



University of Southern California
Center for Systems and Software Engineering



Building Cost Estimating Relationships for Acquisition Support

15th Annual PSM Conference

Brad Clark, Ray Madachy, Barry Boehm, Thomas Tan
Wilson Rosa, Sponsor

15 July 2011



Agenda

- • Background
- Software Resources Data Report Issues
- Data Conditioning Solutions
- Communication Domain Preliminary Analysis



Project Background

- Goal is to improve the quality and consistency of estimating methods across cost agencies and program offices through guidance, standardization, and knowledge sharing.
- Project led by the Air Force Cost Analysis Agency (AFCAA) working with service cost agencies, and assisted by University of Southern California and Naval Postgraduate School
- We will publish the AFCAA Software Cost Estimation Metrics Manual to help analysts and decision makers develop accurate, easy and quick software cost estimates for avionics, space, ground, and shipboard platforms.

Stakeholder Communities

- Research is collaborative across heterogeneous stakeholder communities who have helped us in refining our data definition framework, domain taxonomy and providing us project data.
 - Government agencies
 - Tool Vendors
 - Industry
 - Academia

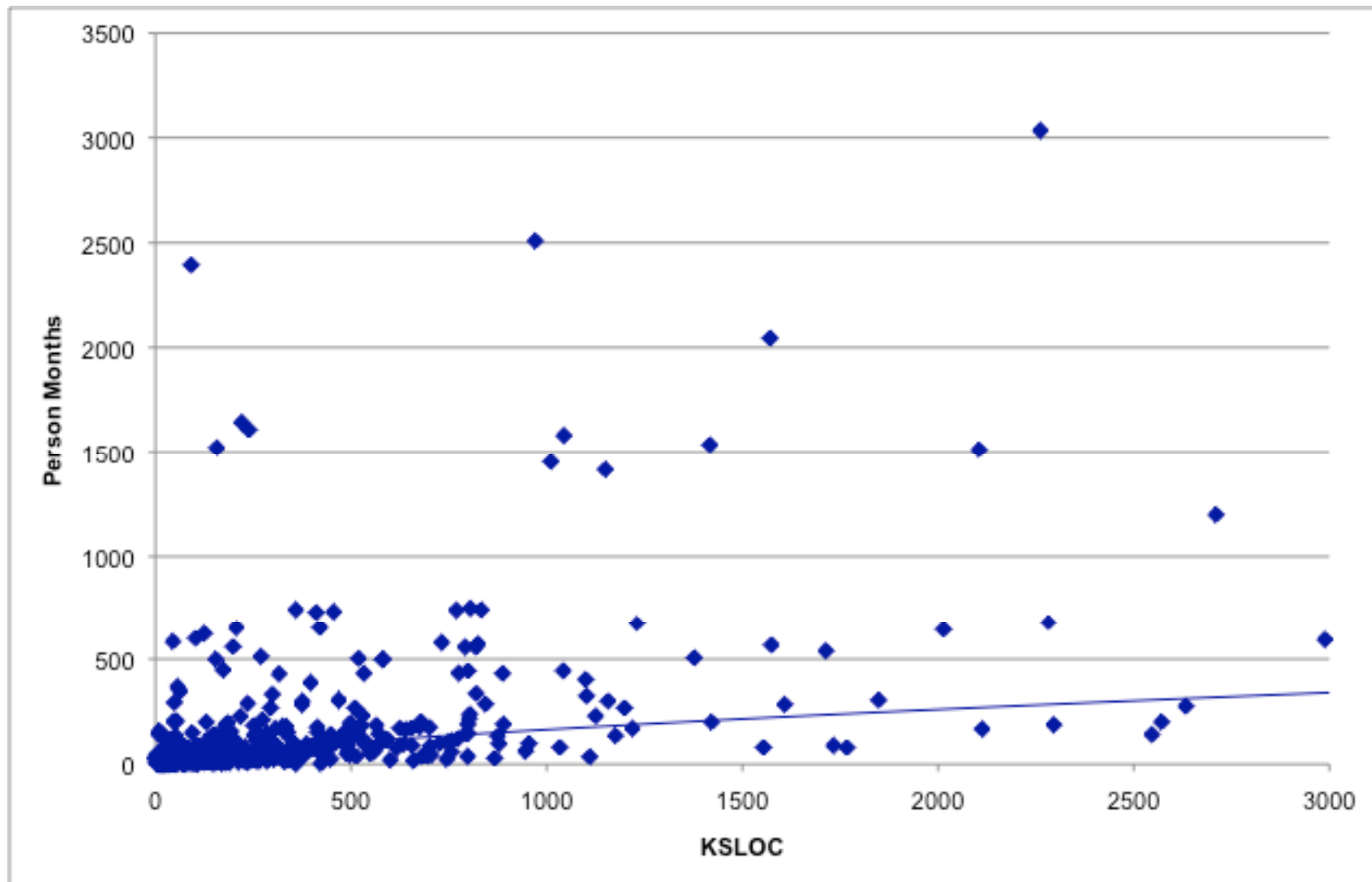




Data Source

- The **Software Resources Data Report** (SRDR) is used to obtain both the estimated and actual characteristics of new software developments or upgrades.
- All contractors, developing or producing any software development element with a projected software effort greater than \$20M (then year dollars) on major contracts and subcontracts within ACAT I and ACAT IA programs, regardless of contract type, must submit SRDRs.
- Reports mandated for
 - Initial Government
 - Initial Developer
 - Final Developer

SRDR Raw Data (520 observations)





Data Challenges

- Inadequate information on modified code (only size provided)
- Inadequate information on size change or growth
- Size measured inconsistently

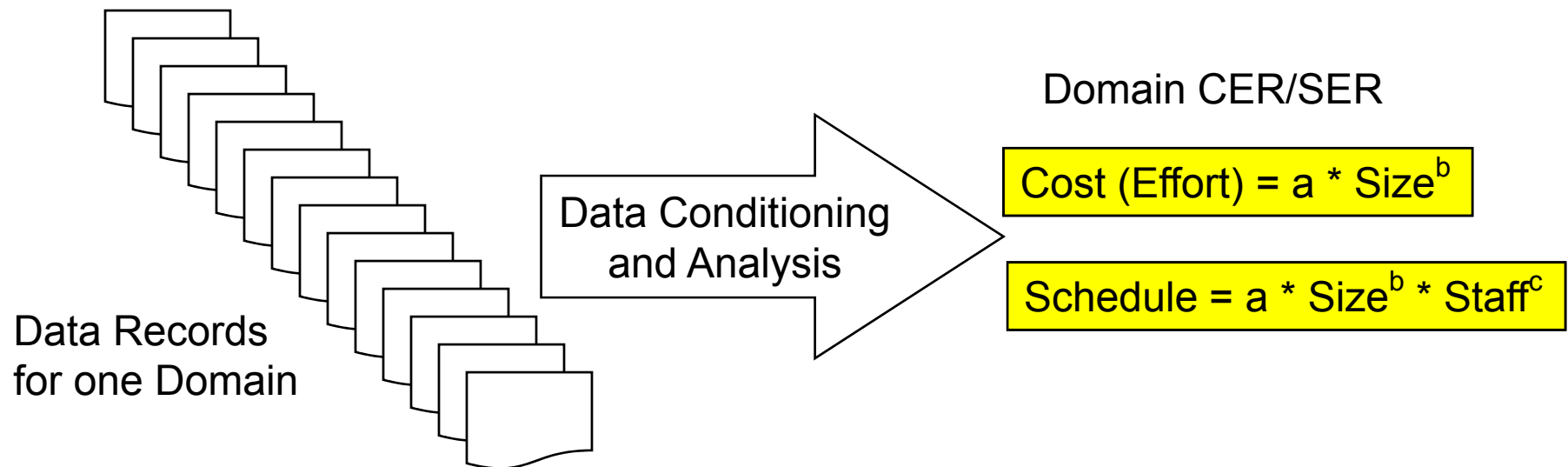
- Inadequate information on average staffing or peak staffing
- Inadequate information on personnel experience
- Inaccurate effort data in multi-build components
- Missing effort data

- Replicated duration for multi-build components
- Inadequate information on schedule compression
- Missing schedule data

- No quality data

Research Objectives

- Make collected data useful to oversight and management entities
 - Provide guidance on how to condition data to address challenges
 - Segment data into different Application Domains and Operating Environments
 - Analyze data for simple Cost Estimating Relationships (CER) and Schedule Estimating Relationships (SER) within each domain
 - Develop rules-of-thumb for missing data



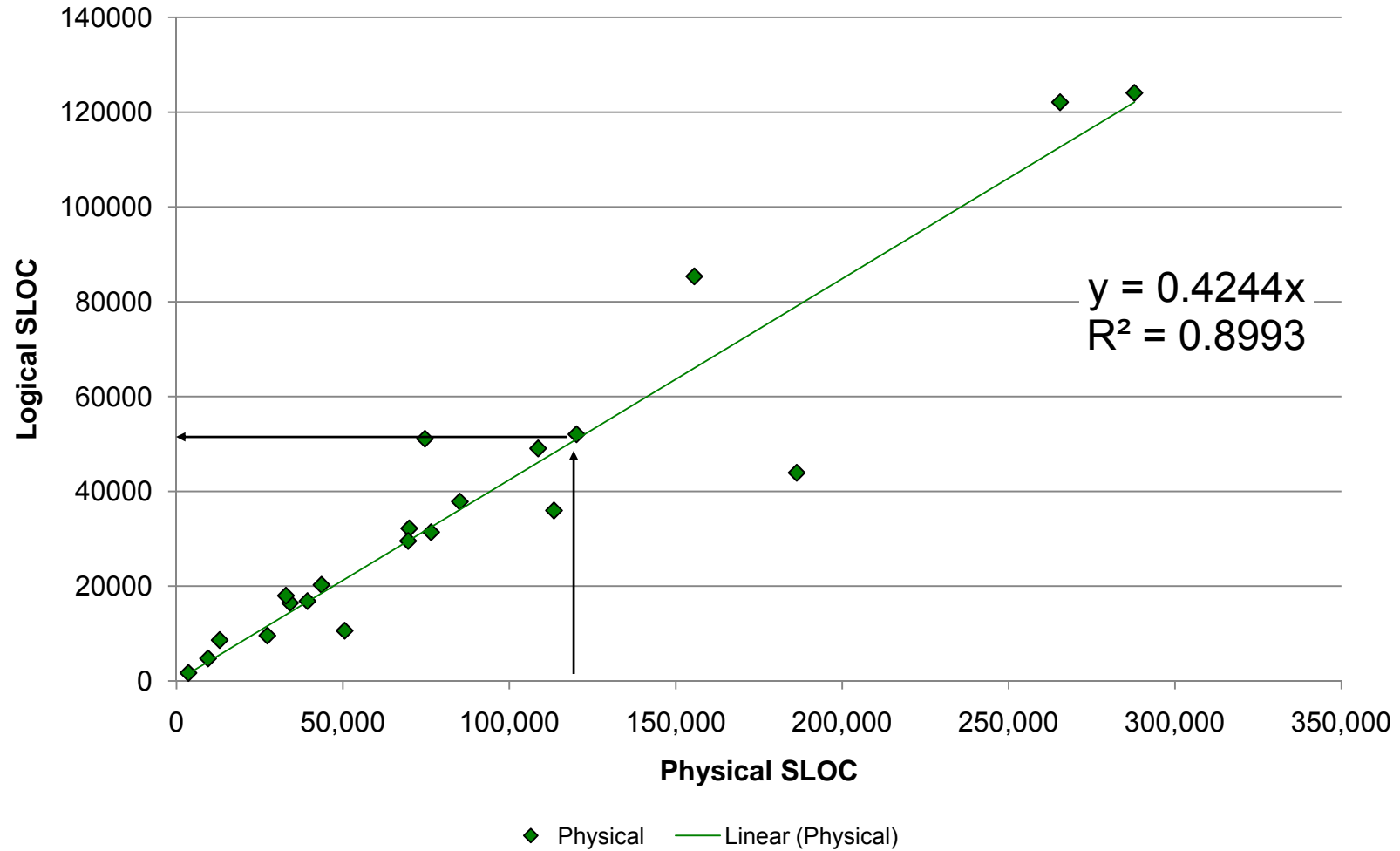
Data Segmentation

App Domain	Avionics	Fixed Ground	Missile	Mobile Ground	Shipboard	Unmanned Airborne	Unmanned Space	Total
Business Systems		6		4	2			12
Command & Control	1	41		16	35			93
Communications	4	77			17		2	100
Controls & Displays	8	6		2	5			21
Executive		4			3			7
Information Assurance		1						1
Infrastructure or Middleware		11			23			34
Maintenance & Diagnostic	1				5			6
Mission Management	42	2	3	2		1		50
Mission Planning	1	17						18
Modeling & Simulation		1						1
Process Control		3		6	1			10
Scientific Systems					3			3
Sensor Control and Processing	12	15			18			45
Simulation & Modeling		19			17			36
Spacecraft BUS							9	9
Spacecraft Payload							16	16
Test & Evaluation		2			2			4
Tool & Tool Systems		6	1					7
Training				2	6			8
Weapons Delivery and Control	11		19		9			39
Total	80	211	23	32	146	1	27	520

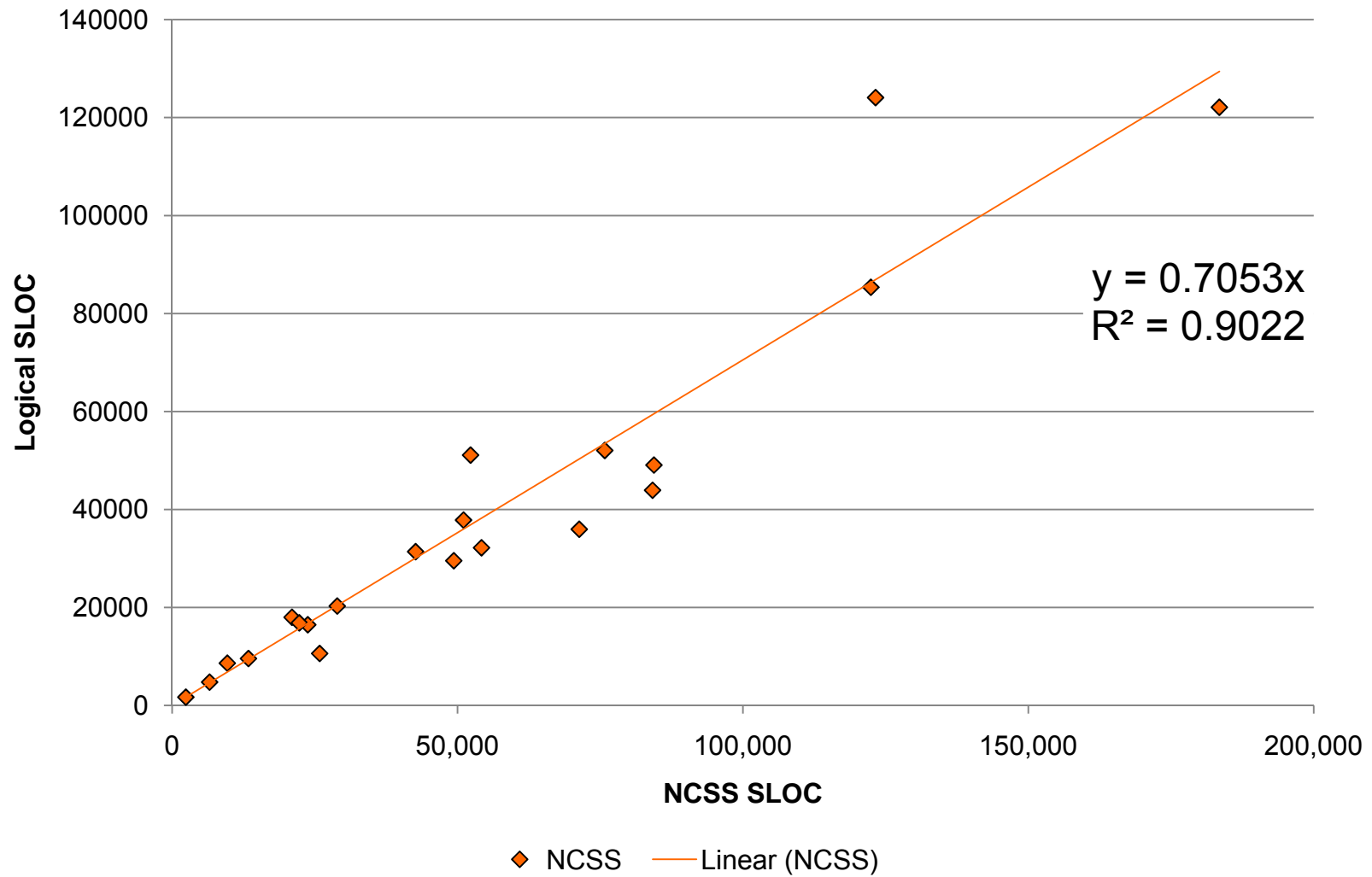
Size Data Conditioning -1

- For analysis, the definition of a source line of code needs to be as consistent as possible to eliminate noise in the data
 - A logical source line of code has been selected as the baseline SLOC definition
- If a source line of code has been defined as either Physical or Non-Commented Source Statements (NCSS), these counts need to be converted to a logical SLOC
 - Physical: a line in a file, e.g. carriage returns including blanks and comment lines
 - NCSS: a line in a file that is not a blank or comment line

Physical to Logical

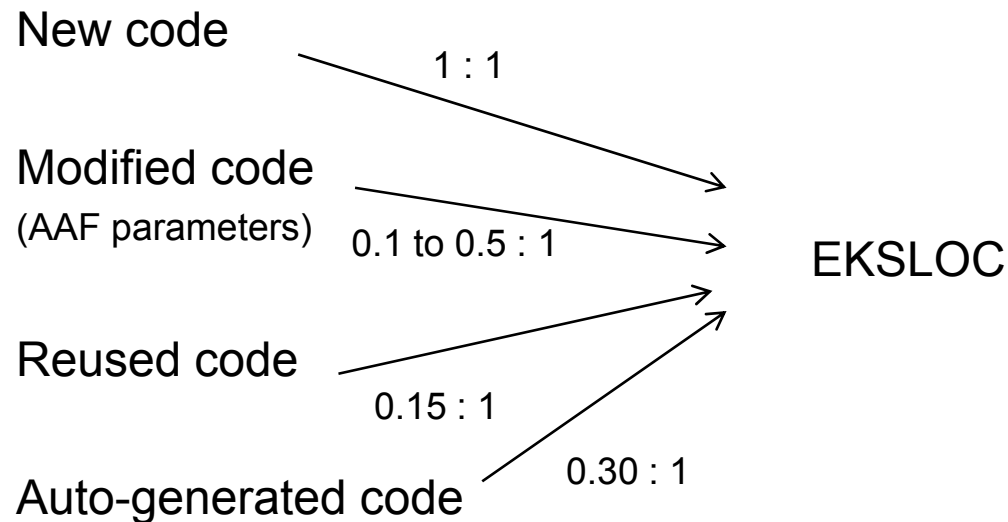


NCSS to Logical



Size Data Conditioning -2

- Normalize sizing data
 - Different approaches to counting lines of code
 - Create “Equivalent” KSLOC (EKSLOC)



AAF Parameters

- Communication Domain (18 observations)

	DM%	CM%	IM%
Median	15	28	64
Low 90% CL	14	20	49
Mean	25	31	62
High 90% CL	36	42	75

- Mission Management (19 observations)

	DM%	CM%	IM%
Median	100	100	100
Low 90% CL	58	69	76
Mean	75	83	88
High 90% CL	92	97	100

Missing Effort Data

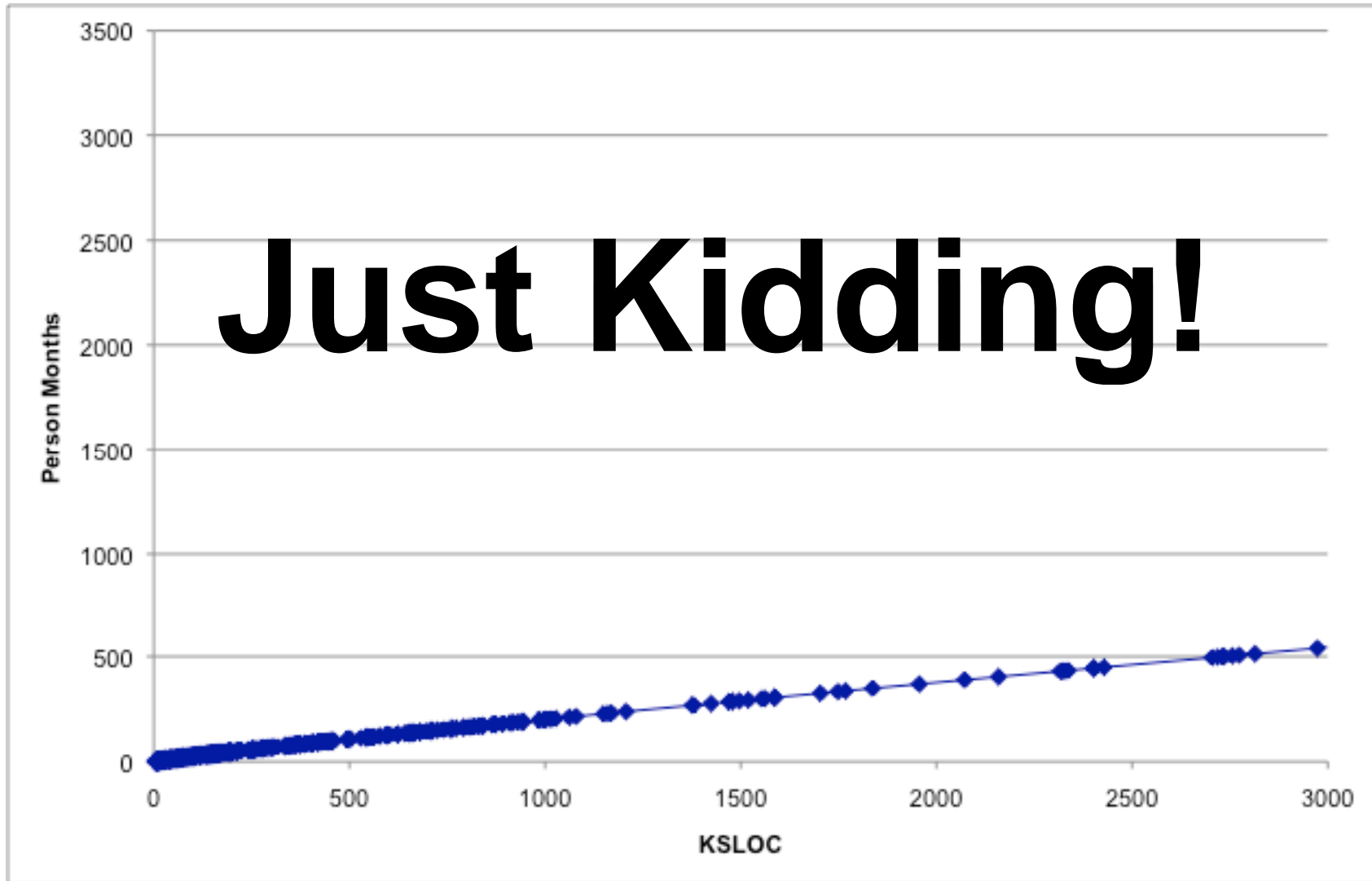
- Communication Domain (27 observations)

	Req't%	Arch%	Code%	I&T%	QT%
Median	16	27	32	21	4
Low 90% CL	14	23	29	17	4
Mean	17	27	32	20	7
High 90% CL	20	30	35	23	10

- Mission Management Domain (16 observations)

	Req't%	Arch%	Code%	I&T%	QT%
Median	24	13	34	17	11
Low 90% CL	18	11	27	12	6
Mean	24	14	32	17	13
High 90% CL	30	19	37	22	20

Data Conditioning Results



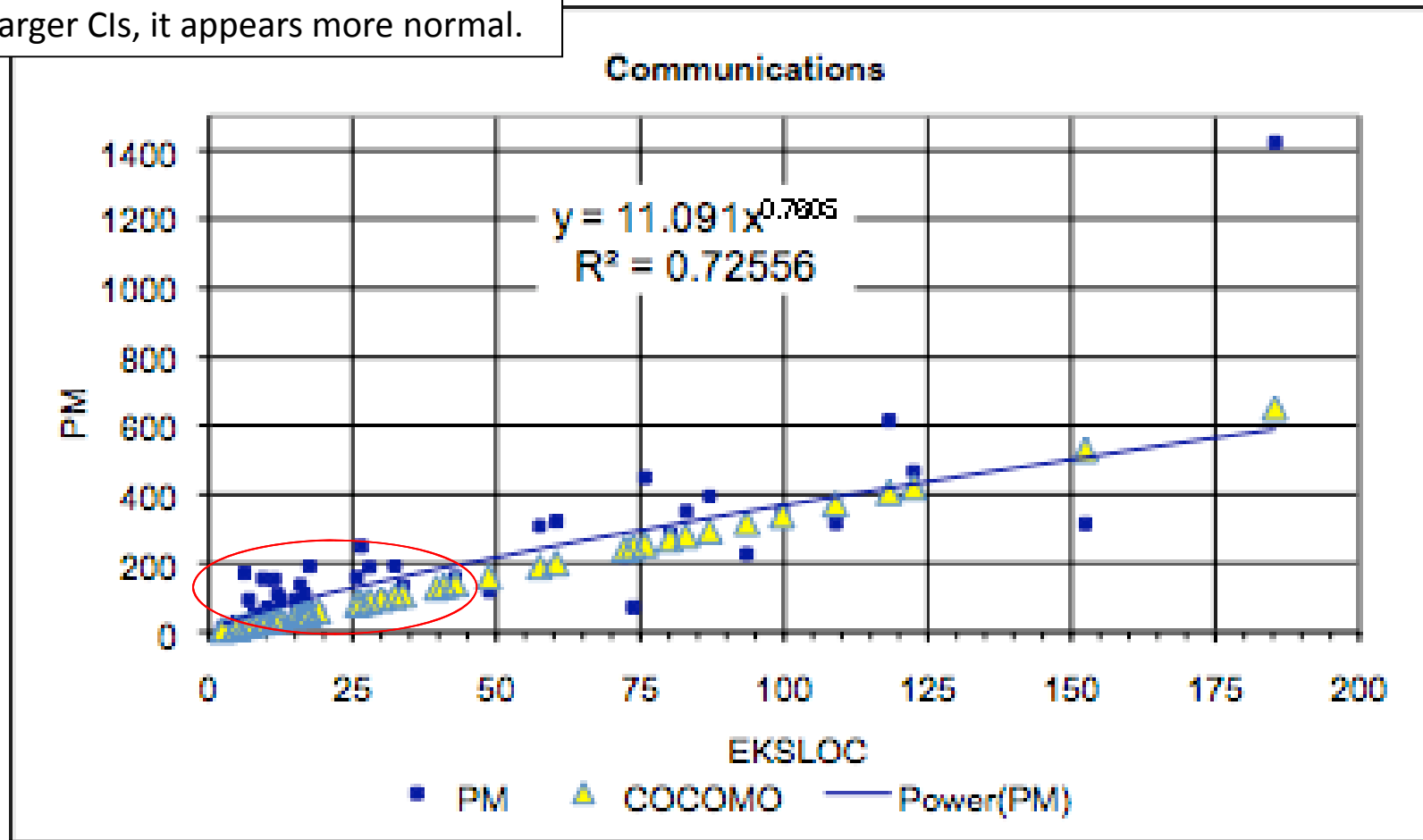


Agenda

- Background
- Software Resources Data Report Issues
- Data Conditioning Solutions
- • Communication Domain Preliminary Analysis

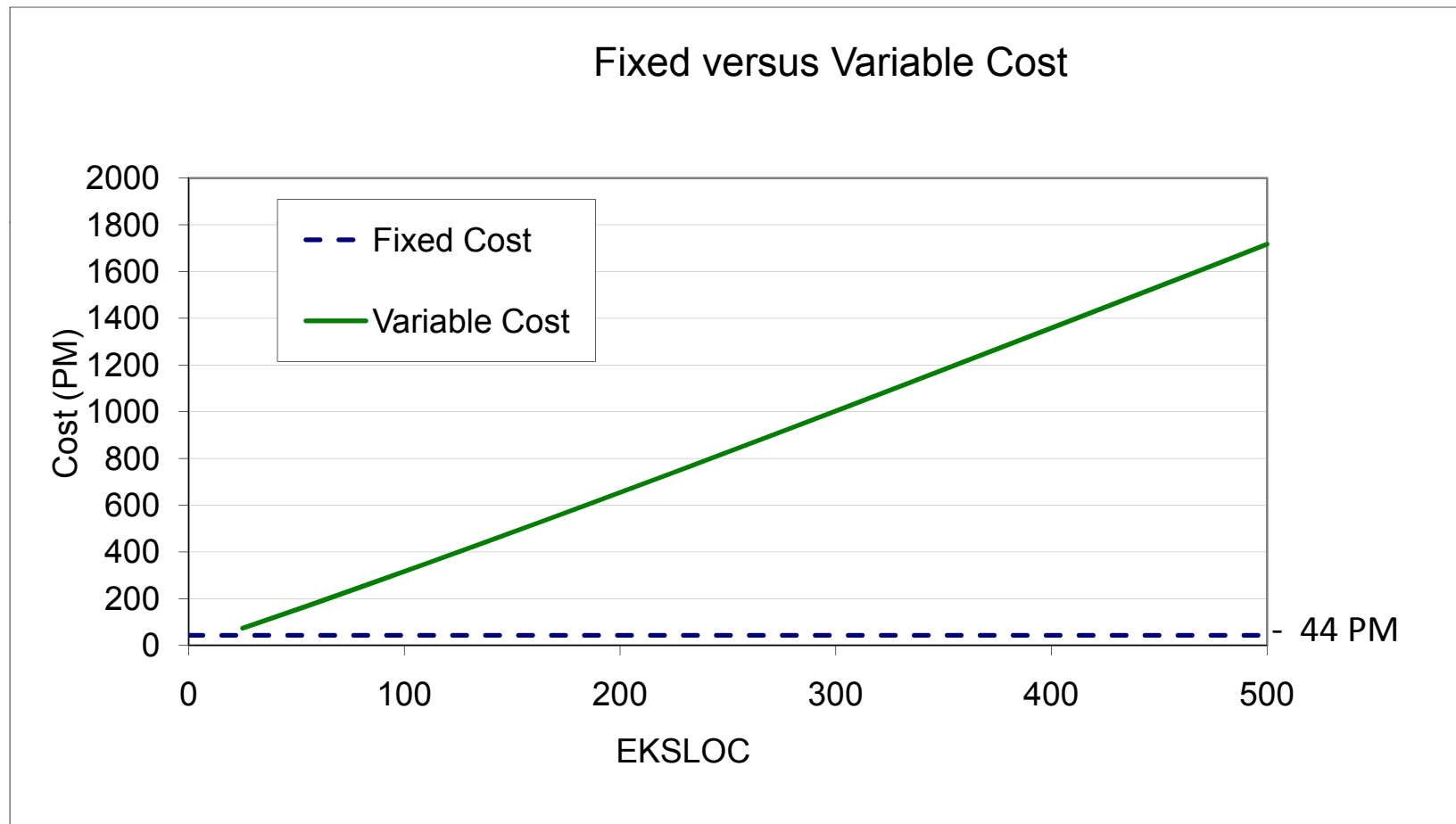
Size - Effort

Effort for small software configuration items (CI) appears high.
 For larger CIs, it appears more normal.



Two Cost Functions

The data shows a combination of **Fixed** and **Variable** costs



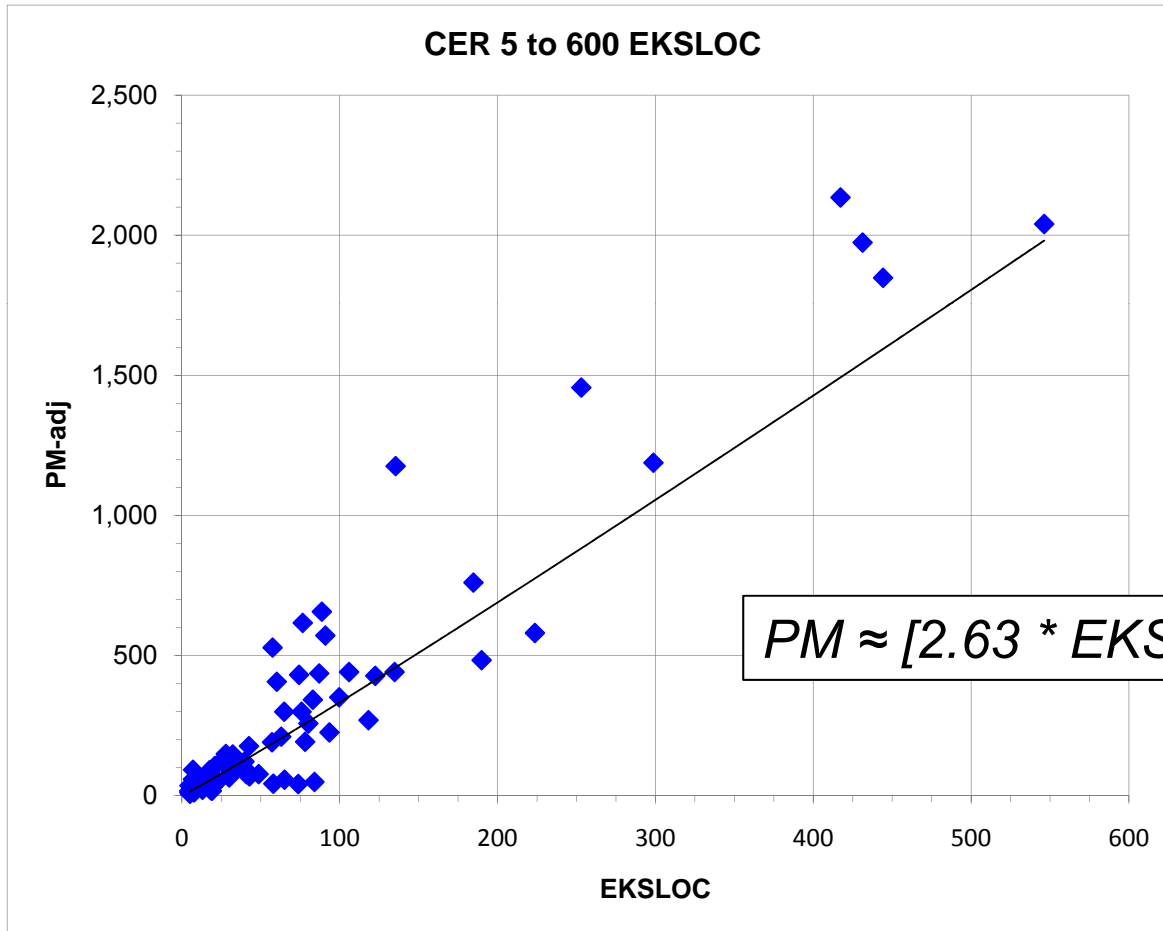


Possible Sources of Fixed Costs

- SEIT/PM* costs reported with the items they are supporting:
 - Project management
 - Business management (operations, scheduling, accounting & finance)
 - Project control & planning
 - Configuration management
 - Quality assurance
 - Contract management
 - Security management
 - Delivered data
 - Technical publications and technical manuals
 - Cost and schedule reporting
 - Training plans, manuals, guides
 - Data management: Engineering & Management

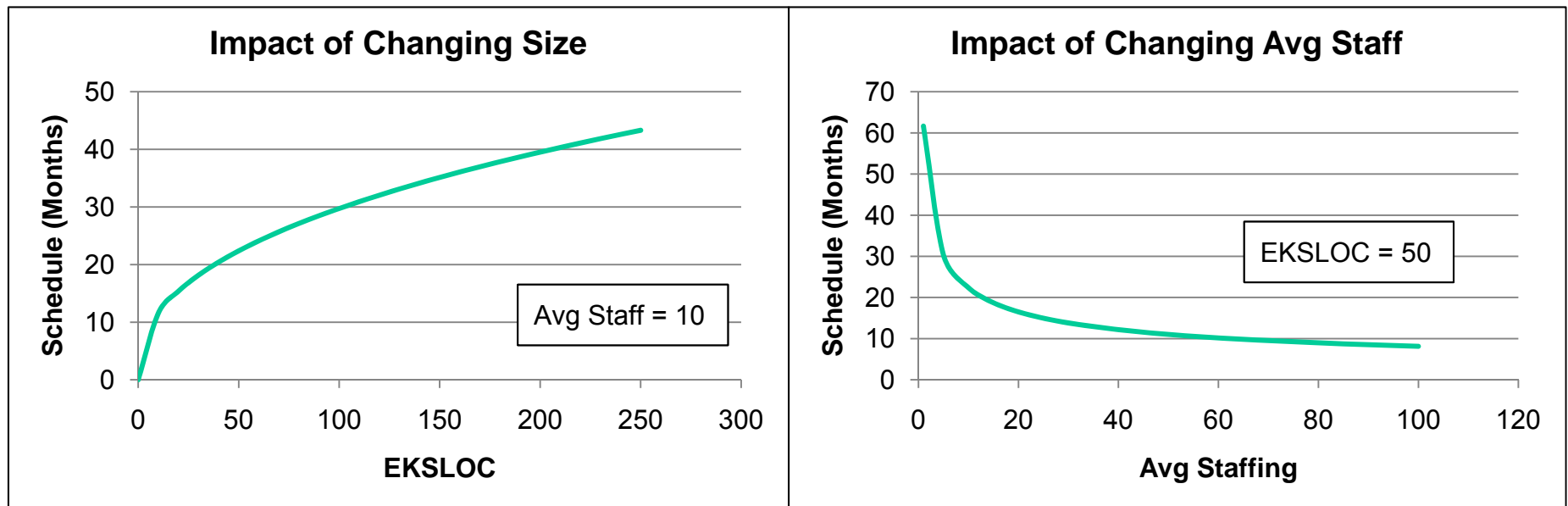
* SEIT/PM: Systems Engineering, Integration, Test and Program Management

Cost Estimating Relationship (CER)

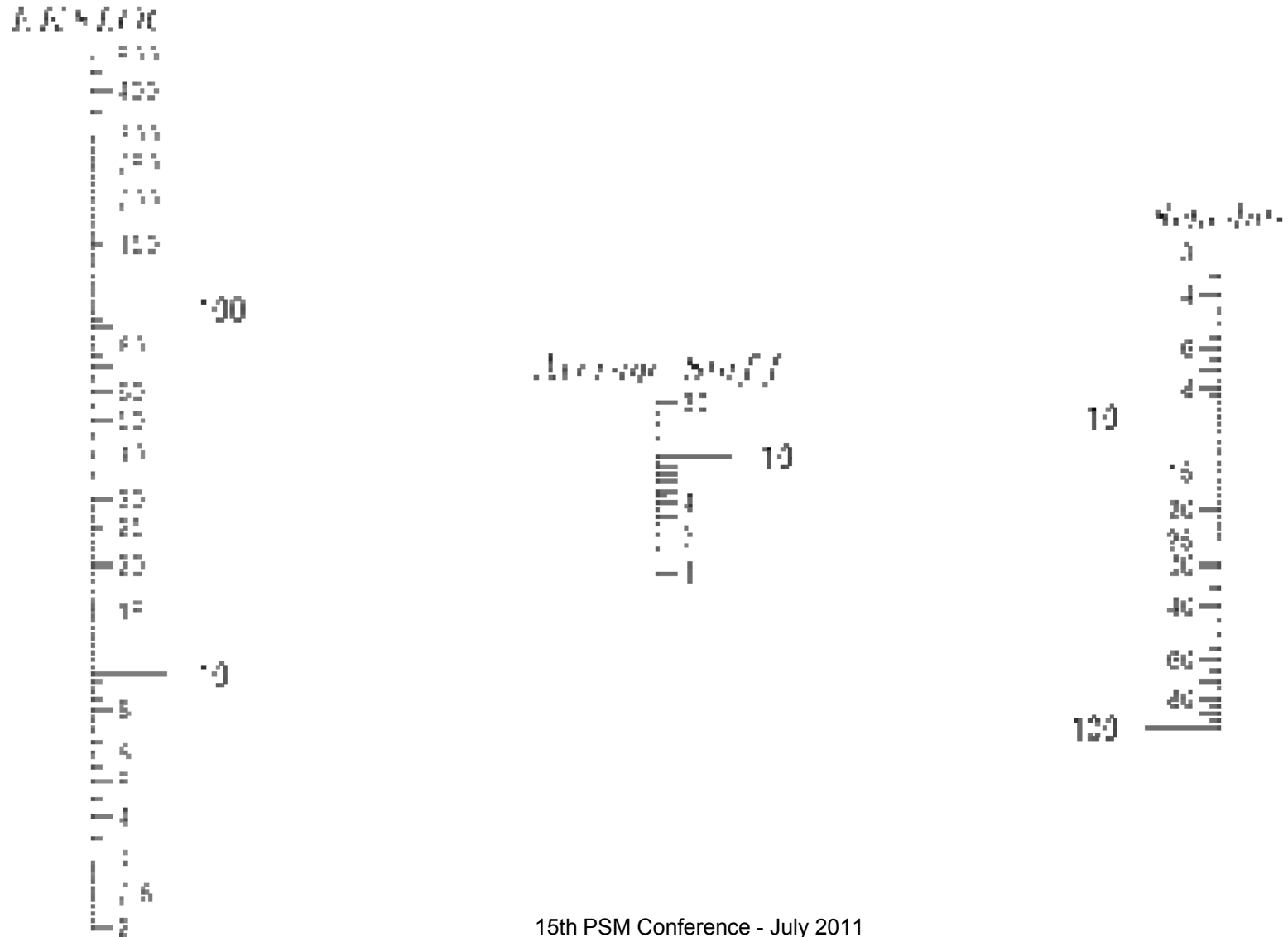


Schedule Estimating Relationship (SER)

$$\text{Schedule} \approx 12.4 * \text{EKSLOC}^{0.41} * \text{AvgStaff}^{-0.42}$$



SER Nomograph





Project Status

- Revise Application Domain and Operating Environment definitions
 - Mil-Std 881C
- Reclassify data
- Data analysis for different application domains
- Publish draft of results
 - http://csse.usc.edu/afcaa/manual_draft/



Questions?

For more information, contact:

Wilson Rosa

Wilson.Rosa@pentagon.af.mil

240-612-5593

Or

Brad Clark

bkclark@csse.usc.edu

703-754-0115