



Using Benchmarking to Quantify the Benefits of Process Improvement

Larry Putnam, Jr.

Co-CEO

**Quantitative Software Management
Inc.**

Agenda

- Benchmarking Overview
- Data Collection
 - What to Measure
 - Dataset Demographics?
- Organizational Baselineing
 - Internal Benchmark
 - Outlier Analysis
 - Opportunities for Process Improvements
 - Business Case for Productivity & Quality Improvement
 - Using multiple baselines for improvement trending
- External Benchmarking
 - Selecting Appropriate External Datasets
 - Benchmark Comparisons

What is Benchmarking?

Benchmarking is the process of **comparing** one's **business processes** and **performance metrics** to **industry** bests and/or **best practices** from other industries. Dimensions typically measured are **quality, time, and cost**. **Improvements** from learning mean doing things **better, faster, and cheaper**.

en.wikipedia.org/wiki/Benchmarking

Why Do Organizations Benchmark?

Typical Motivations:

- Assess internal and external competitive position
- To support a process assessment (CMMI, ISO, etc.)
- Assess capability of and negotiate with suppliers
- Provide baseline to measure improvements
- Quantify the benefits of improvement
- Support realistic estimation processes

Solutions:

- QSM Benchmarking Services
 - High Performance Benchmark Service
 - SLIM-Metrics & SLIM-DataManager Products
 - SLIM WebServices (Closeout & Benchmark Services)
 - Performance Benchmark Tables -
<http://www.qsm.com/resources/performance-benchmark-tables>

Best Practices Core Quantitative Metrics

- Size and Scope
 - Abstraction of measurements change depending on where you are in the lifecycle
 - Tells you the overall size of system
 - Tells how much of the product is completed at different points in time
- Schedule
 - Measured in calendar time like days, weeks, months or years
 - Overall duration of the project
 - Tells how much of the overall schedule has been consumed and identifies key task/milestone completion dates
- Effort
 - Measured in labor units like person hours, days, weeks, months or years
 - Proportional to the overall cost
 - Tells how labor/money is being spent
- Defects
 - Deviations from specification usually measured by severity level
 - Tells you how many defects you are likely to find and when
 - Tells how stable the product is at different points in time and helps project when the product will be reliable enough deploy

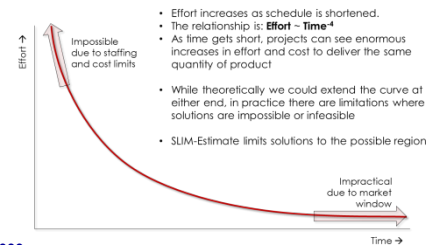
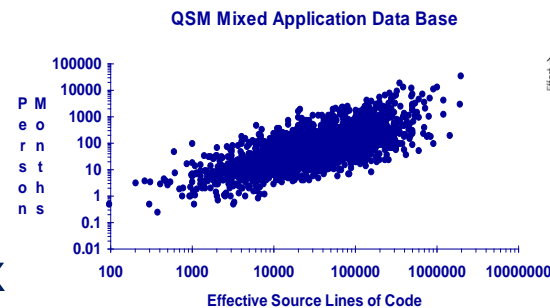
Benchmark Qualitative Information

- Project Demographics
 - Application type
 - Development classification (new, enhancement, maintenance, etc...)
 - Industry sector
 - Technology used
- Tools & Methods Used
 - Requirements, design construction, test, PM, etc...
- Technical Complexity
 - Security requirements
 - Data, algorithm, logic & interface complexity
 - Customer interface
 - Documentation requirements
 - Platform stability
- Personnel Profile
 - Overall team skills & experience, cohesion & motivation, management effectiveness

We Need Comprehensive Productivity Metrics

Typical Productivity Metrics:

- Ratio based measures
 - Size/Effort
 - LOC/PM, FP/PM
- They are not sufficient to explain software development's non-linear behavior
 - As size increases
 - As change staffing levels
- Solution
 - QSM's Productivity Index
 - Takes schedule, effort & size into account
 - Also ties in quality/reliability



Production Equation



Delivered
System
Size

proportional
to

Effort

over

Time

at some

Productivity

A measure of
Value Delivered

A measure of
Resources
Expended

A measure of
Duration Required

A measure of
Capability and
Difficulty of the
task

Determining Development Team Efficiency



Developer Efficiency (PI)

=

Functionality Developed

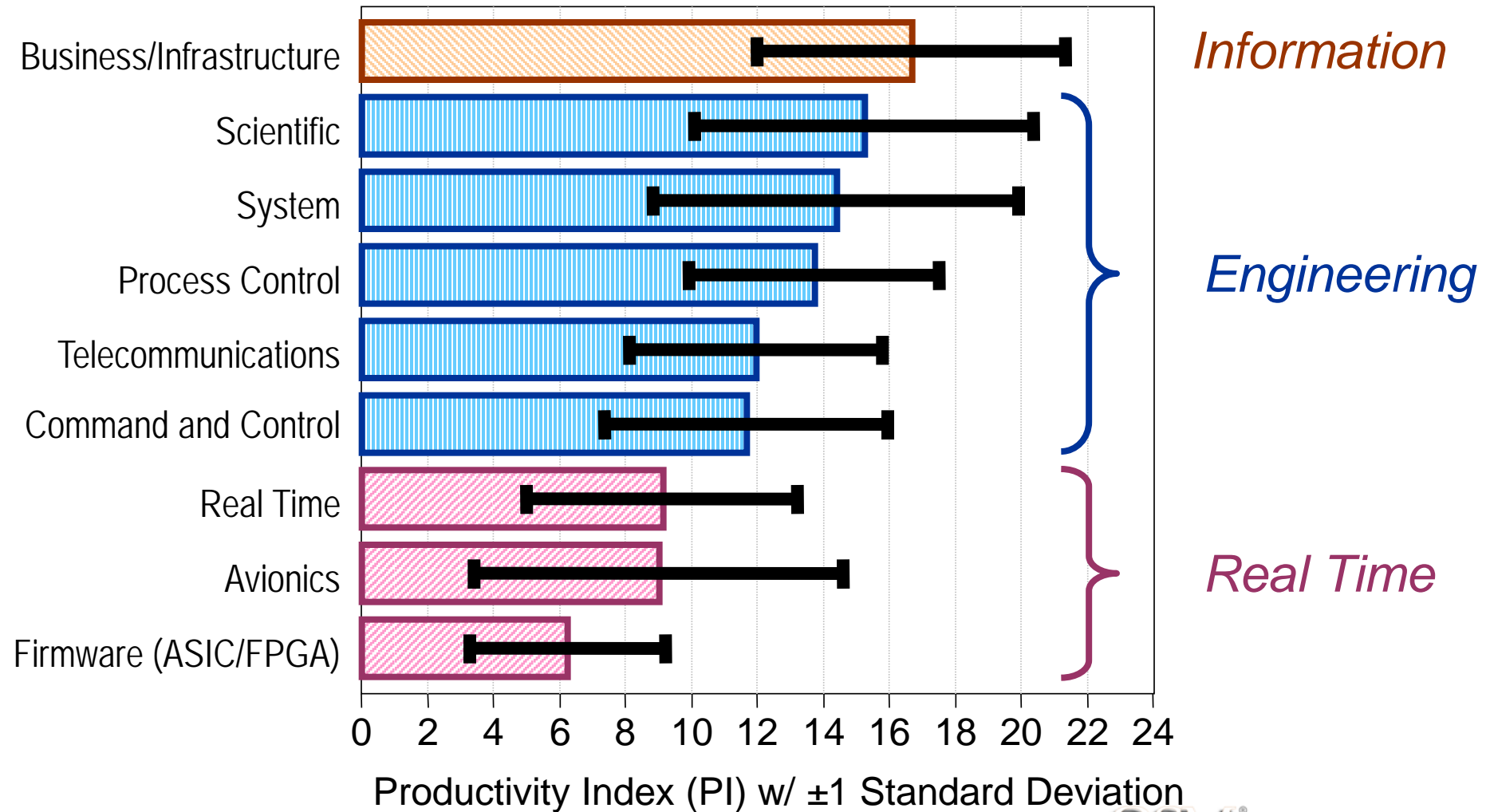
Effort x Time



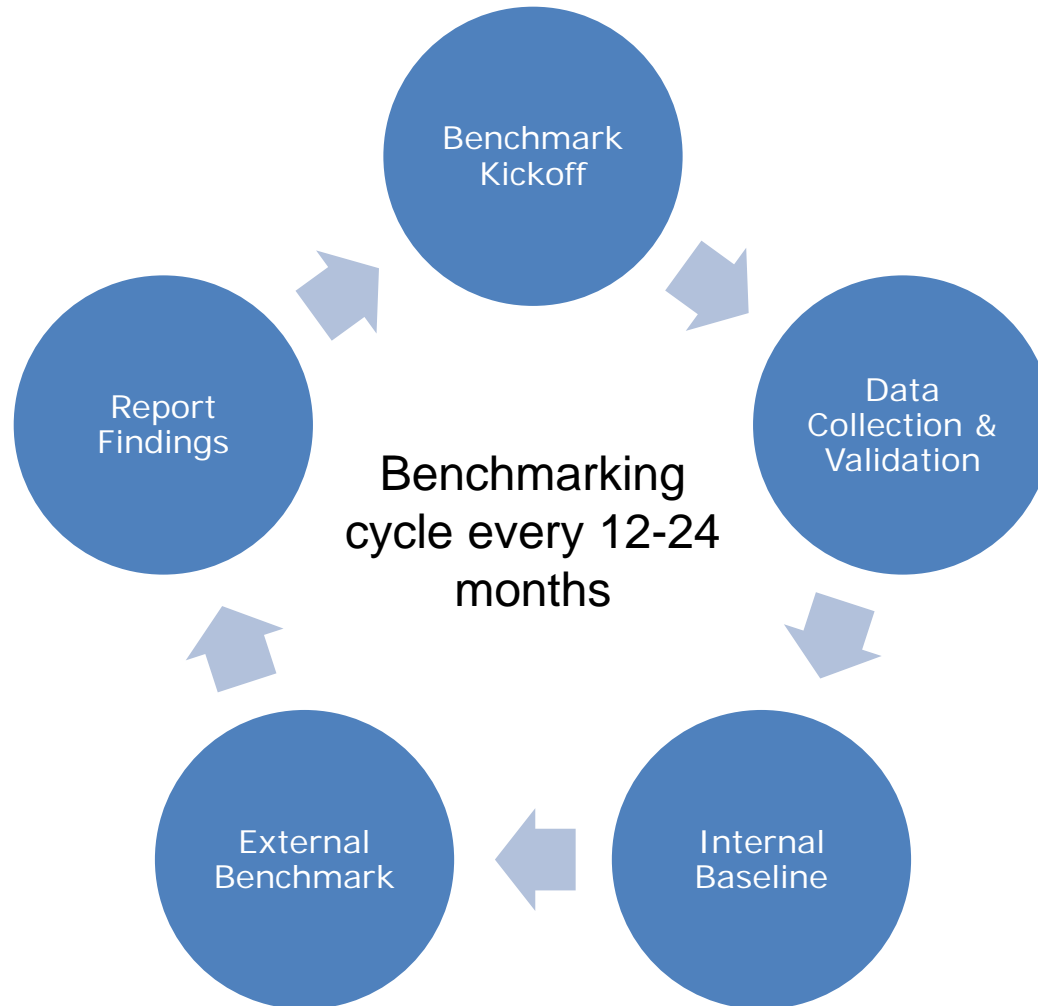
A Higher value means less time and effort

Productivity Index (PI)

(industry values by application type)



QSM Benchmarking Process



Good Start Kick-off Meeting

Objectives:

- Identify individuals who will participate
 - Coordinate benchmark analysis
 - Identify data sources
- Explain data requirements & definitions
- Establish schedule and deliverables
- Brief out benchmark process and objectives
 - Overview of QSM benchmarking process
 - Outline benchmark objectives
 - Provides a “stake in the ground” for future comparison
 - Focused on positive organizational improvement
 - Provide a quantitative basis for providing realistic future estimates

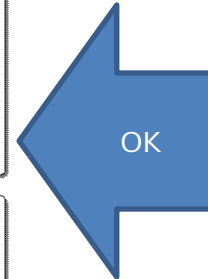
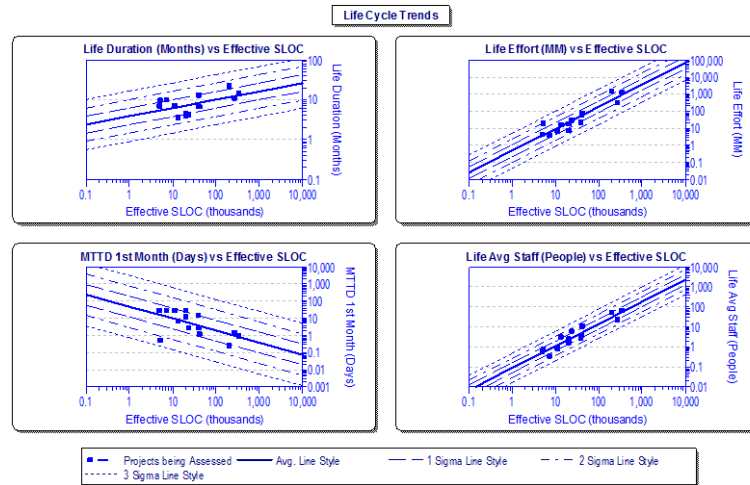
Data Collection and Validation

Data Collection Form

Phase	Start Date	End Date	Hours	PHR	1000E	Peak Staff	Staffing Shape
PLANS	01/15/1998	04/30/1998	142	107	16,700	15	Front Load Backlog
ANAL	03/05/1998	07/05/1998	148	226	7,700	1	Front Load Backlog
IMP	06/17/1998	06/17/1998	143	17,324	23,900	2	Median Front Load Backlog
DEVEL	06/23/1998	10/23/1998	153	103	12,700	1	Front Load Backlog
UAT/CR	01/19/1999	10/23/1999	214	22,294	44,652	200	Front Load Backlog



Analysis Begins



Coordinator



Benchmark Consulting Team



Data Validation Interview (2 hours)



Benchmark Consulting Team



Company XYZ Internal Project Baseline & Improvement Analysis

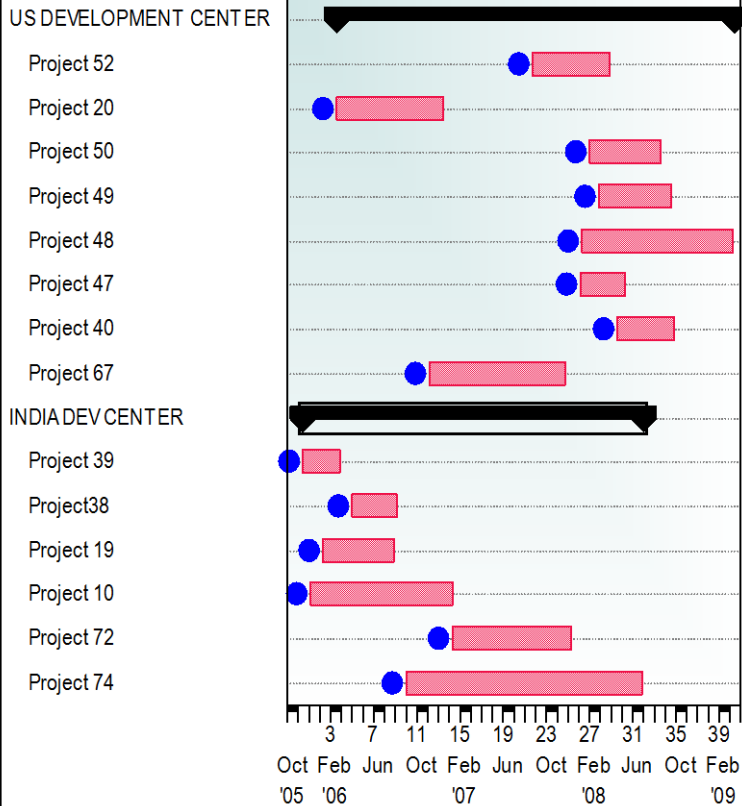
Method of Analysis

- Step 1: Reconstruct a time series resource profile of all projects being assessed.
 - Assess phase allocations of time, effort and overlap percentages.
- Step 2: Curve fit trend lines from the data to establish the average behavior for schedule, effort, staffing, reliability and statistical variation within the project set.
- Step 3: Identify outliers on the high and low end performance spectrum.
- Step 4: High and low performers are evaluated in detail to understand the significant factors influencing productivity and quality performance.
- Step 5: Recommendations for improvement are formulated.
- Step 6: Business case for improvement is developed.

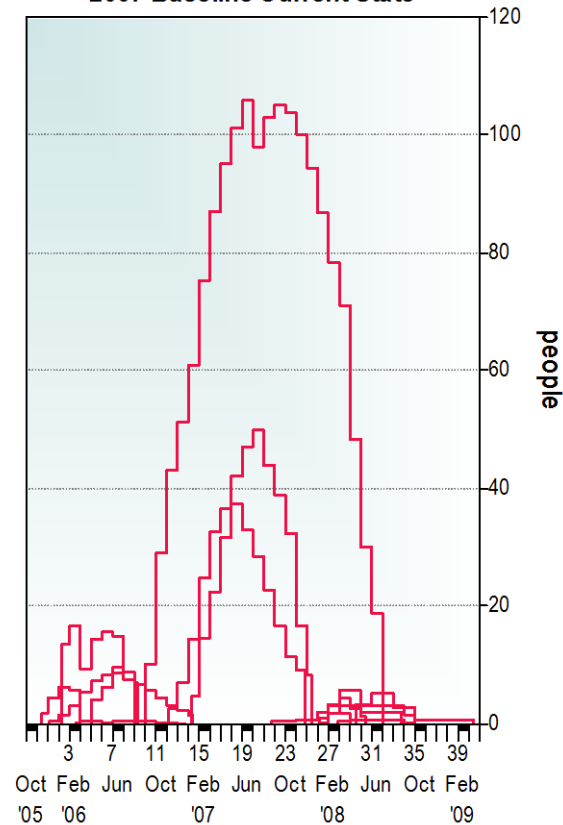
Project Reconstructions

Staffing & Schedule

Monthly Gantt Chart (L3)
 < 2007 Baseline Current State >



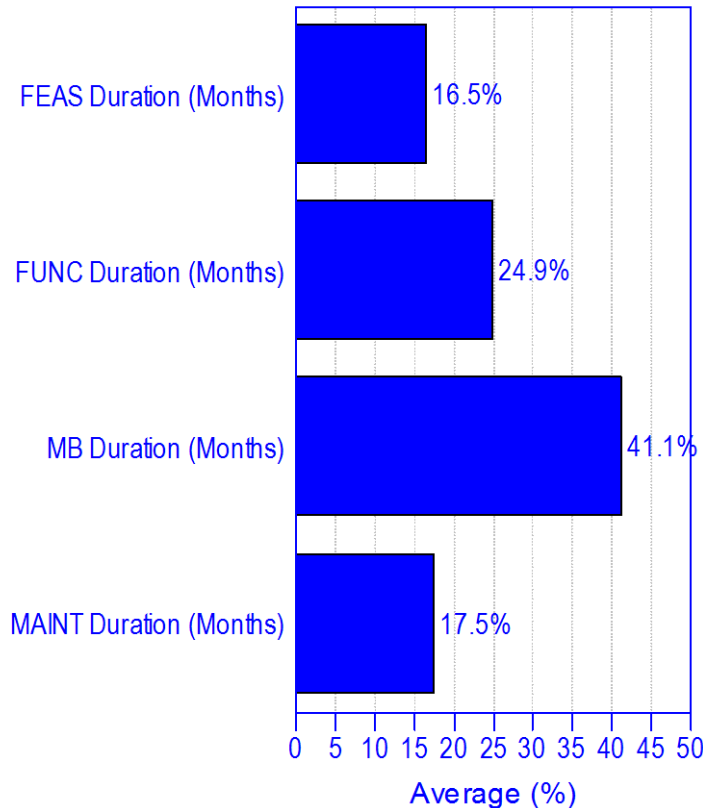
Monthly Avg Staff (L3)
 < 2007 Baseline Current State >



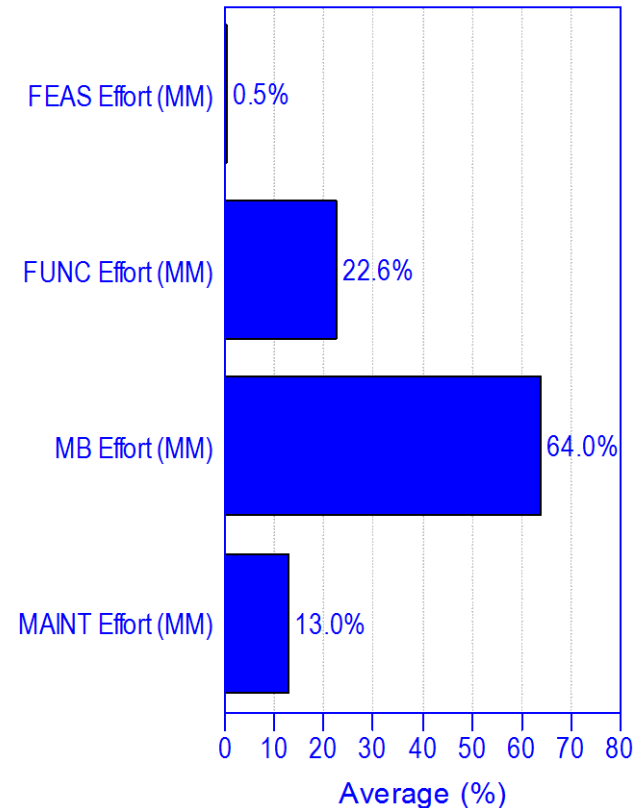
Average Schedule and Effort Distribution

Schedule & Effort Distributions

Average Value of Metrics (%)



Average Value of Metrics (%)

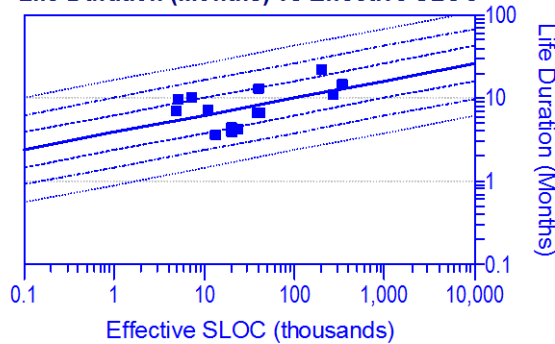


■ Projects being Assessed

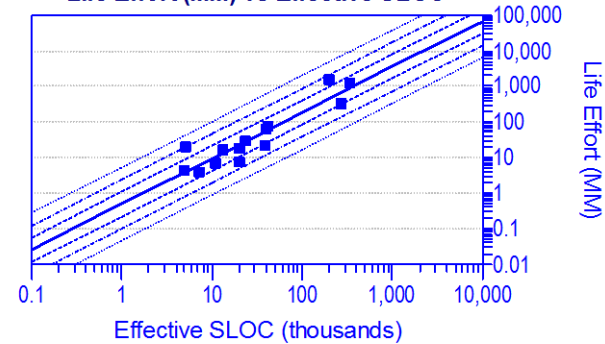
Complete Project Curve Fit Reference Trend lines

Life Cycle Trends

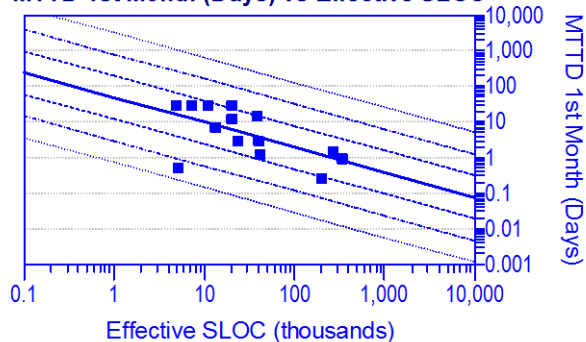
Life Duration (Months) vs Effective SLOC



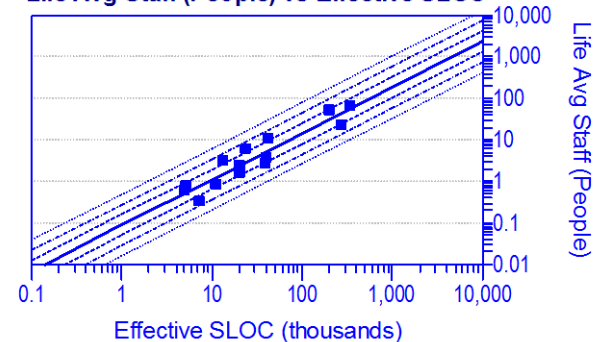
Life Effort (MM) vs Effective SLOC



MTTD 1st Month (Days) vs Effective SLOC



Life Avg Staff (People) vs Effective SLOC



Schedule, effort and staffing all increase as the size of the delivered product increases.

Reliability at delivery decreases as the size get bigger.

Take away – Keep the new and modified functionality as small as possible on new projects to contain cost and schedule and deliver high quality.

■ — Projects being Assessed — Avg. Line Style 1 Sigma Line Style 2 Sigma Line Style
 3 Sigma Line Style

Identification of High and Low Performing Projects

Internal Comparison Report

Star Mappings: 1 star (bottom 20%), 2 stars (bottom 20% to 45%), 3 stars (bottom 45% to top 30%), 4 stars (top 30% to top 10%), 5 stars (top 10%)

Deviations Calculated From: Projects being Assessed

Projects	FUNC Effort (MM)	FUNC Duration (Months)	MB Effort (MM)	MB Duration (Months)	PI	MTD 1st Month (Days)	Composite Project Star Rating
Project 10	☆☆☆	★	☆☆	★	★	☆☆	☆☆
Project 38	☆☆☆☆	☆☆☆☆	☆☆	☆☆☆☆	☆☆☆	☆☆	☆☆☆
Project 39	☆☆☆☆	☆☆☆☆	☆☆	☆☆☆☆	☆☆☆	☆☆☆	☆☆☆
Project 74	★	★	★	★	★	★	★
Project 20	☆☆☆	☆☆☆	☆☆☆	★	☆☆	☆☆☆☆	☆☆☆
Project 19	☆☆	☆☆☆☆	☆☆☆	☆☆☆☆	☆☆☆☆	☆☆	☆☆☆
Project 40	☆☆☆	☆☆☆	☆☆☆	☆☆☆	☆☆☆	☆☆☆	☆☆☆
Project 47	☆☆☆☆	☆☆☆☆	☆☆☆☆	☆☆☆☆	☆☆☆☆	☆☆☆☆	☆☆☆☆
Project 49	☆☆	☆☆	☆☆	☆☆	☆☆	☆☆☆	☆☆
Project 50	☆☆☆☆	☆☆☆	☆☆☆☆	☆☆☆☆	☆☆☆☆	☆☆☆☆	☆☆☆☆
Project 48	★	★	★	★	★	★	★
Project 52	☆☆☆	☆☆☆	☆☆☆☆	☆☆☆	☆☆☆	☆☆☆☆	☆☆☆☆
Project 72	☆☆☆☆	☆☆	☆☆☆☆	☆☆☆	☆☆☆	☆☆☆	☆☆☆
Project 67	☆☆	☆☆	☆☆	☆☆	☆☆	☆☆☆	☆☆
Aggregate	☆☆☆	☆☆☆	☆☆☆	☆☆☆	☆☆☆	☆☆☆	☆☆☆

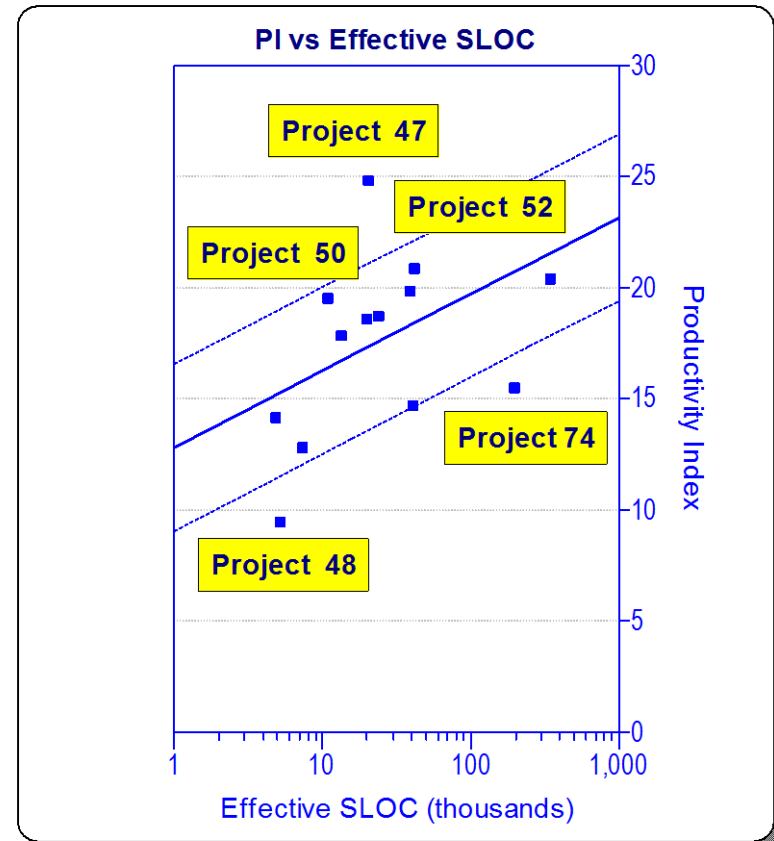
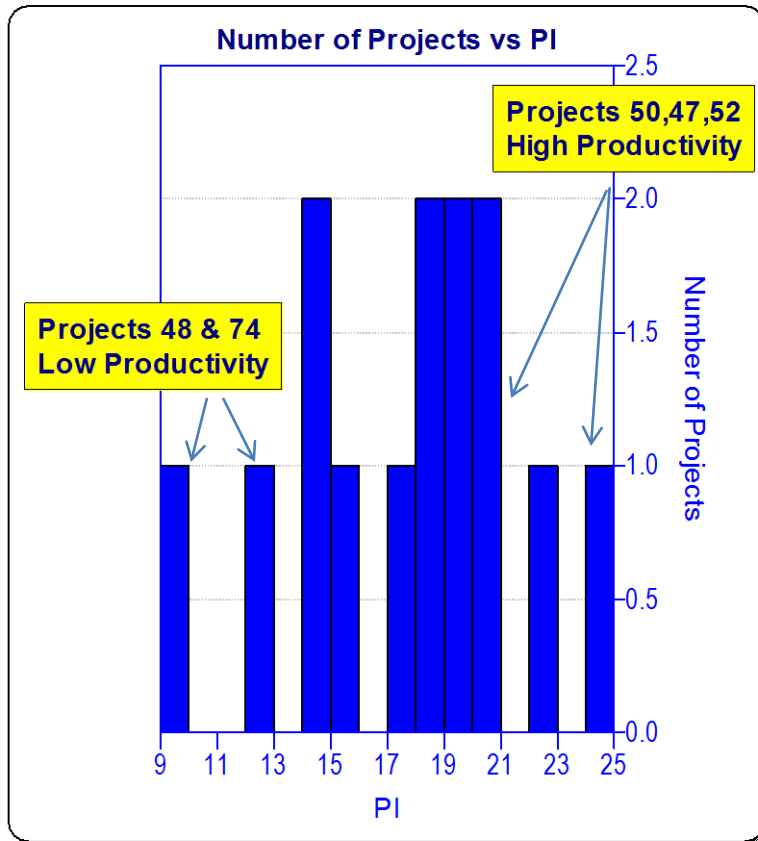
Low Performing Projects

High Performing Projects

The Internal Comparison Report identifies the project's performance relative to the baseline trend used in the comparison - in this case the internal baseline. This technique makes it easy to identify the best and the worst performing projects. The projects identified in yellow shading are the top performers' in this group. A more in-depth analysis of the high and low end performer should identify the key drivers of productivity and quality.

Productivity Index (PI)

PI (Productivity Index)



■ — Projects being Assessed
 — Avg. Line Style
 - - - 1 Sigma Line Style

Most Significant Factors Affecting Productivity & Quality

- Further analysis show that the following factors are the most significant drivers of productivity and quality for XYZ Corporation:
 - Schedule Parallelism/Task Concurrency
 - Extreme Staffing Levels
 - Volume of Requirements Change
- Each factor will be evaluated independently and a prescription for how to improve this activity is provided.



Schedule Parallelism Analysis

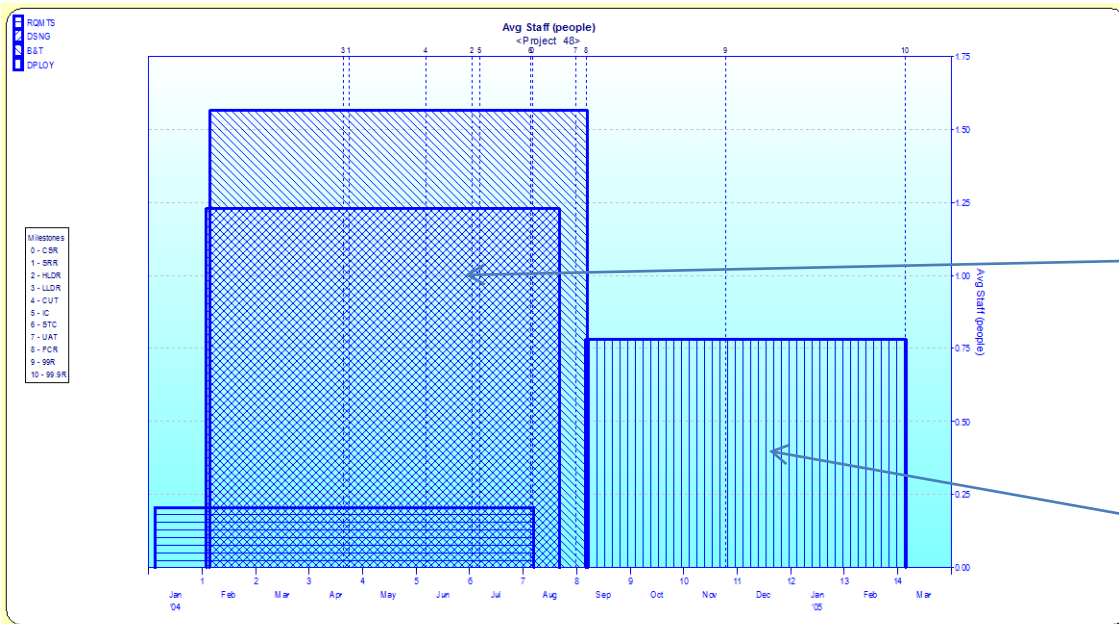
Observed Behavior

In reconstructing the schedule and staffing profiles we noticed that some projects had extreme amounts of schedule overlap between the design, construction and test phases.

Example Project 48:
Staffing and Schedule reconstruction:

Notice the 100% overlap between design & construction.

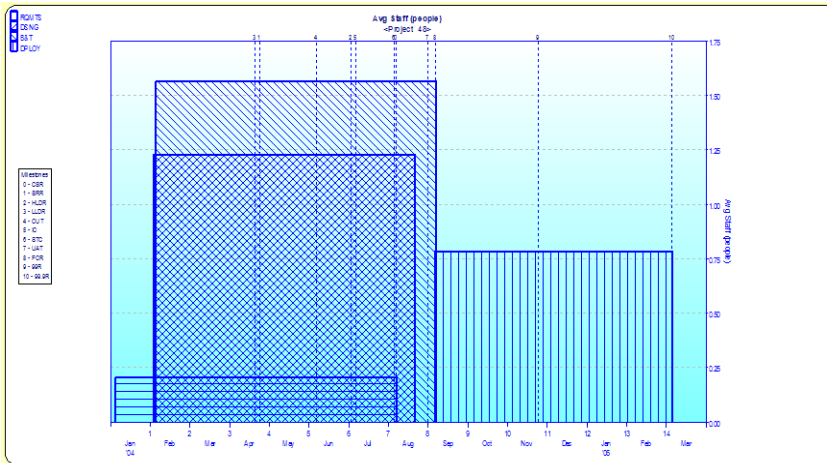
Also notice the long time required to stabilize the product in the perfective maintenance phase.



Why is this Process Problematic?

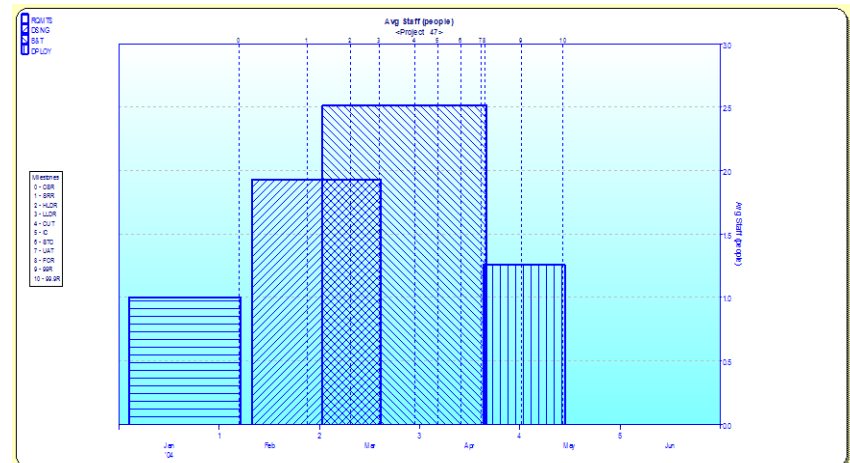
Projects	FUNC Effort (MM)	FUNC Duration (Months)	MB Effort (MM)	MB Duration (Months)	PI	MITD 1st Month (Days)	Composite Project Star Rating
Project 47	★★★★	★★★★★	★★★★★	★★★★★	★★★★★	★★★★★	★★★★★
Project 48	★★	★	★★★	★★	★★	★	★★

Project 48 Low Productivity & Quality



Total parallelism in design and construction results in more defects being introduced into the system, because the team is building from a unspecified design that is likely to change significantly. This causes more rework cycles resulting in low productivity and quality.

Project 47 High Productivity & Quality



A normal degree of overlap reduces the rework because the team is building from a complete and well thought out design. This leads to optimal productivity, schedule compression, and reliability.

Schedule Parallelism

Root Cause

- The US center is constantly battling **unrealistic schedule expectations**. They have no effective way to negotiate with their customers or escalate the issue to a higher level management forum for mediation-resolution.
- The Indian center enjoys a significant labor cost advantage. Because their labor rates are so inexpensive they are using large teams and phase parallelism to try to gain a competitive advantage in schedule performance.

Recommendation

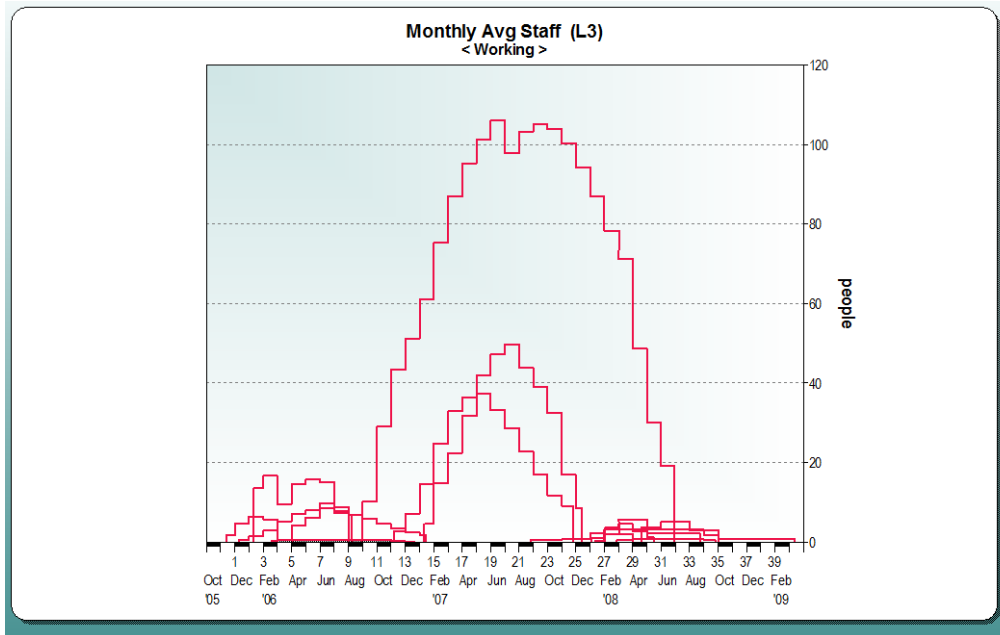
- Use base line trends from this assessment to graphically demonstrate schedule feasibility. Consider setting up a board to reconcile customer expectations with realistic development capabilities.
- Use the one standard deviation trend line for average staffing as a hard constraint for all new projects. Staffing levels above this provide no schedule benefit and are detrimental to cost and reliability.
- **Constrain projects to no more than 50% phase overlap** as a practical limit to schedule compression that will not impact productivity and quality.



Extreme Staffing Analysis

Observed Behavior

- After reconstructing all of the projects we noticed that there were some extreme staffing levels which dramatically impacted schedule, productivity and quality.
 - On the high end reaching over 100 FTE staff
 - On the low end using less than ½ of a FTE



Staffing profile for
Each of the projects
analyzed in this study

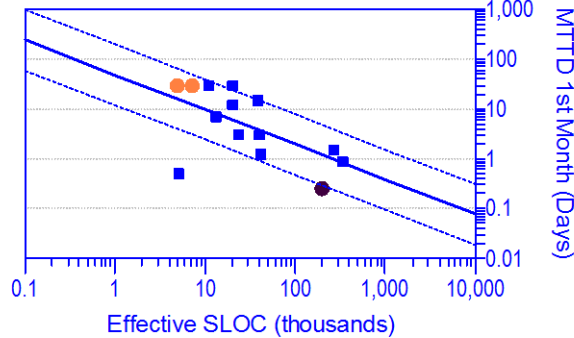
Why is Extreme Staffing Problematic?

- For projects with extremely large teams it has been well documented that:
 - At some point large teams fail to produce any additional schedule reductions.
 - Large teams have more human communication problems which result in a large volume of defects all of which must be eliminated before release.
 - Large teams have large cash flows which cause massive overruns if they do not meet their schedule dates.
- For projects with part time staffing:
 - You lose productivity and continuity as a person juggles two or more projects at one time.
 - Schedules are stretched out which often result in a higher probability of requirements changes.

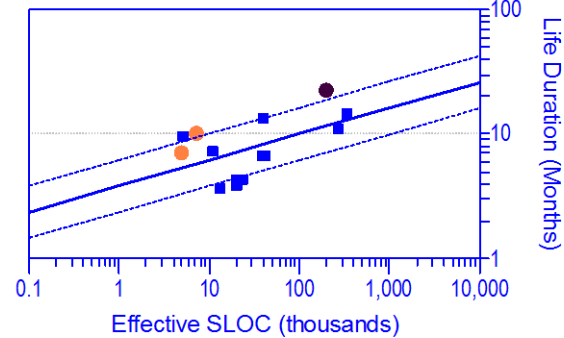
Data Stratified by Staffing Styles

Life Cycle Trends

MTTD 1st Month (Days) vs Effective SLOC

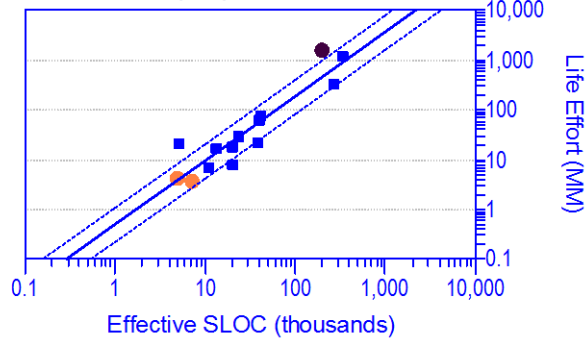


Life Duration (Months) vs Effective SLOC

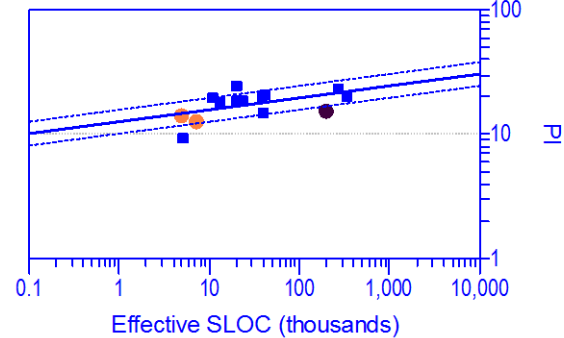


The black circle is a project that used greater than 100 people. Notice that it took longer to develop, was more expensive, had a lower quality at delivery and had a lower productivity.

Life Effort (MM) vs Effective SLOC



PI vs Effective SLOC



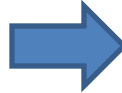
The red circles are the projects that used part time effort. Notice that they took longer to develop, had lower productivity but had better quality and were lower than average cost.

■ — Projects being Assessed
 ● part time effort
 ● High Staffing
 — Avg. Line Style
 - - - 1 Sigma Line Style

Extreme Staffing

Root cause:

- The US development center has some of its staff working several projects at any given time. This practice is abnormal and is a direct result of cost cutting in a recession. The only real penalty for this approach is on schedule performance.
- The Indian development center had a single project that used greater than 100 people. The decision to staff at this level was primarily driven by the sales operation. Market intelligence discovered that the customer's budget would support this funding level on a time and materials contract, so it was staffed accordingly.



Recommendation:

- Re-evaluate resource constraints when economic conditions permit. 1-2 FTE staff will provide better schedule performance and still be cost effective and produce products with high reliability.
- Use a more comprehensive definition of a successful project outcome. Even though the initial engagement may have been profitable, the extreme staffing approach may not be the best strategy for the long term relationship if the schedule and/or reliability performance are poor and the customer is unhappy.



Requirements Change Analysis

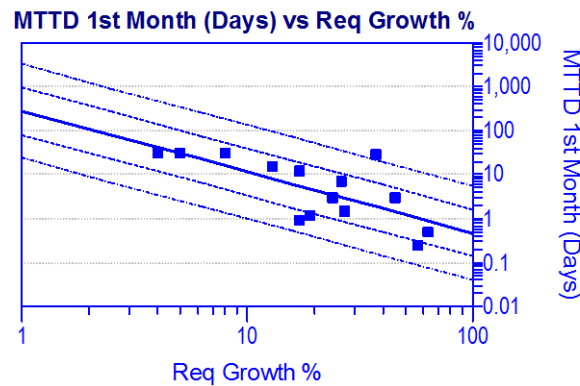
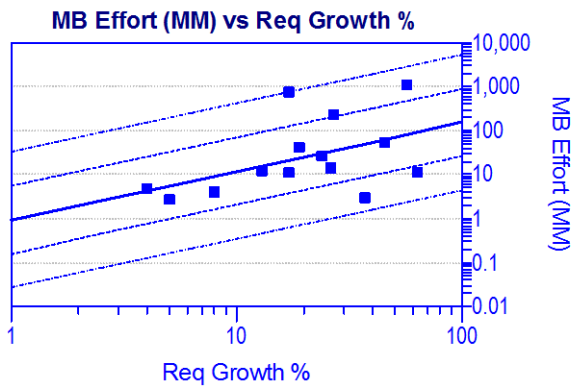
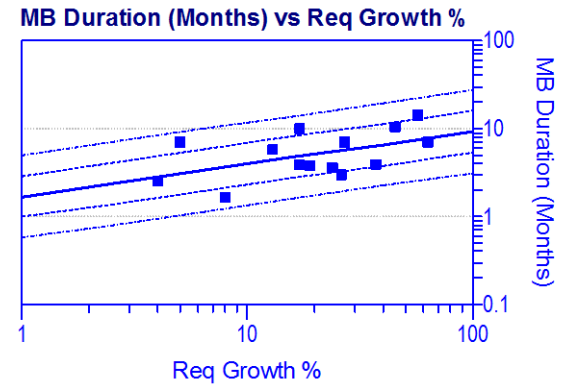
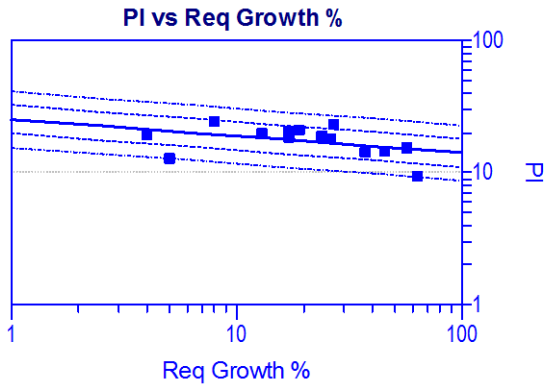
Observed Behavior

- Requirements change during a project varied from 4% on the low end up to a maximum of 64%
- The data showed that as the amount of requirements change increases the following behaviors are observed
 - The schedule increases
 - The effort-cost increases
 - The reliability at delivery decreases
 - The productivity of the project decreases

Observed Behavior

Requirements Growth Impact

As requirements change increases the schedule gets longer, the project effort-cost increases, the reliability at delivery decreases and the overall productivity decreases.

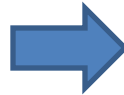


■ — Projects being Assessed — Avg. Line Style - - - - - 1 Sigma Line Style - - - - - 2 Sigma Line Style

Requirement Change Analysis

Root Cause:

Both development centers have problems with requirements change. There is a documented requirements change process in place. Some of the teams simply don't use the process.



Recommendation:

Enforce the existing requirements management process. Establish consequences if the process is not followed.



Economic Opportunities for Change

Economic Opportunities for Change

Key economic assumptions:

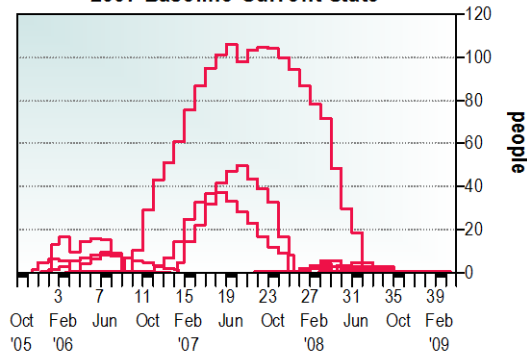
1. Projects with extreme overlap convert to moderate overlap.
2. Extreme Staffing Levels adjusted to recommendations in this study.
3. Requirements change adjusted to 15 percent maximum for poor performers.

As is XYZ Projects Current State						XYZ Projects with Change Implemented			
Project	Productivity Index	Phase overlap	Reqts Change	Average Staffing	Internal 5 star rating	Productivity Index Business case	Phase overlap	Reqts Change	Average Staffing
#52	19.5	Low	13.00%	2.6	3	19.8	Low	13.00%	2.6
#50	19.9	Moderate	4.00%	0.9	4	20.2	Moderate	4.00%	0.9
#49	14.2	None	37.00%	0.6	3	16.2	None	15.00%	1
#48	9.4	Extreme	63.00%	0.8	1	12	Moderate	15.00%	1
#47	24	Moderate	8.00%	1.6	4	24	Moderate	8.00%	1.6
#40	18.6	Moderate	17.00%	2.4	3	19	Moderate	17.00%	2.4
#20	12.8	Moderate	5.00%	0.34	3	15.8	Moderate	5.00%	1
#39	17.9	Moderate	26.00%	3.1	3	18.2	Moderate	26.00%	3.1
#38	18.7	None	24.00%	6	3	19	None	24.00%	6
#19	20.8	None	19.00%	11	3	21.2	None	19.00%	11
#10	14.7	High	45.00%	3.7	2	17	Moderate	15.00%	3.7
#67	22.1	High	17.00%	34.7	2	22.5	Moderate	17.00%	34.7
#72	23	Moderate	27.00%	45	3	23.2	Moderate	27.00%	45
#74	15.5	Extreme	57.00%	105	1	17.5	Moderate	15	70

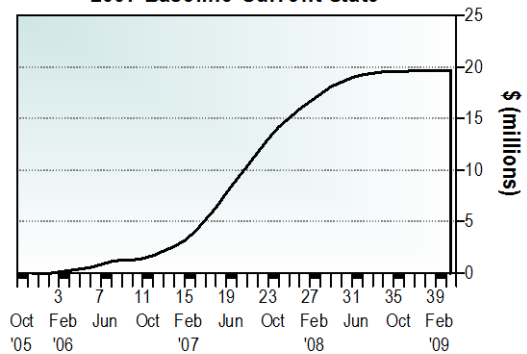
Current Portfolio of XYZ Projects 2007 Baseline

Schedule Staffing & Cost

Monthly Avg Staff (L2)
< 2007 Baseline Current State >



Monthly Cum Cost (L0)
< 2007 Baseline Current State >



Monthly Gantt Chart (L3)
< 2007 Baseline Current State >

US DEVELOPMENT CENTER

Project52

Project20

Project50

Project49

Project48

Project47

Project40

Project67

INDIA DEV CENTER

Project39

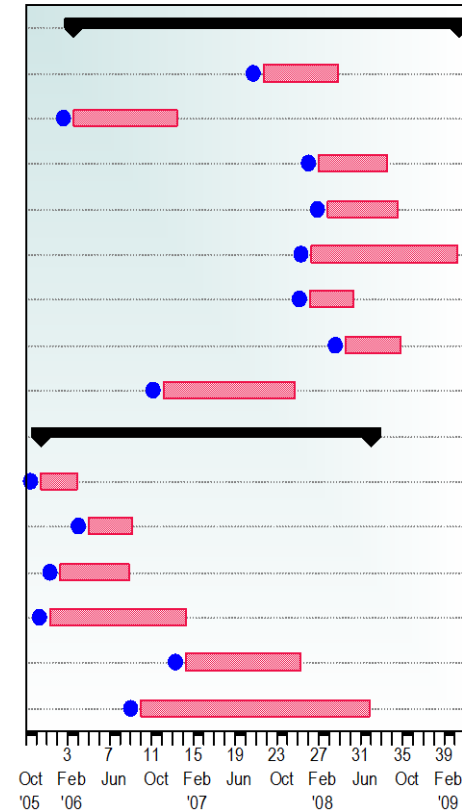
Project38

Project19

Project10

Project72

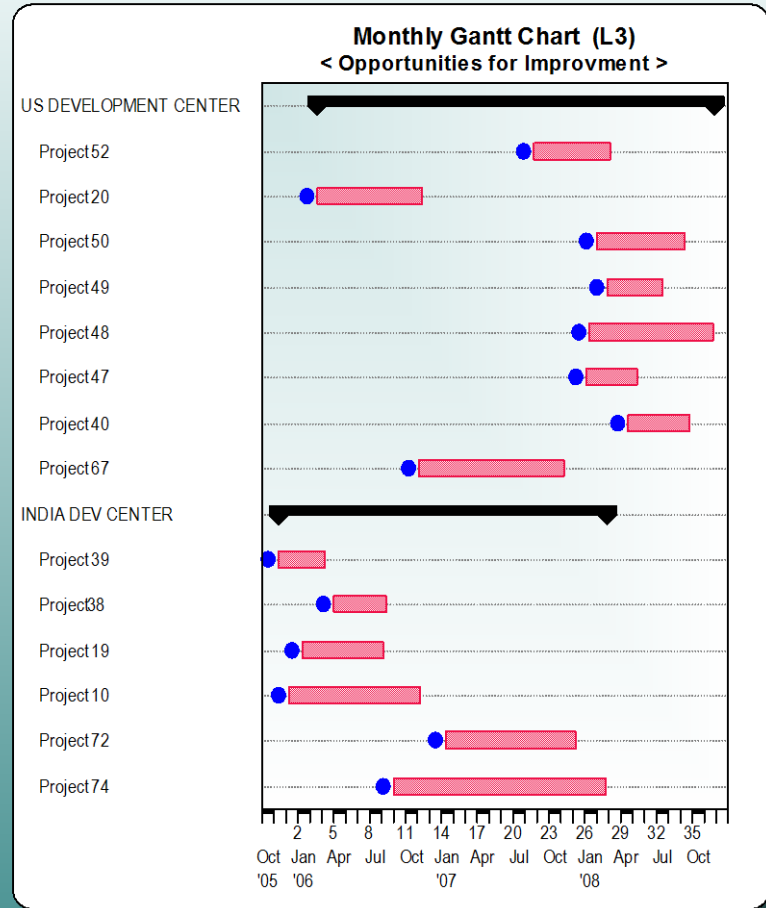
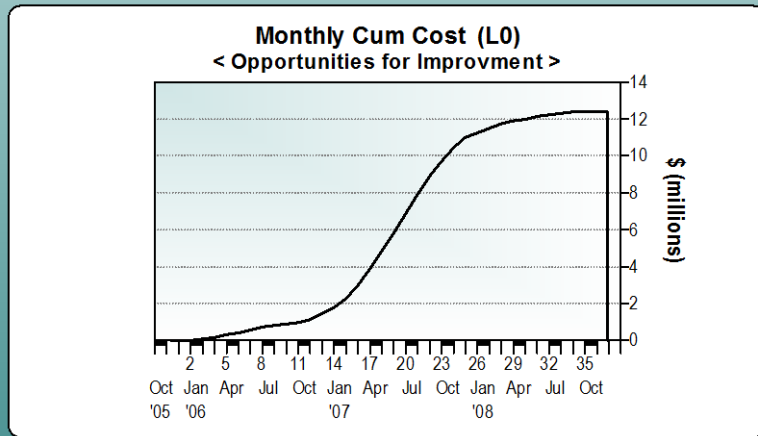
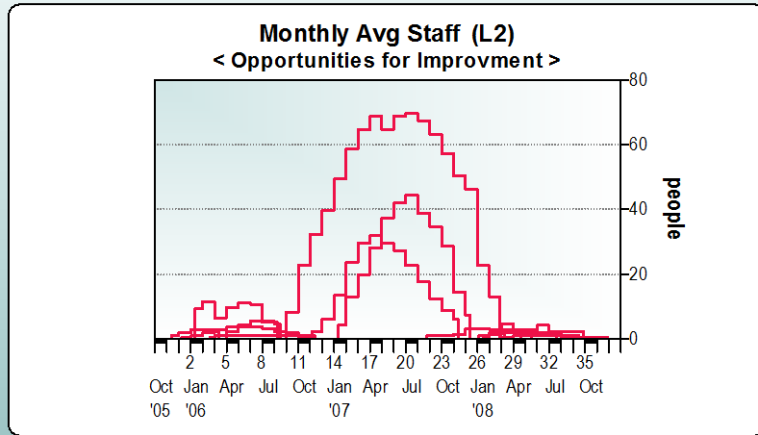
Project74



Current Portfolio of XYZ Projects Improvement Opportunities

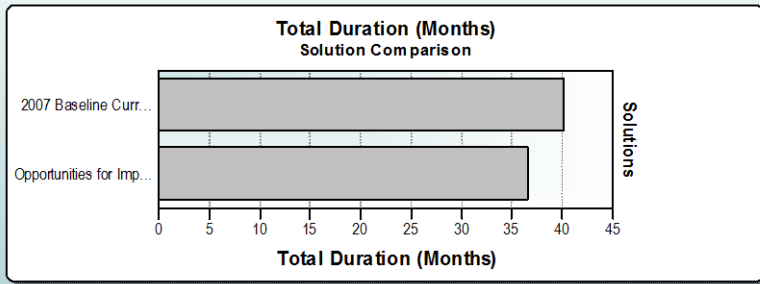
With Recommended Changes

Schedule Staffing & Cost

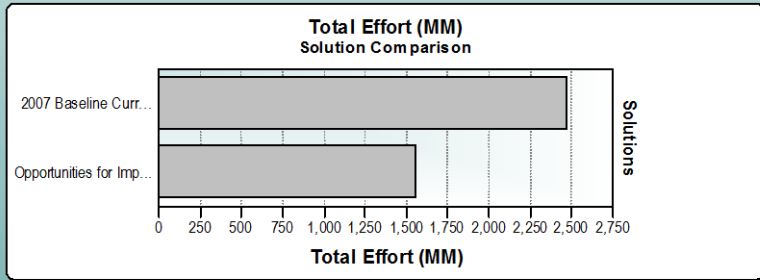


Benefit of Implementing Recommended Changes: \$7.2 Million

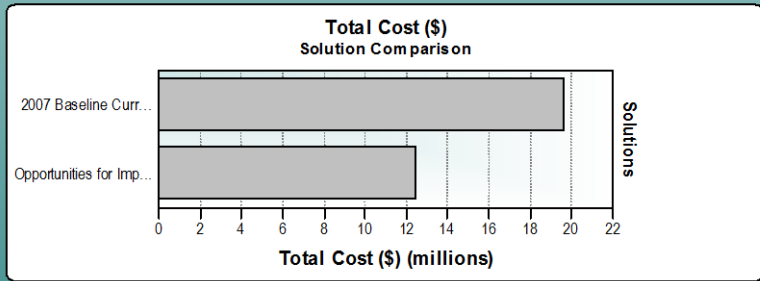
Benefits of Change



Solution	Total Duration (Months)
2007 Baseline Current State	40.05
Opportunities for Improvement	36.57



Solution	Total Effort (MM)
2007 Baseline Current State	2,469.45
Opportunities for Improvement	1,553.01



Solution	Total Cost (\$)
2007 Baseline Current State	19,625,641
Opportunities for Improvement	12,412,558

There are significant benefits to be realized by implementing the changes identified in this report.

This portfolio of projects could have been completed **3.5 months quicker with a cost saving of \$7.2 million.**



Year over Year Assessment

**XYZ 2007 Baseline
Compared to
XYZ 2006 Baseline**

Year over Year Assessment Effort

Dataset Comp - Phase 3 Effort

**Comparison of XYZ 2007 Projects to XYZ 2006
Effective SLOC vs. MB Effort (MM)**

	MB Effort (MM) Values				
	at Min Effective SLOC: 4836	at 25% Quartile Effective SLOC: 9134	at Median Effective SLOC: 21999	at 75% Quartile Effective SLOC: 119851	at Max Effective SLOC: 342258
Benchmark Reference Group: XYZ 2006	3.13	7.16	22.45	203.61	797.13
Comparison Data Set: XYZ 2007 Projects	2.52	5.70	17.65	155.97	600.83
Difference From Benchmark	-0.61	-1.45	-4.80	-47.65	-196.30

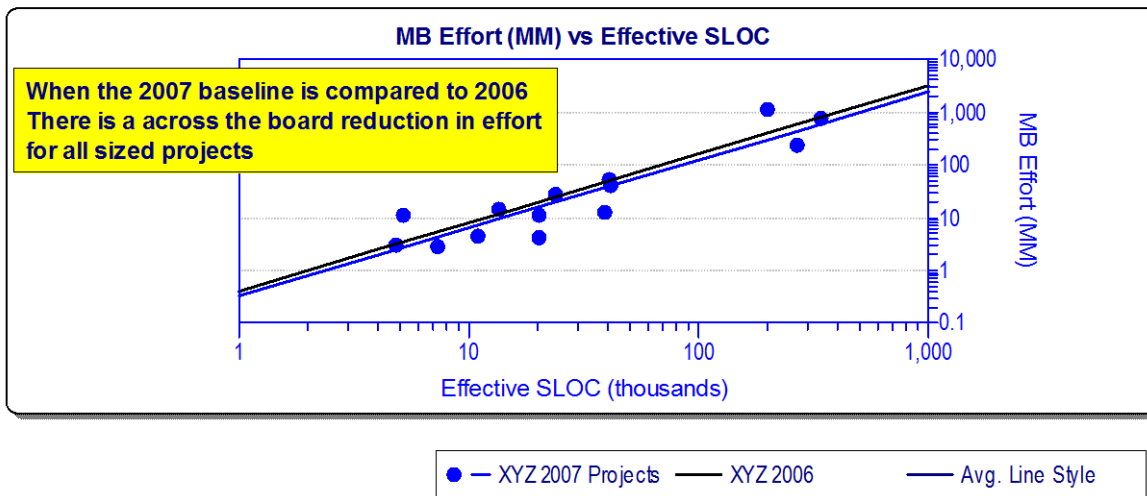
Comparison breakpoints based on min, max, median and quartile values for the data set: XYZ 2007 Projects

Year over Year Performance Effort

The black lines is the baseline that was establish last years consortium report. The time frame from that assessment is 2006. The blue line and the squares are the projects from this years assessment and are baselined at an average completion date of 2007.

The 2007 baseline parallel and shifted lower that the 2006 baseline indicating that there is a reduction in effort to produce a product across the entire size regime of projects in the portfolio.

The effort saving in manmonths is shown in the table to the right. The effort saving are very significant for the larger project.



Year over Year Assessment Duration

Dataset Comp - Phase 3 Duration

Comparison of XYZ 2007 Projects to XYZ 2006
Effective SLOC vs. MB Duration (Months)

	MB Duration (Months) Values				
	at Min Effective SLOC: 4836	at 25% Quartile Effective SLOC: 9134	at Median Effective SLOC: 21999	at 75% Quartile Effective SLOC: 119851	at Max Effective SLOC: 342258
Benchmark Reference Group: XYZ 2006	4.70	5.38	6.49	9.32	11.66
Comparison Data Set: XYZ 2007 Projects	3.39	3.91	4.76	6.96	8.81
Difference From Benchmark	-1.31	-1.47	-1.73	-2.36	-2.84

Comparison breakpoints based on min, max, median and quartile values for the data set: XYZ 2007 Projects

Year over Year Performance Schedule

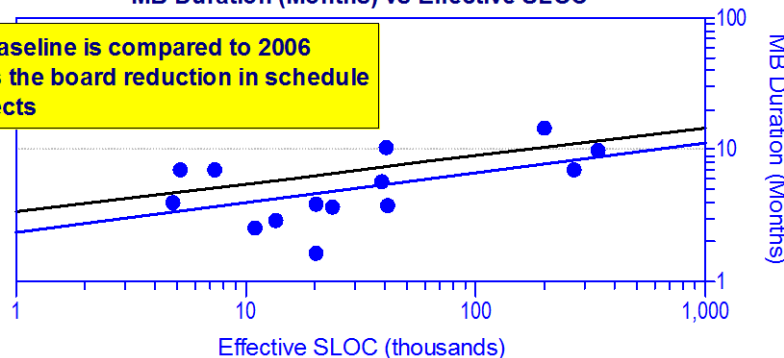
The black lines is the baseline that was established last years consortium report. The time frame from that assessment is 2006. The blue line and the squares are the projects from this years assessment and are baselined at an average completion date of 2007.

The 2007 baseline parallel and shifted lower than the 2006 baseline indicating that there is a reduction in schedule to produce a product across the entire size regime of projects in the portfolio.

The time saving in calendar months is shown in the table to the right. The time saving are approximately 1.3 to 3 months. Depending on the project size

MB Duration (Months) vs Effective SLOC

When the 2007 baseline is compared to 2006
There is a across the board reduction in schedule
for all sized projects



● — XYZ 2007 Projects — XYZ 2006 — Avg. Line Style



External Benchmark

XYZ Corporation Compared with The Financial Services Sector

Method of Analysis

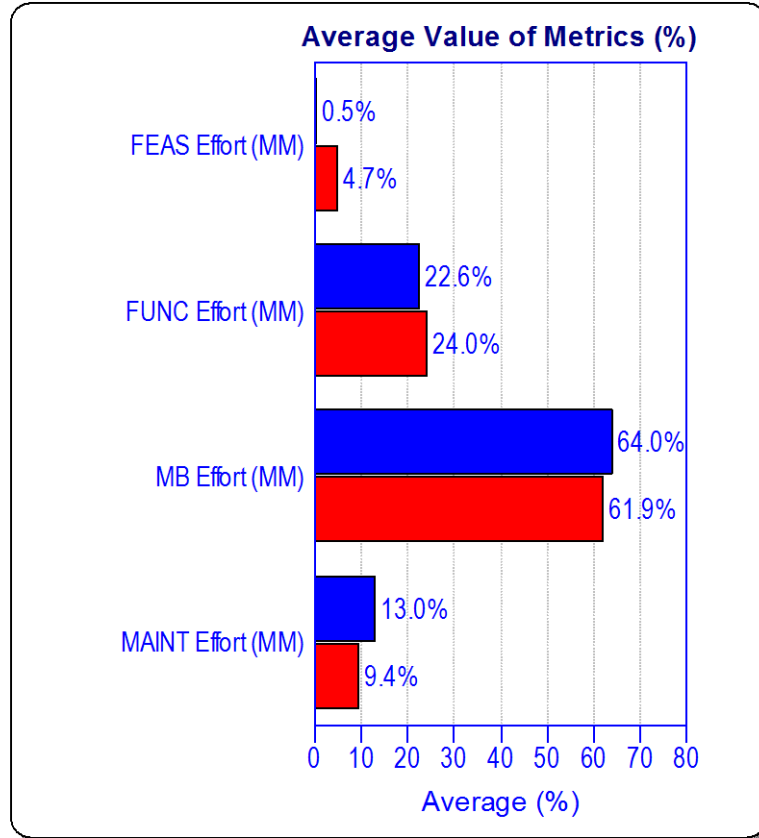
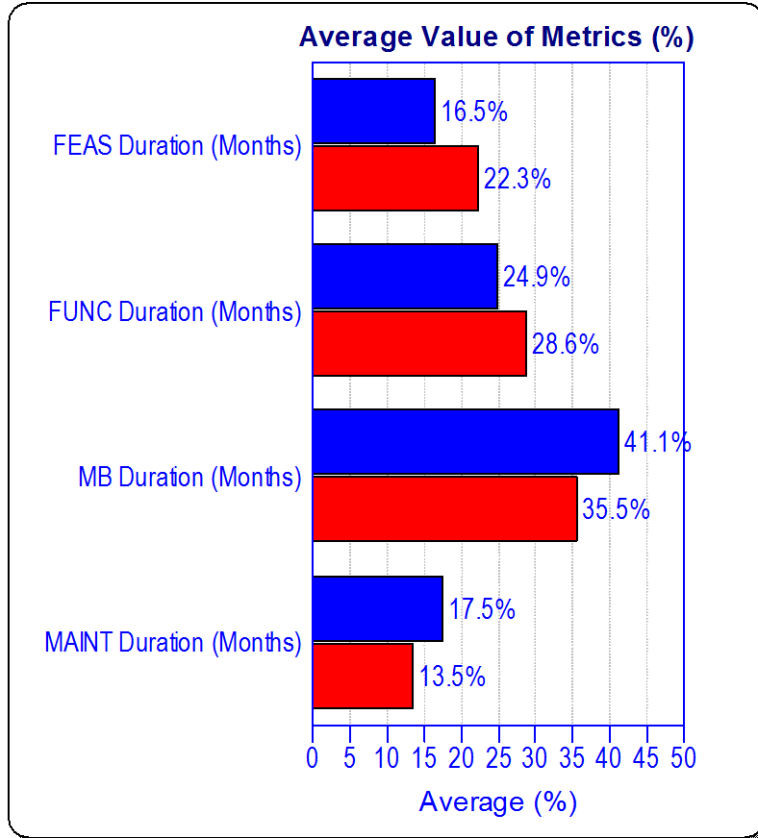
- Step 1: Select an appropriate benchmark comparison sample to be used in the comparison.
 - Completion dates from same time frame
 - Same industry sector
 - Same size regime (SLOC, Staffing, Schedules)
- Step 2: Compare positioning of XYZ corporation's projects relative to benchmark data & trend lines.
- Step 3: Produce Average Difference Comparison Reports
- Step 4: Produce 5 Star Rating Report
- Step 5: Provide Observations & Recommendations

Financial Sector Benchmark Criteria

- The following criteria were used in selecting the projects to be included in this benchmark comparison:
 - The project completion date is between 2006 and 2008.
 - The project must be an IT project developed for the financial services industry.
 - New and Modified SLOC must be greater than 3,000 and less than 500,000.
 - Average staffing must be greater than ½ a person and less than 150.
- This resulted in a comparison sample of 55 projects.

Percentage Schedule & Effort Spent in Development Phase

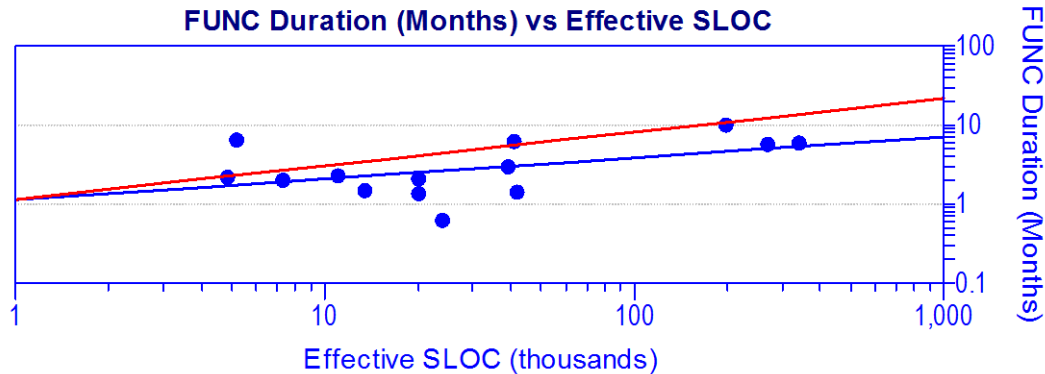
Schedule & Effort Distributions



■ Projects being Assessed ■ Financial Benchmark

Requirements & Design Phase Average Schedule Comparison

Dataset Comp - Phase 2 Duration



XYZ corporation requires less time on average to complete the Requirements and Design phase.

On the small projects the time difference is 2 weeks to 1 month. On the large projects the time difference is 5 to 8.5 months.

Comparison of Projects being Assessed to Financial Benchmark
Effective SLOC vs. FUNC Duration (Months)

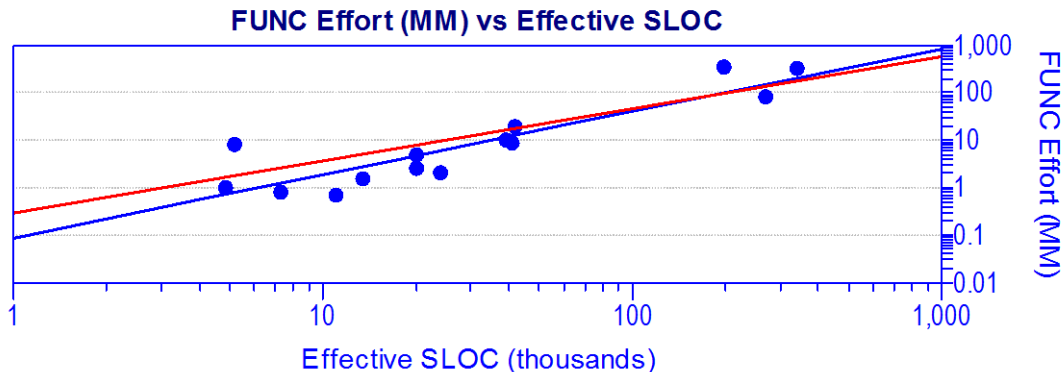
	FUNC Duration (Months) Values				
	at Min Effective SLOC: 4836	at 25% Quartile Effective SLOC: 9134	at Median Effective SLOC: 21999	at 75% Quartile Effective SLOC: 119851	at Max Effective SLOC: 342258
Benchmark Data Set:					
Financial Benchmark	2.24	2.94	4.28	8.83	13.81
Comparison Data Set:					
Projects being Assessed	1.71	2.02	2.56	4.04	5.35
Difference From Benchmark	-0.54	-0.92	-1.72	-4.79	-8.47

Comparison breakpoints based on min, max, median and quartile values for the data set: Projects being Assessed

● — Projects being Assessed — Financial Benchmark — Avg. Line Style

Requirements & Design Phase Average Effort Comparison

Dataset Comp - Phase 2 Effort



XYZ corporation requires less effort on average to complete the Requirements and Design phase except for the very large projects.

On the small projects the effort savings is 1-3 person months. On the large projects the cost is more expensive by 25 person months.

Comparison of Projects being Assessed to Financial Benchmark
Effective SLOC vs. FUNC Effort (MM)

	FUNC Effort (MM) Values				
	at Min Effective SLOC: 4836	at 25% Quartile Effective SLOC: 9134	at Median Effective SLOC: 21999	at 75% Quartile Effective SLOC: 119851	at Max Effective SLOC: 342258
Benchmark Data Set:					
Financial Benchmark	1.66	3.35	8.84	57.38	182.66
Comparison Data Set:					
Projects being Assessed	0.71	1.66	5.35	51.26	207.61
Difference From Benchmark	-0.95	-1.69	-3.49	-6.12	24.95

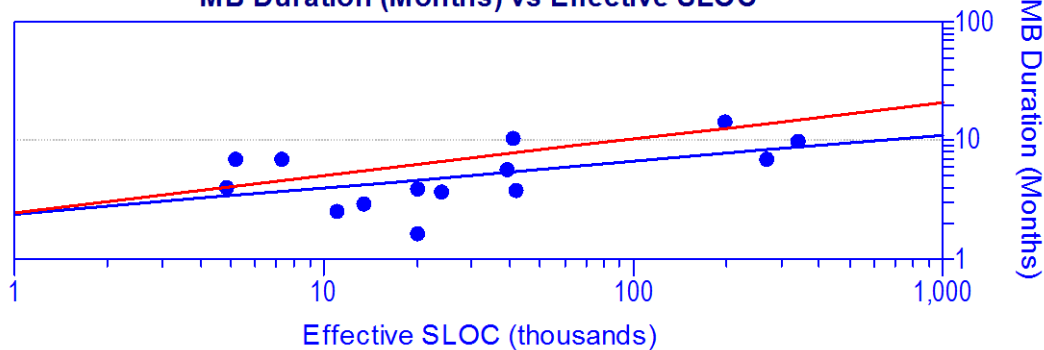
Comparison breakpoints based on min, max, median and quartile values for the data set: Projects being Assessed

● — Projects being Assessed — Financial Benchmark — Avg. Line Style

Construction & Test Phase Average Schedule Comparison

Dataset Comp - Phase 3 Duration

MB Duration (Months) vs Effective SLOC



XYZ Schedule performance is from 2 weeks to 6 months better than the benchmark during this phase.

Comparison of Projects being Assessed to Financial Benchmark
Effective SLOC vs. MB Duration (Months)

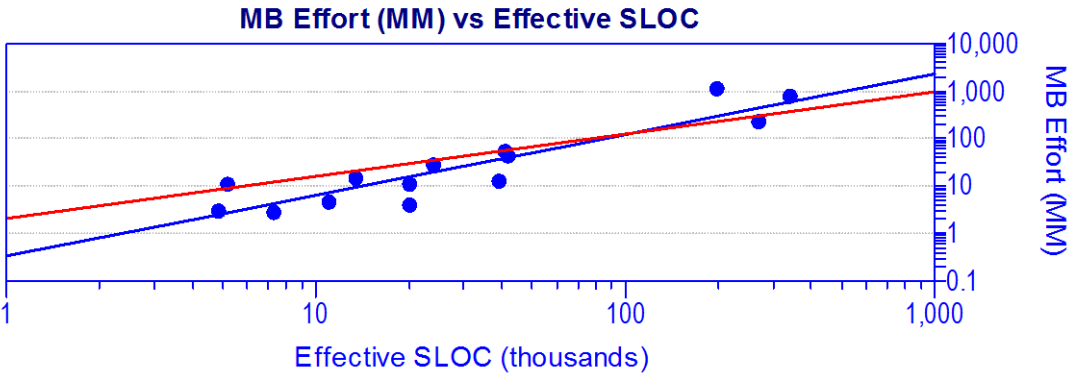
	MB Duration (Months) Values				
	at Min Effective SLOC:	at 25% Quartile Effective SLOC:	at Median Effective SLOC:	at 75% Quartile Effective SLOC:	at Max Effective SLOC:
Benchmark Data Set:					
Financial Benchmark	4836	9134	21999	119851	342258
Comparison Data Set:					
Projects being Assessed	4.02	4.90	6.44	10.92	15.13
Difference From Benchmark	3.39	3.91	4.76	6.96	8.81
	-0.63	-0.99	-1.68	-3.95	-6.32

Comparison breakpoints based on min, max, median and quartile values for the data set: Projects being Assessed

● — Projects being Assessed — Financial Benchmark — Avg. Line Style

Construction & Test Phase Average Effort Comparison

Dataset Comp - Phase 3 Effort



XYZ's effort performance varies from 17 person months better than the benchmark during this phase to 420 person months worse than the benchmark on the largest projects.

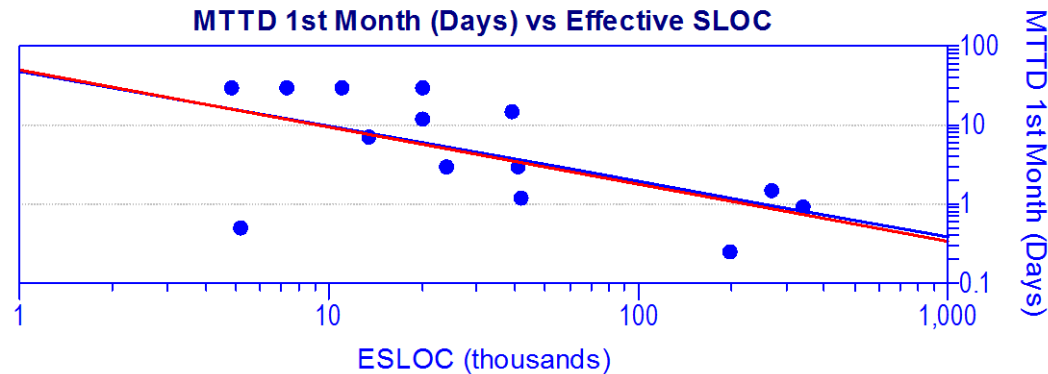
Comparison of Projects being Assessed to Financial
Effective SLOC vs. MB Effort (MM)

	MB Effort (MM) Values				
	at Min Effective SLOC:	at 25% Quartile Effective SLOC:	at Median Effective SLOC:	at 75% Quartile Effective SLOC:	at Max Effective SLOC:
Benchmark Reference Group:					
Financial	4836	9134	21999	119851	342258
Comparison Data Set:					
Projects being Assessed	2.52	5.70	17.65	155.97	600.83
Difference From Benchmark	-11.91	-15.34	-17.81	58.99	420.06

● — Projects being Assessed — Financial Benchmark — Avg. Line Style

Mean Time To Defect at Initial Delivery Average Comparison

Dataset Comp - MTTD



MTTD at delivery is essentially the same for XYZ and the benchmark.

Comparison of Projects being Assessed to Financial Benchmark
Effective SLOC vs. MTTD 1st Month (Days)

	MTTD 1st Month (Days) Values				
	at Min Effective SLOC: 4836	at 25% Quartile Effective SLOC: 9134	at Median Effective SLOC: 21999	at 75% Quartile Effective SLOC: 119851	at Max Effective SLOC: 342258
Benchmark Data Set:					
Financial Benchmark	15.97	10.07	5.33	1.56	0.73
Comparison Data Set:					
Projects being Assessed	16.11	10.32	5.58	1.70	0.82
Difference From Benchmark	0.14	0.25	0.25	0.14	0.09

Comparison breakpoints based on min, max, median and quartile values for the data set: Projects being Assessed

● — Projects being Assessed — Financial Benchmark — Avg. Line Style

5 Star Assessment Comparison Report

1 star = bottom 20%/ 2 star = bottom 20% to 45%/ 3 star = bottom 45% to top 30%/ 4 star = top 30% to top 10%/ 5 star = top 10%

Deviations Calculated From: Financial

Projects	FUNC Effort (MM)	FUNC Duration (Months)	MB Effort (MM)	MB Duration (Months)	PI	MTTD 1st Month (Days)	Composite Project Star Rating
Project 10	☆☆☆	☆☆	☆☆☆	☆☆	☆☆	☆☆	☆☆
Project 38	☆☆☆☆	☆☆☆☆	☆☆☆	☆☆☆☆	☆☆☆☆	☆☆	☆☆☆☆
Project 39	☆☆☆☆	☆☆☆☆	☆☆☆☆	☆☆☆☆	☆☆☆☆	☆☆☆	☆☆☆☆
Project 20	☆☆☆☆	☆☆☆	☆☆☆☆	☆☆	☆☆☆	☆☆☆☆	☆☆☆☆
Project 49	☆☆☆☆	☆☆☆	☆☆☆☆	☆☆☆	☆☆☆	☆☆☆☆	☆☆☆☆
Project 50	☆☆☆☆	☆☆☆	☆☆☆☆	☆☆☆☆	☆☆☆☆	☆☆☆☆	☆☆☆☆
Project 52	☆☆☆	☆☆☆☆	☆☆☆☆	☆☆☆	☆☆☆☆	☆☆☆☆	☆☆☆☆
Project 67	☆	☆☆☆☆	☆	☆☆☆	☆☆☆	☆	☆☆
Project 72	☆☆	☆☆☆☆	☆☆	☆☆☆☆	☆☆☆☆	☆☆	☆☆☆☆
Project 19	☆☆	☆☆☆☆	☆☆☆	☆☆☆☆	☆☆☆☆	☆	☆☆☆☆
Project 40	☆☆☆☆	☆☆☆☆	☆☆☆☆	☆☆☆☆	☆☆☆☆	☆☆☆☆	☆☆☆☆
Project 47	☆☆☆☆	☆☆☆☆	☆☆☆☆	☆☆☆☆	☆☆☆☆	☆☆☆☆	☆☆☆☆
Project 48	☆☆	☆	☆☆☆	☆☆	☆☆	☆	☆☆
Project 74	☆	☆☆	☆	☆☆	☆☆	☆	☆☆
Aggregate	☆☆☆	☆☆☆☆	☆☆☆☆	☆☆☆	☆☆☆☆	☆☆☆	☆☆☆

The composite rating for XYZ's portfolio is 3 stars.

- 8 of the 14 projects achieved a 4 or 5 star rating.
- To bring the overall portfolio up to a 4 star rating you need to improve the 2 stars to a 3 star level.

External Benchmark Observations

- In general XYZ corporation spends less time and effort on the small and medium sized projects. However the large projects use more people and are more expensive.
- 10 of the 14 projects perform at or above the industry benchmark performance level.
- 8 of 14 projects are exceptional performers having achieved a 4 or 5 star composite rating
- Four projects have a 2 star composite rating and are the recommended targets for improvements identified in the previous section of this report.
- The 3 largest XYZ projects expended more effort, had larger staff and found significantly more defects prior to release.

Additional Resources

- www.qsm.com/resources
 - Performance Benchmark Tables
 - Past Webinar Recordings
 - Articles and White Papers
- www.qsm.com/blog
- Follow us on Twitter at @QSM_SLIM

Keith Ciocco
keith_ciocco@qsm.com
703-749-3812

Characteristics Common to Design Projects

- Design project
 - Typically have a “what, how, do” lifecycle
 - Involve human problem solving and iterative rework cycles
- Design project common characteristics:
 - They take time to complete (calendar months)
 - They consume labor over the elapsed time schedule (labor hours)
 - They produce defects of different severity
 - They deliver some size/scope
 - **This is the metric that needs to be tailored to the design process**

What is "Size"?

- A proxy for the value and knowledge content of the delivered system—what the system is worth
- Size can be tailored to the process or development technology being used:

Front end: Unit of Need

Based on characteristics of the statement of needs

- Requirements
- Function Points/Object Points
- IO Counts
- States/Events/Actions
- Use Cases
- Stories/Story Points
- Objects/Classes
- Components
- Design Pages
- Web Pages

Back end: Unit of Work

Based on the characteristics of the system when built

- Lines of Code
- Statements
- Actions
- Modules
- Subsystems
- GUI Components
- Logic Components
- Logic Gates
- Tables
- Business Process Configurations

Size Examples for Non-Software Design Projects

Infrastructure

- Data Center
 - Servers/CPU's
 - Cabling
 - Network & connectivity configurations
 - Storage & power configurations
 - Management consoles
- Call Center
 - Call flow configurations
 - IP telephony & voicemail
 - Email/web collaboration
 - Expert agent configuration
 - Reports & CRM configuration

- Integrated Circuits (ASIC & FPGA's)
 - Logic gates
- Airframe Design
 - CAD/CAM Engineering designs/drawings
- Documentation
 - Pages
 - Chapters
 - Books/Volumes
- Systems Engineering
 - SE Requirements

Year over Year Assessment

- XYZ had trend lines established in last analysis with an average completion date of 2006.
- In the Year over Year analysis we compare the 2006 baseline trends with the 2007 baseline trends for schedule, effort and quality (Mean Time to Defect) to see what improvement, if any, has been realized.

Year over Year Assessment MTTD

Dataset Comp - MTTD

Comparison of XYZ 2007 Projects to XYZ 2006
Effective SLOC vs. MTTD 1st Month (Days)

	MTTD 1st Month (Days) Values				
	at Min Effective SLOC: 4836	at 25% Quartile Effective SLOC: 9134	at Median Effective SLOC: 21999	at 75% Quartile Effective SLOC: 119851	at Max Effective SLOC: 342258
Benchmark Reference Group: XYZ 2006	24.29	14.46	7.06	1.77	0.75
Comparison Data Set: XYZ 2007 Projects	16.11	10.32	5.58	1.70	0.82
Difference From Benchmark	-8.18	-4.14	-1.49	-0.07	0.06

Comparison breakpoints based on min, max, median and quartile values for the data set: XYZ 2007 Projects

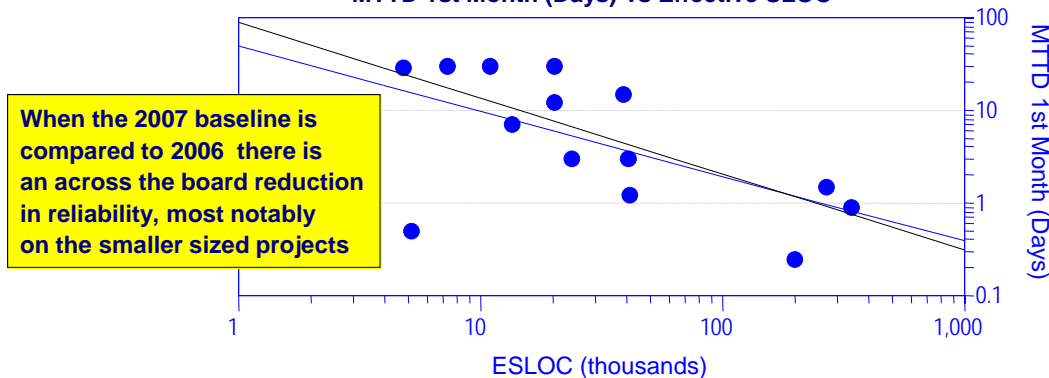
Year over Year Performance Schedule

The black lines is the baseline that was establish from last years consortium report. The time frame from that assessment is 2006. The blue line and the squares are the projects from this years assessment and are baselined at an average completion date of 2007.

The 2007 baseline parallel and shifted lower than the 2006 baseline indicating that there is a reduction in reliability at delivery across the entire size regime of projects in the portfolio.

The quality drop off is most significant on the very small projects losing anywhere from 2 to 8 days. On the large system the difference is very small.

MTTD 1st Month (Days) vs Effective SLOC



● — XYZ 2007 Projects — XYZ 2006 — Avg. Line Style