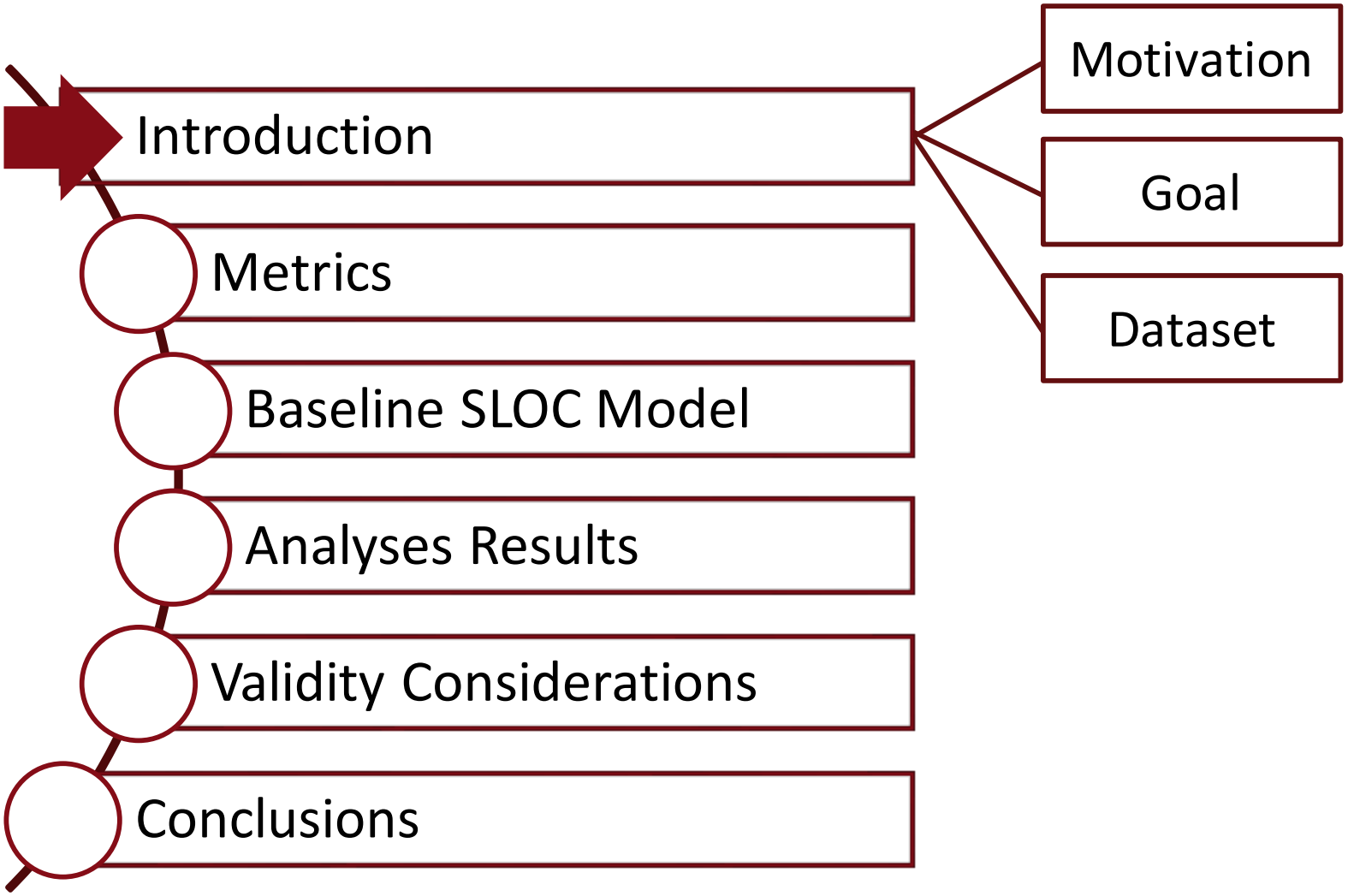
USC Viterbi School of Engineering

# Function Point Analysis for Software Maintenance

Anandi Hira and Barry Boehm

CREST Open Workshop – Predictive Models in Software Engineering: Measures, Models, and Benchmark

# Outline


Introduction
Motivation
Goal
Dataset
Metrics
Baseline SLOC Model
Analyses Results
Validity Considerations
Conclusions

# Motivation

- New development cost models > software maintenance cost models
- Source lines of code (SLOC) most common software size input
  - Difficult to estimate early in lifecycle
- Function points (FPs) represents software size by functions or modifications to functions
  - Easier to calculate earlier in lifecycle
  - Widely used to estimate effort and SLOC

# Goal

Answer following questions with empirical analysis:

1. Can Function Points effectively estimate effort for software maintenance projects?
2. Does using a Function Points to SLOC ratio add a layer of uncertainty to estimates?

# Dataset: Unified Code Count (UCC)

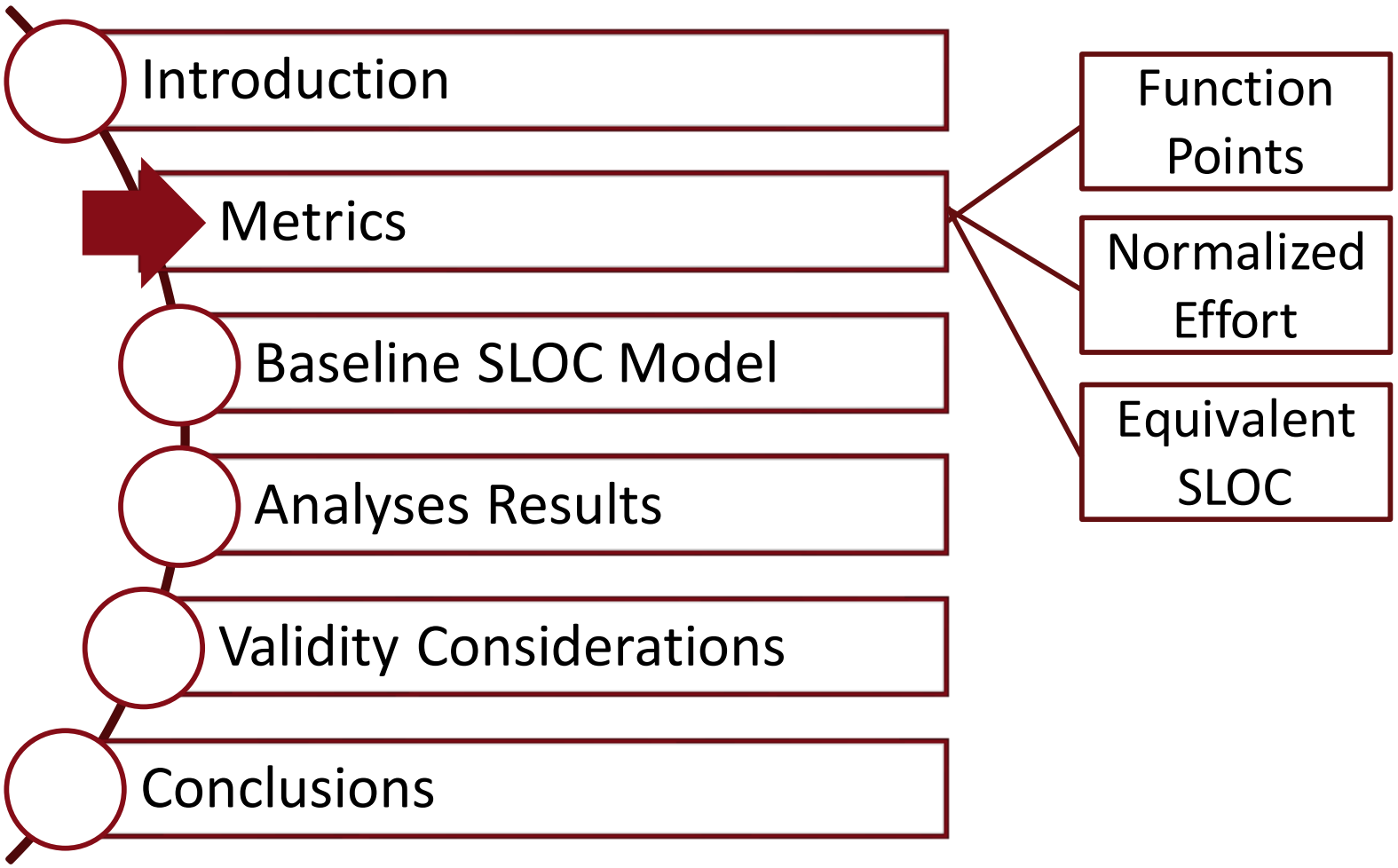
**Project Description**

- Maintained at University of Southern California (USC)
- Code metrics tool (logical SLOC, cyclomatic complexity)
- Implemented in C++
- 45 to 1425 logical SLOC
- 2010 to 2014
- Modularized architecture
- 4-month time-boxed increments

**Project Types**

- Add Functions
  - New language parsers
  - New features, such as GUI front-end
- Modify functions
  - Cyclomatic complexity support (modify existing language parsers with mathematical operation and algorithms)

# Outline


Introduction
Metrics
Baseline SLOC Model
Analyses Results
Validity Considerations
Conclusions
Function Points
Normalized Effort
Equivalent SLOC

# Function Points – 1/2

| Type of Component | Complexity of Components Multiplier Factor | | | |
|---|---|---|---|---|
| | Low | Average | High | Total |
| External Inputs | 3 | 4 | 6 | |
| External Outputs | 4 | 5 | 7 | |
| External Inquiries | 3 | 4 | 6 | |
| Internal Logical Files | 7 | 10 | 15 | |
| External Interface Files | 5 | 7 | 10 | |
| Total Number of Unadjusted Function Points | | | | |

# Function Points – 2/2

**General System Characteristics**

- Data communications
- Distributed data processing
- Performance
- Heavily used configuration
- Transaction rate
- Online data entry
- End-user efficiency
- Online update
- Complex processing
- Reusability

**General System Characteristics Cntd.**

- Installation ease
- Operational ease
- Multiple sites
- Facilitate change

**Equations: Value Adjustment Factor (VAF), Enhancement Project FP (EFP)**

$$VAF = 0.65 + \frac{(\sum C_i)}{100}$$

$$EFP = [(ADD + CHGA) * VAFA] + (DEL * VAFB)$$

# Normalized Effort

COCOMO model:

$$\text{Effort (PM)} = 2.94 \times Size^{1.0997} \times \prod_{i=1}^{17} EM_i$$

$$\text{PM} = 152 \text{ hours}$$

$$\text{Normalized Effort (hours)} = \frac{\text{Total Effort (hours)}}{\left(\prod EM_i\right)}$$

# Equivalent SLOC (ESLOC)

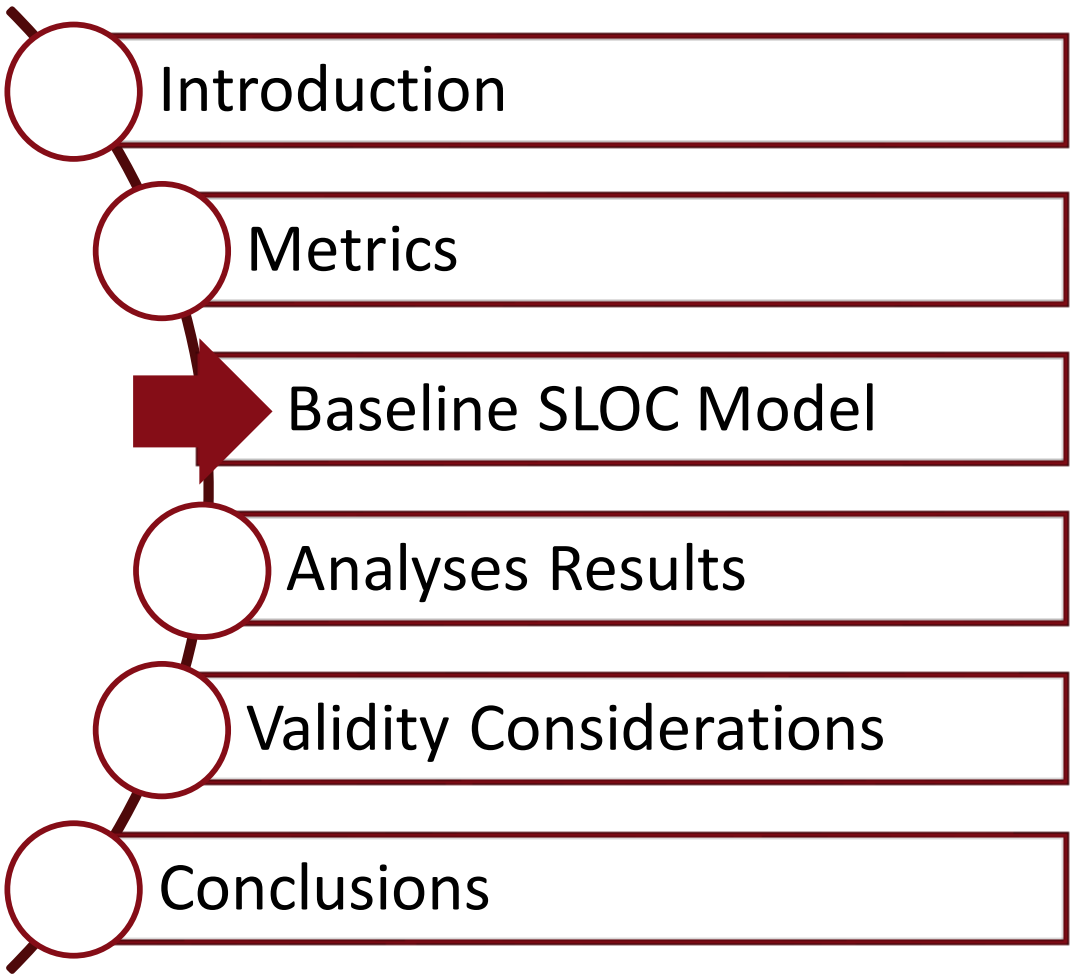
$$AAF = 0.4 * DM + CM + 0.3 * IM$$

$$AAM = \begin{cases} \dfrac{AA + AAF + \left(1 - \left(1 - \dfrac{AAF}{100}\right)^2\right) * SU * UNFM}{100} & \text{if } AAF \le 100 \\ \dfrac{AA + AAF + SU * UNFM}{100} & \text{if } AAF > 100 \end{cases}$$

$$ESLOC = SLOC_{added} + (ASLOC * AAM)$$

| | |
|---|---|
| **DM** | Design Modification |
| **CM** | Code Modification |
| **IM** | Integration and Test |
| **SU** | Software Understanding |
| **UNFM** | Programmer Unfamiliarity |
| **AA** | Assessment and Assimilation |

# Outline

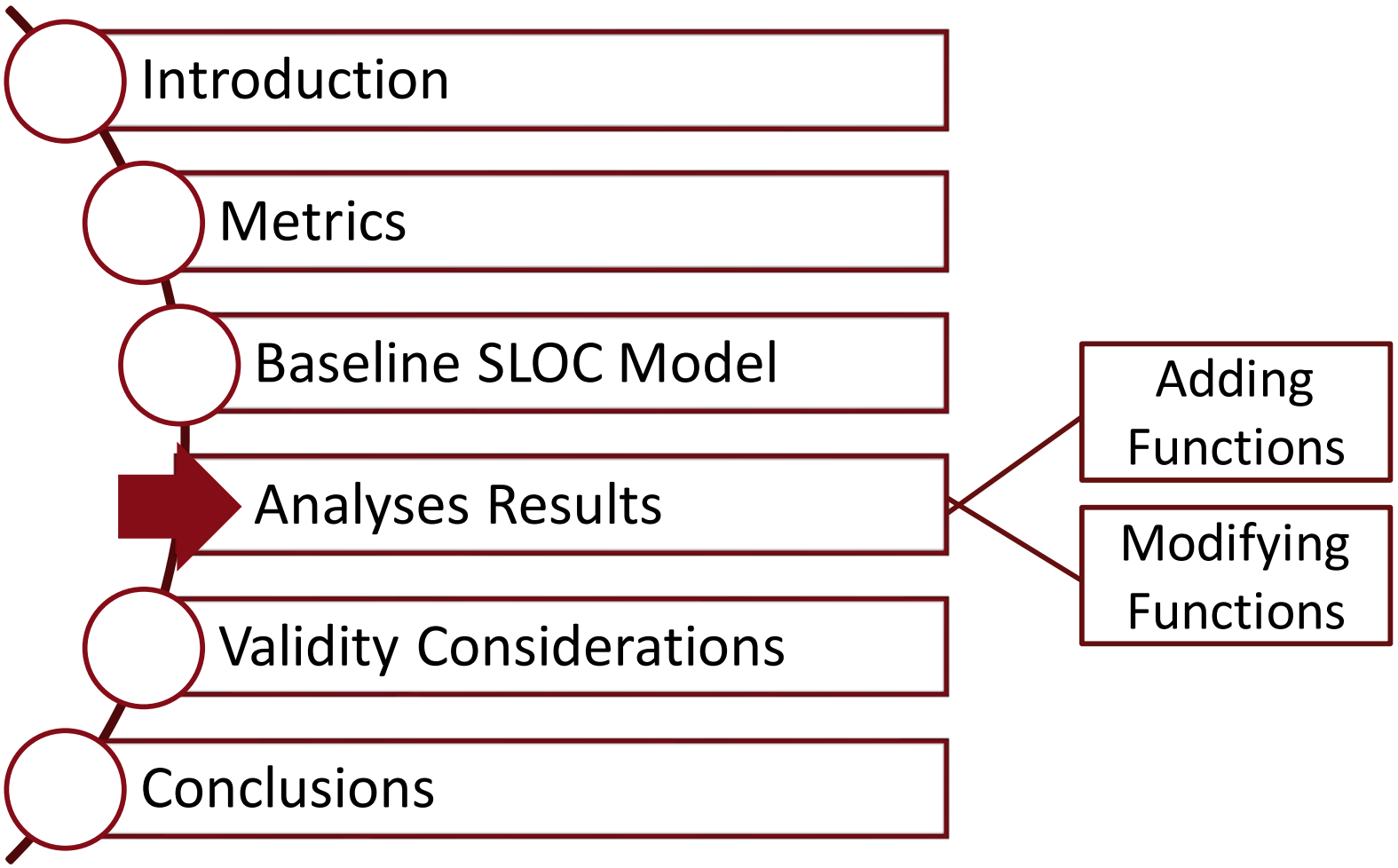

Introduction
Metrics
Baseline SLOC Model
Analyses Results
Validity Considerations
Conclusions

# Baseline SLOC Model

$$Effort = 446.88 \times \left(\frac{ESLOC}{1000}\right)^{1.0997} \times \prod_{i=1}^{17} EM_i$$

| | |
|---|---|
| **$R^2$** | 90% |
| **PRED (20)** | 70% |
| **PRED (25)** | 70% |
| **PRED (30)** | 89% |

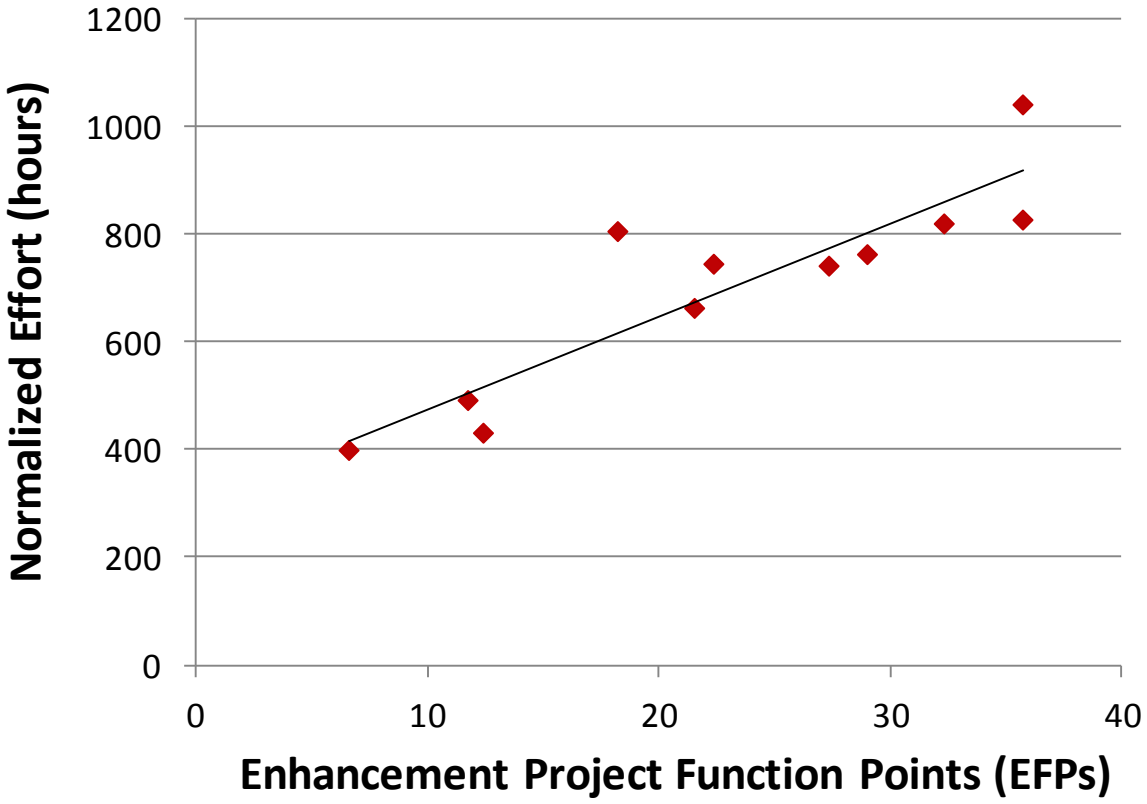
# Outline


Introduction
Metrics
Baseline SLOC Model
Analyses Results
Adding Functions
Modifying Functions
Validity Considerations
Conclusions

# RESULTS: ADDING FUNCTIONS

# FPs vs Normalized Effort

Normalized Effort = 302.06 + (17.312 × EFP )



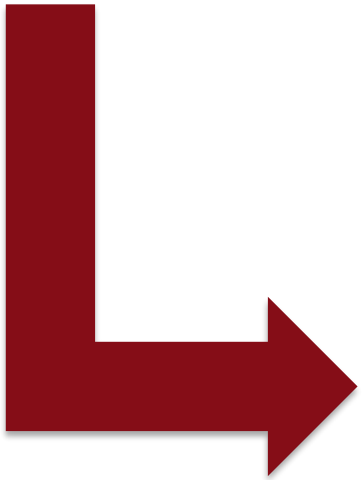

| | |
|---|---|
| **$R^2$** | 80% |
| **PRED (20)** | 82% |
| **PRED (25)** | 82% |
| **PRED (30)** | 100% |

# FPs vs ESLOC

$$ESLOC = -26.081 + (13.607 \times EFP) + (19.316 \times \#\ modified\ modules)$$

## ESLOC Estimates

| | |
|---|---|
| **$R^2$** | 90% |
| **PRED (20)** | 63% |
| **PRED (25)** | 75% |
| **PRED (30)** | 75% |



ESLOC estimates with Baseline SLOC model

## Effort Estimates

| | |
|---|---|
| **PRED (20)** | 50% |
| **PRED (25)** | 50% |
| **PRED (30)** | 75% |

# RESULTS: MODIFYING FUNCTIONS

# FPs vs Normalized Effort

Normalized Effort (hours) =
80.987
– (1.027 × CHGA)
+ (2.433 × CHGB)
– (22.485 × # modified files)
+ (20.703 × # new modules)
+ (24.909 × # modified modules)

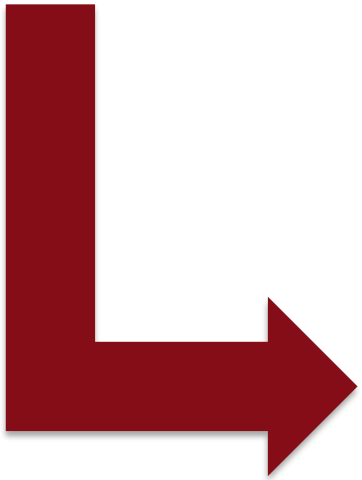
| | |
|---|---|
| **$R^2$** | 73% |
| **PRED (20)** | 21% |
| **PRED (25)** | 47% |
| **PRED (30)** | 63% |

# FPs vs ESLOC

$$ESLOC = 103.630 + 0.396 \times EFP^{[1.336 + (-0.039 \times \#\text{modified files}) + (0.088 \times \#\text{ new modules})]}$$

## ESLOC Estimates

| | |
|---|---|
| **$R^2$** | 71% |
| **PRED (20)** | 38% |
| **PRED (25)** | 38% |
| **PRED (30)** | 62% |

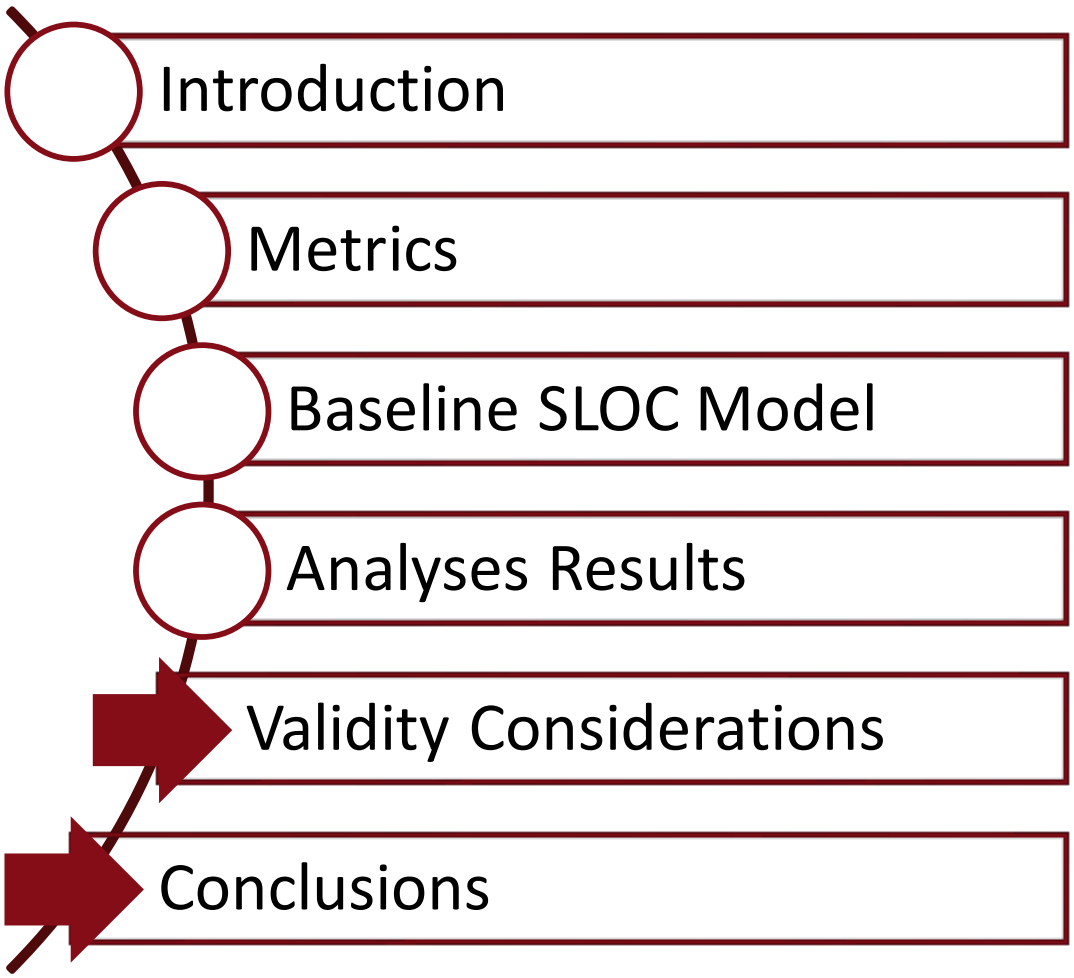


ESLOC estimates with Baseline SLOC model

## Effort Estimates

| | |
|---|---|
| **PRED (20)** | 19% |
| **PRED (25)** | 19% |
| **PRED (30)** | 24% |

# Outline


Introduction
Metrics
Baseline SLOC Model
Analyses Results
Validity Considerations
Conclusions

# Validity Considerations

**Internal**

- Reported effort may not be accurate
  - Forget to update timesheets
  - Show high productivity
  - Show excessive hours

**Mitigation**

- Members evaluated on ability to meet deadlines, adapt to problems, communicate clearly

**External**

- Segregation between adding functions and modifying functions
  - Test and verify on other datasets
- Linear relationship does not account for diseconomies of scale
  - Analyze on datasets with larger projects for scalable results

# Conclusions

- Function Points effective to estimate affect?
  - Add functions: Yes!
  - Modify functions: No (even with additional metrics)
- Function Points to SLOC ratio effective for effort estimates?
  - Adds layer of uncertainty
  - Resulting effort accuracy unsatisfactory