Firehouse Is Underrated: Continued Analysis Of Software Cost Estimation Models Using Normalized, Stratified Data

Dan Strickland
Software SME, Research Division

6 - 9 JUN 2017

Overview

- **Software Resources Data Reports (SRDR) Database and Issues**
- **Normalizing the "Good" SRDR Database**
- **New SRDR Database Metrics**
- **Three Models and Productivity**
- **Model Performance**
- **Model Performance with Boundaries**
- **Paired Strata**
- **Conclusions**

SRDR Database

- **Software Resources Data Reports (SRDR) are the DoD's mechanism for collecting data on software projects for cost analysis**
- **SRDRs are collected by the Office of the Secretary of Defense (OSD) – Cost Assessment and Program Evaluation (CAPE) Organization from government contractors at the beginning and end of software projects**
- **SRDRs contain data like size in Source Lines of Code (SLOC), contract type, hours expended per development phase, and application type of the software**
- **New effort in DoD to normalize and standardize the SRDR further via a new Data Item Description (DID), signed early 2016**
- **Multiple efforts in DoD to collect, normalize, and clean the data:**
  - Naval Air Systems Command (NAVAIR) – initial SRDR Excel database with pairing and "goodness" of data
  - Air Force Cost Analysis Agency (AFCAA), Department of Homeland Security (DHS), Naval Center for Cost Analysis (NCCA) – regression of "good" database with Application Domains and metadata
  - NCCA – matching of "good" database to new DID Application Domain and regressions
- **Many papers and research projects on Cost Estimating Relationships (CERs) and data analysis of the SRDR databases**

SRDR Problems

- Matching Data – many reports have either and Initial or Final Report, but not both
- Data Correctness – many reports have obvious data problems or issues that can't be corrected without finding originators
- Data Fidelity Level – many reports are rolled up to a "system software" level rather than a "software program" level - Computer Software Configuration Item (CSCI) preferred

Establish a set of ground rules on the SRDR data to normalize it in a defined fashion

SRDR Database Ground Rules and Assumptions

- **Starting with the NCCA database of APR 2015 (2624 records)**
  - Final SRDRs only (888 records)
  - “Good” Quality Tag only (401 records)
- **All data items should be of component or CSCI “size” in ESLOC:**
  - Equivalent SLOC (ESLOC) = New + 50% (Modified) + 5% (Reuse) + 30% (AutoGen)
  - CSCI size is greater than 5K ESLOC, less than 200K ESLOC (same as Aerospace study)
  - Reduces database size to 321 records
- **All data items should have defined hours for Software Design, Code, and Test & Integration (DCTI)**
  - **Architecture/Design** hours are SW Design hours
  - **Code and Unit Test** hours are SW Code hours
  - **SW and System Integration, SW Qualification Testing** hours are SW Test and Integration hours
  - **Requirements Analysis** and **SW Developmental Test and Evaluation (DT&E)** hours are **not** part of DCTI hours
  - **Other** hours are distributed proportionally across all active phases
  - Reduces database size to 282 records
- **Duration calculated in months**
  - Maximum Date (DCTI) – Minimum Date (DCTI)

SRDR Database Strata Sample

Application Domains	
Command & Control	Software that allows humans to manage a dynamic situation and respond in real time
Communications	The transmission of information, e.g. voice, data, commands, images, and video across different mediums and distances
Custom AIS	Software needed to build a custom software application to fill a capability gap not captured by COTS/GOTS software packages
Mission Planning	Supports all the mission requirements of the platform and may have the capability to program onboard platform systems with routing, targeting, performance, map, and Intel data
Real-Time Embedded	Interrupt-driven, embedded software in military and consumer appliances, devices, and products, possibly directing and processing sensor inputs/outputs, generally with a very small executive for an operating system interface to basic processor(s).
Scientific/Simulation	Non real time software that involves significant computations and scientific analysis
Signal Processing	Software that requires timing-dependent device coding to enhance, transform, filter, convert, or compress data signals
SW Tools	Software that is used for analysis, design, construction, or testing of computer programs
Systems Software	Layers of software that sit between the computing platform and applications
Test, Measurement, and Diagnostic Equipment	Software used for testing, measuring, diagnosing, emulating, and evaluating operational hardware and software systems
Training	Hardware and software that are used for educational and training purposes
Vehicle Control	Software necessary for the control of vehicle primary and secondary mechanical devices and surfaces
Vehicle Payload	Software which controls and monitors vehicle payloads and provides communications to other vehicle subsystems and payloads

Operating Environments
Air Vehicle, Manned
Air Vehicle, Unmanned
Ordinance System, Unmanned
Sea System, Manned
Sea System, Unmanned
Surface Fixed, Manned
Surface Mobile, Manned
Surface Vehicle, Manned
Surface Vehicle, Unmanned

Development Language
C, C++
Ada
Java
Other

New SRDR Database Metrics

Data Points	
Initial SRDR Database	401
Total With 5K <= ESLOC <= 200K (CSCI size)	321
Has Design, Code, Test Populated	282

Effort Distribution (DCTI Only)	
Design	24%
Code	40%
Test	36%

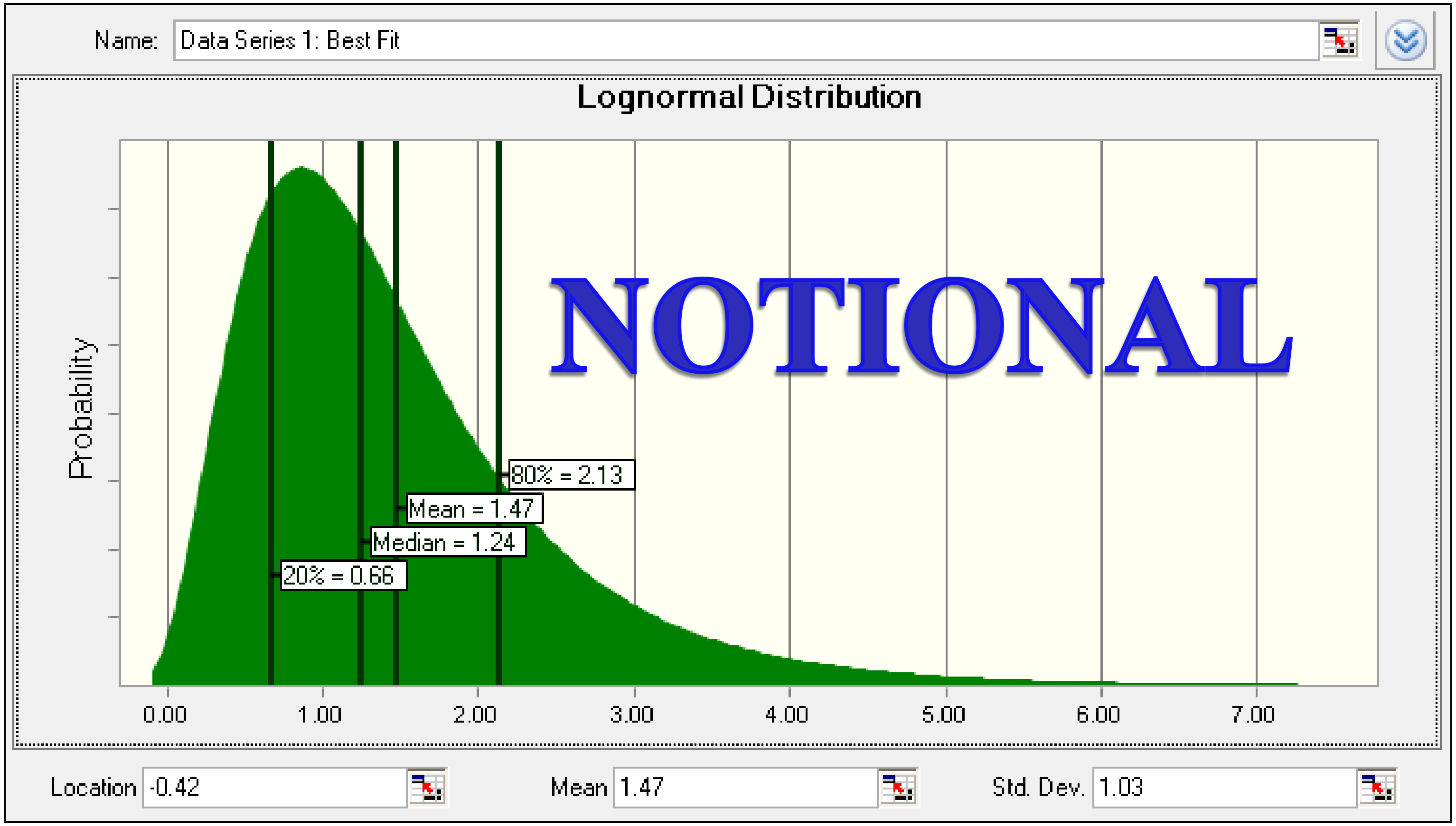
Additional Phases	
Records With Requirements Analysis Hours	263
Records With DT&E Hours	134
Requirements Analysis % (tax)	18%
DT&E % (tax)	32%

		Productivity (ESLOC/hr)*	
Domain	Records	Mean	Median
All	282	1.90	1.29
Command/Control	39	1.46	1.25
Communications	37	1.6	1.56
Custom AIS Software	12	2.82	2.46
Mission Planning	17	2.28	2.17
Real-Time Embedded	59	1.17	1.12
Roll-Up**	1	3.99	3.99
Scientific/Simulation	12	2.93	2.05
Signal Processing	22	1.14	0.61
SW Tools	6	7.72	5.03
Systems Software	35	2.57	1.33
Test/Measurement/Diagnostics	4	0.65	0.55
Training	2	7.39	7.39
Vehicle Control	22	1.84	1.15
Vehicle Payload	14	1.41	0.82

* Productivity is of DCTI with Other hours only
** Roll-Up is not an official SRDR Application Domain and this data point was removed from analysis

SRDR Database Productivity Distribution

Example: Command/Control Productivity Curve Fit



Software Estimating Models

- **In the Feb 2006 *CrossTalk* article "Software Estimating Models: Three Viewpoints", three popular software cost estimation models (Sage/SEER-SEM, SLIM-Estimate, COCOMO II) are described in their base mathematical forms**
- **All three models calculate effort using size and productivity**
- **Two models (SLIM-Estimate, SEER-SEM) also use development time as a factor to calculate**
- **Productivity is expressed as software output over software input, usually in SLOC/hr or SLOC/PM; in calibration, estimators are looking for productivity**
- **SLIM-Estimate model on calculating productivity: "From historic projects, we know the size, effort, and schedule…just put in a consistent set of historic numbers and calculate a Process Productivity Parameter."**
- **Is it really that simple to calculate Productivity?**

Given a database of completed projects and the three default models, can reliable productivity ranges be developed in different development strata?

SLIM-Estimate Model

- **Basic Equation: Size = (Effort / βeta)$^{1/3}$ (Schedule)$^{4/3}$ PP**
  - Size in SLOC
  - Effort is software development effort in Person-Years
  - Beta is a skills factor based on size and ranging from 0.16 to 0.39
  - Schedule is the development time in years
  - PP is the Process Productivity Parameter and has been observed to range from **1,974 to 121,393**
- **Solve for Productivity: PP = Size / ((Effort / βeta)$^{1/3}$ (Schedule)$^{4/3}$))**
  - As PP increases, effort (cost) decreases
- **Using the Normalized SRDR database, solve for PP and observe stratified results**
- **Can a Productivity value be developed that produces accurate results?**
  - Effort (PY) = 15 * βeta* $(td_{min})^3$ (SLIM-Estimate's effort equation)
  - td_{min}(years) = 0.68 * (Size / PP)$^{0.43}$ (SLIM-Estimate's minimum development time equation)

Operating Environment	Application Domai	Primary Langua	ESLOC (MD	Development H	Dev Mo
Surface Fixed, Manned	Mission Planning	Visual Basic	230506	44885	36

SLOC = ESLOC

Effort (PY) = Development Hrs / 152 / 12

Schedule (yrs) = Develop Months / 12

SLIM-Estimate Model – Results

SRDR Database			Productivity Parameter				
Op Env	**App Type**	**Language**	**Records**	**Min**	**Max**	**Mean**	**Median**
All	All	All	281	103.0	76992.5	6779.0	3413.3
All	Cmd/Ctrl	All	39	134.3	68096.0	12331.7	7040.4
All	Communications	All	37	151.6	35315.7	6724.4	3797.3
All	Custom AIS	All	12	741.7	24361.1	9769.1	10207.9
All	Mission Planning	All	17	1826.2	24667.4	9228.8	9626.6
All	Real-Time Embedded	All	59	146.4	20319.1	3435.9	2161.7
All	Scientific/Simulation	All	12	1700.7	76992.5	13960.0	4812.9
All	Signal Processing	All	22	271.1	18173.3	3093.1	1514.3
All	SW Tools	All	6	3413.3	29550.3	11993.7	7905.3
All	Systems Software	All	35	103.0	33784.9	6246.1	3359.6
All	Test/Meas/Diag Equip	All	4	717.9	2080.3	1155.1	911.2
All	Training	All	2	2993.8	6262.0	4627.9	4627.9
All	Vehicle Control	All	22	315.3	11039.9	4420.2	3844.6
All	Vehicle Payload	All	14	316.1	8300.8	4360.8	4296.2
All	All	C/C++	183	103.0	76992.5	6506.7	3262.3
All	All	Ada	53	169.8	68096.0	4274.6	2065.2
All	All	Java	39	709.4	49888.8	11633.2	6928.2
All	All	Other	6	1045.7	15172.9	5652.3	4497.9
Air Veh, Manned	All	All	53	205.5	68096.0	5807.1	2790.8
Air Veh, Unmanned	All	All	21	1208.4	24236.7	6746.2	4907.1
Ord Sys, Unmanned	All	All	27	397.3	45003.1	9414.7	6161.9
Sea Sys, Manned	All	All	28	151.6	16118.2	1954.8	1388.4
Sea Sys, Unmanned	All	All	2	1684.4	8300.8	4992.6	4992.6
Surface Fixed, Man	All	All	100	134.3	76992.5	9506.4	5264.9
Surface Mobile, Manned	All	All	12	169.8	3726.6	1449.1	1231.8
Surface Vehicle, Manned	All	All	34	103.0	33784.9	4710.9	2631.1
Surface Vehicle, Unmanned	All	All	4	775.8	3590.7	2079.9	1976.6

Example (Mission Plan, Median PP):
Effort (PY) = 15 * β * $(td_{min})^3$
td_{min}(years) = 0.68 * $(Size / 9626.6)^{0.43}$

Several PP values below the minimum observed values by SLIM-Estimate

SLIM-Estimate Model – Accuracy

				Mean PP		Median PP	
Op Env	**App Type**	**Language**	**Records**	**MMRE**	**PRED(30)**	**MMRE**	**PRED(30)**
All	All	All	281	1.52	20.6%	4.00	13.2%
All	Cmd/Ctrl	All	39	0.89	15.4%	1.62	25.6%
All	Communications	All	37	1.09	27.0%	2.48	24.3%
All	Custom AIS	All	12	0.85	16.7%	0.77	16.7%
All	Mission Planning	All	17	2.13	11.8%	1.99	23.5%
All	Real-Time Embedded	All	59	1.94	18.6%	3.87	11.9%
All	Scientific/Simulation	All	12	1.11	16.7%	5.04	0.0%
All	Signal Processing	All	22	2.33	18.2%	6.35	22.7%
All	SW Tools	All	6	2.57	50.0%	4.86	16.7%
All	Systems Software	All	35	1.60	17.1%	3.71	14.3%
All	Test/Meas/Diag Equip	All	4	4.18	0.0%	6.03	0.0%
All	Training	All	2	11.00	0.0%	11.00	0.0%
All	Vehicle Control	All	22	1.99	18.2%	2.46	9.1%
All	Vehicle Payload	All	14	1.26	28.6%	1.30	21.4%
All	All	C/C++	183	1.65	18.7%	4.31	12.6%
All	All	Ada	53	2.60	11.3%	7.06	20.8%
All	All	Java	39	0.72	15.4%	1.20	25.6%
All	All	Other	6	3.28	16.7%	4.57	16.7%
Air Veh, Manned	All	All	53	2.04	15.1%	5.69	18.9%
Air Veh, Unmanned	All	All	21	1.31	14.3%	2.08	19.0%
Ord Sys, Unmanned	All	All	27	1.82	18.5%	3.24	25.9%
Sea Sys, Manned	All	All	28	3.51	7.1%	5.89	3.6%
Sea Sys, Unmanned	All	All	2	3.95	0.0%	3.95	0.0%
Surface Fixed, Man	All	All	100	1.02	22.0%	2.32	15.0%
Surface Mobile, Manned	All	All	12	4.47	16.7%	5.67	8.3%
Surface Vehicle, Manned	All	All	34	2.00	20.6%	4.45	23.5%
Surface Vehicle, Unmanned	All	All	4	0.92	25.0%	1.02	25.0%

- **Comparing the estimate vs. actual from SRDR data using SLIM-Estimate formula**
- **Mean Magnitude of Relative Error (MMRE) measures the average relative error of all the predictions to their actuals, independent of scale and sign – lower is better**
- **Prediction Level (PRED) measures the percentage of all the predictions that fall within a defined error bounds of the actual, here we used PRED(30) or within 30% – higher is better**
- **Overall, the SLIM-Estimate model using either the mean or median productivity values is not accurate**
- **What about those data points that had calculated PP outside of the published limits for SLIM-Estimate? What happens when those data points are removed?**

SLIM-Estimate Model – In-Bounds Results

SRDR Database	In-Bounds		Productivity Parameter				
Op Env	**App Type**	**Language**	**Records**	**Min**	**Max**	**Mean**	**Median**
All	All	All	188	1989.4	76992.5	9632.1	6045.8
All	Cmd/Ctrl	All	29	1997.0	68096.0	16232.5	13720.5
All	Communications	All	21	2366.7	35315.7	10973.5	6235.1
All	Custom AIS	All	11	2206.4	24361.1	10589.8	10970.9
All	Mission Planning	All	16	2758.9	24667.4	9691.5	10081.8
All	Real-Time Embedd	All	32	1989.4	20319.1	5420.4	3926.9
All	Scientific/Simulat	All	11	2325.9	76992.5	15074.5	5884.0
All	Signal Processing	All	10	2005.3	18173.3	5815.0	3137.6
All	SW Tools	All	6	3413.3	29550.3	11993.7	7905.3
All	Systems Software	All	21	2082.2	33784.9	9830.9	4254.5
All	Test/Meas/Diag Eq	All	1	2080.3	2080.3	2080.3	2080.3
All	Training	All	2	2993.8	6262.0	4627.9	4627.9
All	Vehicle Control	All	17	2118.6	11039.9	5450.6	5441.1
All	Vehicle Payload	All	11	2065.2	8300.8	5296.3	6161.9
All	All	C/C++	123	1989.4	76992.5	9173.7	6161.9
All	All	Ada	27	2065.2	68096.0	7553.7	4540.3
All	All	Java	33	2366.7	49888.8	13504.2	8736.8
All	All	Other	5	3359.6	15172.9	6573.6	4926.9
Air Veh, Manned	All	All	37	1997.0	68096.0	7894.4	5339.8
Air Veh, Unmanned	All	All	20	2005.3	24236.7	7023.1	4917.0
Ord Sys, Unmanned	All	All	20	3146.4	45003.1	12347.8	7034.3
Sea Sys, Manned	All	All	5	2364.9	16118.2	5636.8	2742.1
Sea Sys, Unmanned	All	All	1	8300.8	8300.8	8300.8	8300.8
Surface Fixed, Man	All	All	78	1989.4	76992.5	11876.6	8675.1
Surface Mobile, Manned	All	All	4	2080.3	3726.6	2590.1	2276.8
Surface Vehicle, Manned	All	All	21	2118.6	33784.9	7250.2	5616.3
Surface Vehicle, Unmanned	All	All	2	2252.5	3590.7	2921.6	2921.6

Removed 93 data points with calculated PP values below threshold aka "out of bounds"

SLIM-Estimate Model – In-Bounds Accuracy

				Mean PP			Median PP		
Op Env	App Type	Language	Records	MMRE	PRED(30)	Delta	MMRE	PRED(30)	Delta
All	All	All	188	1.17	28.2%	7.6% ▲	2.33	20.2%	7.0% ▲
All	Cmd/Ctrl	All	29	0.74	20.7%	5.3% ▲	0.84	17.2%	-8.4% ▼
All	Communications	All	21	0.80	28.6%	1.5% ▲	1.78	9.5%	-14.8% ▼
All	Custom AIS	All	11	0.71	18.2%	1.5% ▲	0.65	18.2%	1.5% ▲
All	Mission Planning	All	16	2.08	18.8%	7.0% ▲	1.96	18.8%	-4.8% ▼
All	Real-Time Embedded	All	32	1.28	12.5%	-6.1% ▼	2.33	9.4%	-2.5% ▼
All	Scientific/Simulation	All	11	1.03	18.2%	1.5% ▲	4.01	0.0%	0.0% ▼
All	Signal Processing	All	10	1.74	40.0%	21.8% ▲	4.47	10.0%	-12.7% ▼
All	SW Tools	All	6	2.57	50.0%	0.0% ▼	4.86	16.7%	0.0% ▼
All	Systems Software	All	21	1.23	14.3%	-2.9% ▼	3.92	23.8%	9.5% ▲
All	Test/Meas/Diag Equip	All	1	3.92	0.0%	0.0% ▼	3.92	0.0%	0.0% ▼
All	Training	All	2	11.00	0.0%	0.0% ▼	11.00	0.0%	0.0% ▼
All	Vehicle Control	All	17	1.72	29.4%	11.2% ▲	1.73	29.4%	20.3% ▲
All	Vehicle Payload	All	11	1.03	36.4%	7.8% ▲	0.83	36.4%	14.9% ▲
All	All	C/C++	123	1.28	28.5%	9.8% ▲	2.34	17.9%	5.2% ▲
All	All	Ada	27	1.83	44.4%	33.1% ▲	4.11	14.8%	-5.9% ▼
All	All	Java	33	0.63	21.2%	5.8% ▲	0.93	30.3%	4.7% ▲
All	All	Other	5	3.05	20.0%	3.3% ▲	4.70	20.0%	3.3% ▲
Air Veh, Manned	All	All	37	1.62	37.8%	22.7% ▲	2.99	18.9%	0.1% ▲
Air Veh, Unmanned	All	All	20	1.26	15.0%	0.7% ▲	2.14	20.0%	1.0% ▲
Ord Sys, Unmanned	All	All	20	1.50	25.0%	6.5% ▲	3.31	35.0%	9.1% ▲
Sea Sys, Manned	All	All	5	1.53	20.0%	12.9% ▲	5.17	0.0%	-3.6% ▼
Sea Sys, Unmanned	All	All	1	3.23	0.0%	0.0% ▼	3.23	0.0%	0.0% ▼
Surface Fixed, Man	All	All	78	0.83	28.2%	6.2% ▲	1.26	25.6%	10.6% ▲
Surface Mobile, Manned	All	All	4	4.39	0.0%	-16.7% ▼	5.36	0.0%	-8.3% ▼
Surface Vehicle, Manned	All	All	21	1.54	19.0%	-1.5% ▼	2.23	14.3%	-9.2% ▼
Surface Vehicle, Unmanned	All	All	2	0.69	50.0%	25.0% ▲	0.69	50.0%	25.0% ▲

- **Removing the data points with "out of bounds" Productivity calculations improves the accuracy, especially using the mean PP**
- **Increases in accuracy are meager and overall, accuracy is poor**

SEER-SEM Model

- **Basic Equation: C_{te} = (Size)/(K)$^{1/2}$ Schedule**
  - Size in ESLOC
  - K is software life-cycle effort in person-years
  - Schedule is the development time in years
  - C_{te} is the Effective Technology Constant ranges from **2.7 to 22,184.1**
  - C_{te} increases, effort (cost) decreases
  - Software development effort is 0.3945 of the total life-cycle effort (K)
- **Using the SRDR database, solve for C_{te} and observe stratified results**
- **Can a Productivity value be developed that produces accurate results?**
  - Effort (PM) = ((Size / (C_{te}*Schedule))^2 (0.3945) * 12

Operating Environment	Application Domai	Primary Langua	ESLOC (MD	Development H	Dev Mo
Surface Fixed, Manned	Mission Planning	Visual Basic	230506	44885	36

ESLOC = ESLOC

K (PY) = Development Hours/(152)(12)(0.3945)

Schedule (yrs) = Develop Months / 12

SEER-SEM Model – Results

			Cte (IB)				
Op Env	**App Type**	**Language**	**Records**	**Min**	**Max**	**Mean**	**Median**
All	All	All	274	127.27	18217.54	4042.54	2747.26
All	Cmd/Ctrl	All	36	166.65	14790.35	5323.79	4988.58
All	Communications	All	37	199.32	18217.54	4643.22	2943.35
All	Custom AIS	All	12	1132.10	13966.77	6642.88	7159.86
All	Mission Planning	All	17	1618.57	17008.37	6771.15	5381.59
All	Real-Time Embedded	All	59	176.32	11466.70	2565.30	1811.96
All	Scientific/Simulation	All	10	1696.09	7414.01	3810.53	2825.08
All	Signal Processing	All	22	297.16	11727.20	2325.44	1286.61
All	SW Tools	All	5	3618.52	13677.00	7352.00	4825.08
All	Systems Software	All	34	127.27	17340.64	4094.34	2534.38
All	Test/Meas/Diag Equip	All	4	658.04	1586.65	973.67	824.99
All	Training	All	2	3629.15	7226.20	5427.67	5427.67
All	Vehicle Control	All	22	493.80	8678.98	3450.97	3318.18
All	Vehicle Payload	All	14	525.79	6681.74	3008.60	2939.25
All	All	C/C++	176	127.27	17340.64	4043.72	2621.58
All	All	Ada	52	193.04	11460.88	2375.12	1636.28
All	All	Java	38	733.99	18217.54	6340.03	5013.16
All	All	Other	6	935.96	11030.57	4726.72	3966.44
Air Veh, Manned	All	All	52	273.74	17340.64	3689.63	2660.97
Air Veh, Unmanned	All	All	21	874.26	16668.81	4988.23	3408.88
Ord Sys, Unmanned	All	All	24	519.46	11727.20	3949.80	3329.60
Sea Sys, Manned	All	All	28	199.32	11792.59	1942.37	1527.83
Sea Sys, Unmanned	All	All	2	1621.83	6681.74	4151.78	4151.78
Surface Fixed, Man	All	All	98	166.65	18217.54	5466.06	3991.24
Surface Mobile, Manned	All	All	12	209.11	2895.29	1224.68	1114.07
Surface Vehicle, Manned	All	All	33	127.27	11466.70	2903.52	2466.62
Surface Vehicle, Unmanned	All	All	4	821.67	2827.99	1842.80	1860.78

Example (Mission Plan, Median C_{te}):
Effort (PM) = ((Size / (5381.6 * Schedule))^2* (0.3945) * 12

Seven calculated C_{te} values were over 22,184 - "out of bounds" with model thresholds and removed

SEER-SEM Model – Accuracy

				Mean Cte		Median Cte	
Op Env	**App Type**	**Language**	**Records**	**MMRE**	**PRED(30)**	**MMRE**	**PRED(30)**
All	All	All	274	1.82	10.6%	3.72	13.1%
All	Cmd/Ctrl	All	36	1.43	16.7%	1.62	11.1%
All	Communications	All	37	1.97	18.9%	4.66	13.5%
All	Custom AIS	All	12	0.83	25.0%	0.70	33.3%
All	Mission Planning	All	17	1.20	0.0%	1.86	5.9%
All	Real-Time Embedded	All	59	1.55	11.9%	2.93	23.7%
All	Scientific/Simulation	All	10	0.96	0.0%	1.59	40.0%
All	Signal Processing	All	22	2.73	18.2%	8.15	0.0%
All	SW Tools	All	5	1.12	0.0%	2.33	20.0%
All	Systems Software	All	34	2.38	8.8%	5.69	17.6%
All	Test/Meas/Diag Equip	All	4	0.69	25.0%	0.81	50.0%
All	Training	All	2	0.66	0.0%	0.66	0.0%
All	Vehicle Control	All	22	1.02	18.2%	1.09	13.6%
All	Vehicle Payload	All	14	0.80	21.4%	0.84	21.4%
All	All	C/C++	177	1.78	9.0%	4.01	11.3%
All	All	Ada	52	1.78	5.8%	3.56	15.4%
All	All	Java	38	1.40	10.5%	2.13	15.8%
All	All	Other	6	1.07	33.3%	1.39	50.0%
Air Veh, Manned	All	All	52	1.75	17.3%	3.22	13.5%
Air Veh, Unmanned	All	All	21	1.48	19.0%	2.99	14.3%
Ord Sys, Unmanned	All	All	24	1.34	20.8%	1.85	16.7%
Sea Sys, Manned	All	All	28	1.92	28.6%	3.03	17.9%
Sea Sys, Unmanned	All	All	2	1.22	0.0%	1.22	0.0%
Surface Fixed, Man	All	All	98	1.53	13.3%	2.77	9.2%
Surface Mobile, Manned	All	All	12	1.06	8.3%	1.28	0.0%
Surface Vehicle, Manned	All	All	33	1.62	15.2%	2.21	12.1%
Surface Vehicle, Unmanned	All	All	4	0.63	50.0%	0.62	50.0%

- **Using the mean or median Cte values, SEER-SEM model performs poorly for most strata**
- **Performs well in the TMDE Application Domain, Java-based developments, and Surface Vehicle, Unmanned**
- **Identify another value at work in the SEER-SEM model that can indicate possible candidates for removal**

SEER-SEM Model – Calculating D

- The Staffing Complexity factor, D, represents the difficulty on terms of the rate at which staff can be added to a software product
- D ranges from 4 to 28, where higher values equate to very complex software that is difficult to staff (missile algorithms) and lower values equate to simple software that can be broken up and staffed easily (data entry)
- D is interactive with the schedule and effort
- Formula: $D = K/(\text{Schedule})^3$
- Calculate D for the database and remove the values that are out of bounds

SEER-SEM Model – In-Bounds Accuracy

				Mean Cte			Median Cte		
Op Env	**App Type**	**Language**	**Records**	**MMRE**	**PRED(30)**	**Delta**	**MMRE**	**PRED(30)**	**Delta**
All	All	All	69	1.12	18.8%	8.3%▲	1.46	20.3%	7.2%▲
All	Cmd/Ctrl	All	9	0.71	22.2%	5.6%▲	0.56	22.2%	11.1%▲
All	Communications	All	8	1.08	25.0%	6.1%▲	1.43	50.0%	36.5%▲
All	Custom AIS	All	7	0.24	57.1%	32.1%▲	0.22	71.4%	38.1%▲
All	Mission Planning	All	6	1.08	16.7%	16.7%▲	1.49	0.0%	-5.9%▼
All	Real-Time Embedded	All	8	1.00	25.0%	13.1%▲	1.14	25.0%	1.3%▲
All	Scientific/Simulation	All	1	0.00	100.0%	100.0%▲	0.00	100.0%	60.0%▲
All	Signal Processing	All	5	2.17	0.0%	-18.2%▼	5.90	60.0%	60.0%▲
All	SW Tools	All	1	0.00	100.0%	100.0%▲	0.00	100.0%	80.0%▲
All	Systems Software	All	11	1.27	36.4%	27.5%▲	1.66	9.1%	-8.6%▼
All	Test/Meas/Diag Equip	All	0						
All	Training	All	0						
All	Vehicle Control	All	7	0.68	28.6%	10.4%▲	0.49	14.3%	0.6%▲
All	Vehicle Payload	All	6	0.32	50.0%	28.6%▲	0.28	50.0%	28.6%▲
All	All	C/C++	51	1.09	19.6%	10.6%▲	1.45	17.6%	6.3%▲
All	All	Ada	7	0.60	28.6%	22.8%▲	0.60	28.6%	13.2%▲
All	All	Java	11	1.08	27.3%	16.7%▲	1.32	36.4%	20.6%▲
All	All	Other	0						
Air Veh, Manned	All	All	12	1.39	0.0%	-17.3%▼	2.83	25.0%	25.0%▲
Air Veh, Unmanned	All	All	7	1.77	14.3%	-4.8%▼	5.50	28.6%	28.6%▲
Ord Sys, Unmanned	All	All	10	0.94	20.0%	-0.8%▼	1.41	20.0%	20.0%▲
Sea Sys, Manned	All	All	2	0.91	0.0%	-28.6%▼	0.91	0.0%	0.0%▼
Sea Sys, Unmanned	All	All	1	0.00	100.0%	100.0%▲	0.00	100.0%	100.0%▲
Surface Fixed, Man	All	All	28	0.87	14.3%	1.0%▲	1.12	21.4%	21.4%▲
Surface Mobile, Manned	All	All	0						
Surface Vehicle, Manned	All	All	8	0.97	12.5%	-2.7%▼	1.56	50.0%	50.0%▲
Surface Vehicle, Unmanned	All	All	1	0.00	100.0%	50.0%▲	0.00	100.0%	100.0%▲

- **Staffing Complexity boundaries remove 205 data points (69 records remaining)**
- **Removing the data points with "out of bounds" Staffing Complexity values improves accuracy**
- **Mass reduction of records eliminates some strata and makes many a single data point**
- **Custom AIS, Signal Processing, and Vehicle Payload strata are examples of most accurate strata**

COCOMO II Model

- **Basic Equation: Effort (PM) = 2.94 * (EAF) (Size)^E**
  - Size in KESLOC
  - EAF is the Effort Adjustment Factor, used to calculate productivity
  - E is the exponential scaling factor; default value of **1.0997**
  - EAF ranges from **0.0569 to 80.8271**
  - As EAF increases, effort (cost) increases
- **Using the SRDR database, solve for EAF and observe stratified results**

Operating Environment	Application Domai	Primary Langua	ESLOC (MD	Development H	Dev Mo
Surface Fixed, Manned	Mission Planning	Visual Basic	230506	44885	36

Size = ESLOC / 1000

Effort = development Hours / 152

Schedule is not used in the equation

COCOMO II – Results

COCOMO EAF (Frozen B)			EAF				
Op Env	App Type	Language	Records	Min	Max	Mean	Median
All	All	All	281	0.05	13.31	1.54	0.99
All	Cmd/Ctrl	All	39	0.08	12.73	1.59	1.01
All	Communications	All	37	0.23	8.23	1.26	0.72
All	Custom AIS	All	12	0.25	0.92	0.55	0.55
All	Mission Planning	All	17	0.18	2.00	0.67	0.60
All	Real-Time Embedded	All	59	0.36	10.54	1.77	1.17
All	Scientific/Simulation	All	12	0.06	1.77	0.75	0.67
All	Signal Processing	All	22	0.19	7.66	2.54	2.17
All	SW Tools	All	6	0.05	0.59	0.34	0.34
All	Systems Software	All	35	0.06	13.31	2.07	1.15
All	Test/Meas/Diag Equip	All	4	1.45	3.03	2.47	2.69
All	Training	All	2	0.10	0.39	0.25	0.25
All	Vehicle Control	All	22	0.17	4.79	1.44	1.32
All	Vehicle Payload	All	14	0.29	2.22	1.44	1.53
All	All	C/C++	182	0.05	13.31	1.51	0.93
All	All	Ada	53	0.08	12.73	2.01	1.45
All	All	Java	39	0.23	3.99	1.16	0.88
All	All	Other	6	0.22	2.37	0.75	0.49
Air Veh, Manned	All	All	53	0.08	5.19	1.26	1.08
Air Veh, Unmanned	All	All	21	0.19	4.79	1.23	0.85
Ord Sys, Unmanned	All	All	27	0.05	4.77	1.15	1.08
Sea Sys, Manned	All	All	28	0.28	7.25	1.64	0.82
Sea Sys, Unmanned	All	All	2	0.29	2.26	1.27	1.27
Surface Fixed, Man	All	All	100	0.12	12.73	1.41	0.87
Surface Mobile, Manned	All	All	12	0.71	8.23	2.97	2.02
Surface Vehicle, Manned	All	All	34	0.06	13.31	2.25	1.32
Surface Vehicle, Unmanned	All	All	4	1.15	3.33	1.99	1.75

Example (Mission Plan, Median EAF):
Effort = 2.94 * (0.60) (Size)$^{1.0997}$

All calculated EAF values are "in-bounds" with model thresholds

COCOMO II – Accuracy

				Mean EAF		Median EAF	
Op Env	App Type	Language	Records	MMRE	PRED(30)	MMRE	PRED(30)
All	All	All	281	1.76	22.4%	1.07	27.4%
All	Cmd/Ctrl	All	39	1.74	30.8%	1.01	38.5%
All	Communications	All	37	1.12	16.2%	0.47	45.9%
All	Custom AIS	All	12	0.33	58.3%	0.33	58.3%
All	Mission Planning	All	17	1.13	35.3%	1.00	17.6%
All	Real-Time Embedded	All	59	0.88	28.8%	0.53	28.8%
All	Scientific/Simulation	All	12	1.37	33.3%	1.22	41.7%
All	Signal Processing	All	22	1.42	40.9%	1.19	45.5%
All	SW Tools	All	6	1.28	16.7%	1.29	16.7%
All	Systems Software	All	35	3.08	14.3%	1.59	20.0%
All	Test/Meas/Diag Equip	All	4	0.26	75.0%	0.26	75.0%
All	Training	All	2	0.89	0.0%	0.89	0.0%
All	Vehicle Control	All	22	1.42	45.5%	1.30	40.9%
All	Vehicle Payload	All	14	0.67	78.6%	0.70	71.4%
All	All	C/C++	182	1.88	22.0%	1.08	25.8%
All	All	Ada	53	1.96	30.2%	1.37	37.7%
All	All	Java	39	0.82	28.2%	0.56	41.0%
All	All	Other	6	0.99	16.7%	0.52	50.0%
Air Veh, Manned	All	All	53	1.58	26.4%	1.35	24.5%
Air Veh, Unmanned	All	All	21	1.22	19.0%	0.79	33.3%
Ord Sys, Unmanned	All	All	27	2.26	37.0%	2.11	37.0%
Sea Sys, Manned	All	All	28	1.31	21.4%	0.58	32.1%
Sea Sys, Unmanned	All	All	2	1.92	0.0%	1.92	0.0%
Surface Fixed, Man	All	All	100	1.42	23.0%	0.78	33.0%
Surface Mobile, Manned	All	All	12	0.98	33.3%	0.65	8.3%
Surface Vehicle, Manned	All	All	34	2.81	20.6%	1.57	20.6%
Surface Vehicle, Unmanned	All	All	4	0.35	50.0%	0.25	50.0%

- **COCOMO II performs average to good for many strata using mean and median EAF values**
- **Performs well in the Custom AIS, TMDE, Vehicle Payload strata**
- **Difficult to remove any data points as all the calculated EAF values are "in-bounds"**
- **Identify another value in the COCOMO II model that can indicate possible candidates for removal**

COCOMO II Exponent

- **The effort equation's scaling exponent is also used in the COCOMO II Schedule Equation**
  - Schedule = 3.67 * (Effort)^F
  - F = 0.28 + 0.2 * (E – 0.91)
- **Where the COCOMO II effort equation does not use schedule as an input, the data in the SRDR database could be used to solve for E**
  - E = [(ln(Schedule/3.67)/ln(Effort)) - 0.098)] / 0.2
- **E ranges from 0.91 to 1.2262**
- **Calculate E from the Schedule and Effort values of the SRDR database and remove values that are out of bounds**

COCOMO II Model – In-Bounds Accuracy

				Mean EAF			Median EAF		
Op Env	App Type	Language	Records	MMRE	PRED(30)	Delta	MMRE	PRED(30)	Delta
All	All	All	38	1.06	18.4%	-4.0%▼	1.05	18.4%	-9.0%▼
All	Cmd/Ctrl	All	8	2.07	25.0%	-5.8%▼	2.19	25.0%	-13.5%▼
All	Communications	All	4	1.74	25.0%	8.8%▲	1.71	25.0%	-20.9%▼
All	Custom AIS	All	5	0.68	40.0%	-18.3%▼	0.65	40.0%	-18.3%▼
All	Mission Planning	All	3	1.44	0.0%	-35.3%▼	1.45	0.0%	-17.6%▼
All	Real-Time Embedded	All	5	0.83	0.0%	-28.8%▼	0.80	0.0%	-28.8%▼
All	Scientific/Simulation	All	0						
All	Signal Processing	All	1	0.25	100.0%	59.1%▲	0.25	100.0%	54.5%▲
All	SW Tools	All	0						
All	Systems Software	All	4	0.60	25.0%	10.7%▲	0.60	25.0%	5.0%▲
All	Test/Meas/Diag Equip	All	0						
All	Training	All	0						
All	Vehicle Control	All	3	0.45	0.0%	-45.5%▼	0.44	0.0%	-40.9%▼
All	Vehicle Payload	All	5	0.40	20.0%	-58.6%▼	0.40	20.0%	-51.4%▼
All	All	C/C++	26	1.00	15.4%	-6.6%▼	0.97	15.4%	-10.4%▼
All	All	Ada	5	1.46	20.0%	-10.2%▼	1.45	20.0%	-17.7%▼
All	All	Java	7	1.00	28.6%	0.4%▲	1.02	28.6%	-12.5%▼
All	All	Other	0						
Air Veh, Manned	All	All	6	1.63	16.7%	-9.7%▼	1.60	16.7%	-7.9%▼
Air Veh, Unmanned	All	All	4	0.97	0.0%	-19.0%▼	0.96	0.0%	-33.3%▼
Ord Sys, Unmanned	All	All	7	0.67	14.3%	-22.8%▼	0.65	14.3%	-22.8%▼
Sea Sys, Manned	All	All	1	4.55	0.0%	-21.4%▼	4.55	0.0%	-32.1%▼
Sea Sys, Unmanned	All	All	0	0.00	0.0%	0.0%▼	0.91	8.2%	8.2%▲
Surface Fixed, Man	All	All	17	0.94	23.5%	0.5%▲	0.95	23.5%	-9.5%▼
Surface Mobile, Manned	All	All	0						
Surface Vehicle, Manned	All	All	3	0.56	33.3%	12.7%▲	0.56	33.3%	12.7%▲
Surface Vehicle, Unmanned	All	All	0						

- **Scaling Exponent boundaries remove 243 data points**
- **Calculating the Scaling Exponent from Schedule and Effort and removing the data points with "out of bounds" values eliminates almost all accuracy**
- **Prediction values almost universally drop to unusable values**

Approved for Public Release
16-MDA-8881 (4 October 16)

Paired Data

- **A popular method with this type of database is stratifying by Operating Environment and Application Domain together to analyze data**
- **As more strata are introduced, less values are available**
- **What is the accuracy in terms of PRED(30) of the mean/median productivity calculations of paired data for both full and "in-bounds" data sets?**
- **Key assumption: Need at least 5 data points in a paired strata to be applicable**

SLIM-Estimate Paired Data Performance

Unbounded:

		Cmd/Ctrl	Comms	Custom AIS	Miss Plng	RTE	Sci/Sim	Sig Proc	SW Tools	Sys SW	TMDE	Trng	Veh Ctrl	Veh Pay
Air Veh, Manned	Mean	12.5%				11.8%			0.0%	20.0%				33.3%
	Median	25.0%				23.5%			20.0%	40.0%				33.3%
Air Veh, Unmanned	Mean	20.0%				0.0%								
	Median	20.0%				0.0%								
Ord Sys, Unmanned	Mean	20.0%				16.7%		16.7%						
	Median	40.0%				16.7%		16.7%						
Sea Sys, Manned	Mean		0.0%			0.0%								
	Median		0.0%			0.0%								
Sea Sys, Unmanned	Mean													
	Median													
Surface Fixed, Man	Mean	15.8%	26.3%	11.1%	11.8%	23.5%	20.0%			27.3%				
	Median	15.8%	15.8%	33.3%	23.5%	5.9%	0.0%			18.2%				
Surface Mobile, Manned	Mean													
	Median													
Surface Vehicle, Manned	Mean					0.0%	30.8%						18.2%	
	Median					28.6%	7.7%						18.2%	
Surface Vehicle, Unmanned	Mean													
	Median													

- **Very few populated pairs (5 or more records)**
- **Accuracy is still low (none over 50%)**

SLIM-Estimate Paired Data Performance – In-Bounds

Bounded:

		Cmd/Ctrl	Comms	Custom AIS	Miss Plng	RTE	Sci/Sim	Sig Proc	SW Tools	Sys SW	TMDE	Trng	Veh Ctrl	Veh Pay
Air Veh, Manned	Mean	0.0%				11.0%			0.0%	20.0%				
	Median	0.0%				0.0%			20.0%	40.0%				
Air Veh, Unmanned	Mean	20.0%				0.0%								
	Median	20.0%				0.0%								
Ord Sys, Unmanned	Mean	20.0%												
	Median	40.0%												
Sea Sys, Manned	Mean													
	Median													
Sea Sys, Unmanned	Mean													
	Median													
Surface Fixed, Man	Mean	30.8%	29.4%	50.0%	18.8%	22.2%	20.0%			12.5%				
	Median	30.8%	35.3%	37.5%	18.8%	22.2%	0.0%			37.5%				
Surface Mobile, Manned	Mean													
	Median													
Surface Vehicle, Manned	Mean					20.0%							11.1%	
	Median					0.0%							11.1%	
Surface Vehicle, Unmanned	Mean													
	Median													

- **Fewer populated pairs (5 or more records)**
- **Accuracy is not improved (one at 50%)**

SEER-SEM Paired Data Performance

Unbounded:

		Cmd/Ctrl	Comms	Custom AIS	Miss Plng	RTE	Sci/Sim	Sig Proc	SW Tools	Sys SW	TMDE	Trng	Veh Ctrl	Veh Pay
Air Veh, Manned	Mean	14.3%				17.6%			0.0%	0.0%				16.7%
	Median	42.9%				29.4%			20.0%	40.0%				0.0%
Air Veh, Unmanned	Mean	20.0%				40.0%								
	Median	40.0%				20.0%								
Ord Sys, Unmanned	Mean					16.7%		0.0%						
	Median					50.0%		33.3%						
Sea Sys, Manned	Mean		42.9%			20.0%								
	Median		42.9%			20.0%								
Sea Sys, Unmanned	Mean													
	Median													
Surface Fixed, Man	Mean	5.6%	5.3%	33.3%	0.0%	5.9%				0.0%				
	Median	0.0%	36.8%	44.4%	5.9%	17.6%				18.5%				
Surface Mobile, Manned	Mean													
	Median													
Surface Vehicle, Manned	Mean					14.3%				8.3%			36.4%	
	Median					14.3%				16.7%			27.3%	
Surface Vehicle, Unmanned	Mean													
	Median													

- **Fewer populated pairs (5 or more records)**
- **Accuracy is not improved (one at 50%)**

SEER-SEM Paired Data Performance – In-Bounds

Bounded:

SEER (IB)		Cmd/Ctrl	Comms	Custom AIS	Miss Plng	RTE	Sci/Sim	Sig Proc	SW Tools	Sys SW	TMDE	Trng	Veh Ctrl	Veh Pay
Air Veh, Manned	Mean													
	Median													
Air Veh, Unmanned	Mean													
	Median													
Ord Sys, Unmanned	Mean													
	Median													
Sea Sys, Manned	Mean													
	Median													
Sea Sys, Unmanned	Mean													
	Median													
Surface Fixed, Man	Mean		0.0%	66.7%	16.7%	5.9%				0.0%				
	Median		66.7%	66.7%	0.0%	17.6%				20.0%				
Surface Mobile, Manned	Mean													
	Median													
Surface Vehicle, Manned	Mean													
	Median													
Surface Vehicle, Unmanned	Mean													
	Median													

- **Very few populated pairs – total dataset reduced to 69 records using Staffing Complexity boundary**
- **Accuracy does improve in populated pairs (some instances of 66.7%)**

COCOMO II Paired Data Performance

Unbounded:

COCOMO II (All)		Cmd/Ctrl	Comms	Custom AIS	Miss Plng	RTE	Sci/Sim	Sig Proc	SW Tools	Sys SW	TMDE	Trng	Veh Ctrl	Veh Pay
Air Veh, Manned	Mean	50.0%				23.5%			40.0%	40.0%				83.3%
	Median	62.5%				29.4%			60.0%	60.0%				83.3%
Air Veh, Unmanned	Mean	40.0%				0.0%								
	Median	40.0%				0.0%								
Ord Sys, Unmanned	Mean	40.0%				33.3%		33.3%						
	Median	60.0%				33.3%		33.3%						
Sea Sys, Manned	Mean		7.1%			60.0%								
	Median		64.3%			60.0%								
Sea Sys, Unmanned	Mean													
	Median													
Surface Fixed, Man	Mean	15.8%	42.1%	44.4%	35.3%	0.0%	60.0%			18.2%				
	Median	31.6%	36.8%	55.6%	17.6%	29.4%	60.0%			36.4%				
Surface Mobile, Manned	Mean													
	Median													
Surface Vehicle, Manned	Mean					14.3%	23.1%						36.4%	
	Median					28.6%	23.1%						63.6%	
Surface Vehicle, Unmanned	Mean													
	Median													

- **Several pairs with a PRED(30) of 50% or higher**
- **Median value for EAF in many pairs is more accurate than other models**

- **NOTE: Bounding the dataset by calculating the E exponent from schedule reduces the number of records to 38. There are no pairs of 5 records or more.**

Conclusions

- SRDR database can be normalized to be “CSCI-like” in size and limited to Design, Code, and Test phases
- Overall performance by using a calculated productivity variable (mean and median) for popular models does not always produce credible results
- Pairing strata is beneficial, but limited by the number of strata represented
- COCOMO II outperforms SEER-SEM and SLIM-Estimate models using this limited methodology
- COCOMO II loses all predictive capability when schedule is integrated into the selection
- Schedule integration in a cost model using SRDR data impacts prediction negatively

Future Research

- Performance against tool-calibrated model (Calico, SEER-SEM calibration, etc.)
- Regression Analysis on New SRDR Dataset
- Schedule Variable Impact

Questions