	single	full	16,090	26,730	294	169	292	695	1,195	1,204	5,294	1,218	3,305	516	200	781	928	16,090
9	ground	multi	partial	77,519	27,183	12,868	645	1,115	2,652	21,497	1,248	4,805	7,680	7,182	2,407	448	1,745	3,571	67,862
10	ground	single	full	64,931	28,926	7,604	786	1,274	3,774	12,673	1,450	5,584	6,331	16,693	943	1,564	6,099	1,558	66,334
11	ground	single	partial	14,010	32,214	577	258	446	1,060	2,940	1,860	7,161	2,092	1,709	665	124	482	986	20,360
12	air	single	full	38,132	44,730	953	404	698	1,661	2,446	2,978	17,433	3,177	3,799	1,557	506	1,972	548	38,132
13	air	single	full	41,999	46,358	1,023	492	849	2,020	2,740	3,630	21,982	2,400	3,500	681	333	1,299	1,051	41,999
14	ground	multi	partial	42,994	51,956	1,507	2,184	3,773	8,978	6,726	4,571	6,270	362	296	1,369	575	2,243	2,031	40,885
15	air	single	full	38,094	54,060	866	400	691	2,544	6,753	2,986	11,495	3,359	2,744	1,903	354	993	2,823	37,911
16	ground	single	partial	37,208	61,186	1,451	676	1,167	2,777	7,296	5,082	19,564	5,717	4,670	1,673	311	1,213	2,482	54,079
17	air	single	full	64,132	64,383	1,321	617	1,066	3,990	10,440	4,657	17,930	5,239	4,280	3,883	722	3,947	5,761	63,854
18	ground	single	partial	39,669	71,365	1,526	718	1,240	2,950	7,651	5,453	20,995	6,135	5,012	1,761	327	1,277	2,612	57,657
19	ground	multi	partial	39,483	73,279	770	679	1,173	2,792	5,558	3,977	1,795	1,500	1,225	1,220	3,564	13,901	1,811	39,967
20	air	single	full	74,651	73,880	1,124	530	915	5,892	15,239	4,036	15,537	4,540	3,709	5,782	1,075	7,457	8,578	74,414
21	ground	multi	partial	20,764	77,986	2,628	1,243	2,147	1,917	4,937	624	16,020	4,681	3,824	1,505	247	963	2,234	42,971
22	ground	single	partial	53,420	88,061	2,018	962	1,661	3,952	10,072	7,409	28,523	8,335	6,809	2,329	433	1,689	3,455	77,645
23	air	multi	partial	81,325	92,887	384	2,609	4,506	10,721	16,998	3,752	14,445	6,267	5,120	2,597	483	1,883	3,853	73,619
24	air	single	full	174,537	129,324	2,608	1,878	3,244	7,719	19,045	14,841	97,263	6,370	6,709	2,337	1,152	4,492	6,879	174,537
25	ground	single	full	73,910	122,843	8,414	581	1,004	2,389	31,427	3,008	11,579	8,781	7,173	1,826	437	1,705	3,489	81,815
26	ground	multi	partial	91,047	153,126	5,151	3,599	6,216	14,791	30,926	2,541	9,784	2,859	5,303	2,077	386	1,506	3,082	88,221
27	ground	multi	partial	542,601	235,527	42,696	13,640	23,560	56,060	131,525	13,475	51,878	15,159	12,383	75,390	14,019	54,673	22,223	526,681
28	ground	multi	partial	400,258	239,128	48,392	12,648	21,847	51,984	121,807	9,628	37,069	10,832	8,848	25,129	4,673	18,224	16,753	387,835
29	ground	single	full	144,307	252,827	4,957	2,526	4,363	10,382	24,213	20,863	80,323	23,471	19,173	5,732	1,066	4,157	8,504	209,729
30	ground	multi	full	472,193	534,768	4,257	2,275	5,544	58,678	129,521	57,345	189,629	73,148	59,753	16,234	3,019	11,773	24,085	635,261

Conclusions and Recommendations

- This approach requires significant analysis
  - Good data dictionaries are crucial
  - Additional data points allows refinement of variables and incorporation of additional factors
- Ideally, data collected and used that is *facility and product-line specific*, will be easier to normalize
- Some advocate a more rigid and standardized SRDR data form to minimize this labor intensive normalization effort
  - What happens when the data requested does not fit with a contractor's internal accounting and metrics systems?
  - What happens when SW tools, processes, and technology changes?
  - Do standardized data collection forms risk creating a mirage of clean data?

Contacts

- Technomics (www.technomics.net)
  - Mike Gallo, Vice President
    Phone (571) 366-1405
    E-mail: mgallo@technomics.net
- DASA-CE
  - Jim Judy, Chief, Networks, Information, Software & Electronics Costing (NISEC) Division
    Phone (703) 601-4168
    E-mail: James.Judy@us.army.mil
  - Noel Bishop, Software Team Leader
    Phone (703) 601-4136
    E-mail noel.bishop@hqda.army.mil