NATIONAL RECONNAISSANCE OFFICE

# Software Maintenance Cost Estimating Relationship Development for Space Systems*

*This is actually an EER (effort estimating relationship) but will be referred to as a CER throughout this document.

Ryan Timm – Booz Allen Hamilton

# About the NRO

- The National Reconnaissance Office (NRO) is the national program to meet the U.S. Government's intelligence needs through spaceborne reconnaissance
- A Department of Defense (DoD) agency and an element of the Intelligence Community
- Funded through the National Intelligence Program and the Military Intelligence Program portions of the federal budget
- The NRO's existence was declassified by the Deputy Secretary of Defense on September 18, 1992

NRO and DoD acquisition processes are closely aligned

# What is Echelon 2 Software Maintenance?

- The objectives of Echelon 2 (E2) Software (SW) maintenance functions are to perform corrective, perfective, and adaptive maintenance actions
- Develop, verify, and distribute configuration changes to the product baseline resulting from Discrepancy Reports (DR) actions
- Develop, integrate, verify, and distribute applications software upgrades and enhancements as authorized by approved Requests for Changes (RFCs)
- Maintain configuration control of the operational Computer Software Configuration Items (CSCIs)

# Why are we estimating effort not cost?

- Effort in FTEs is used to estimate software maintenance as labor cost is highly variable and would create a noisy dataset due to well understood variables such as:
  - Geographic location
  - Overhead (Factory site vs. government site)
  - Contractor
  - Wage inflation
- We have an extensive database of labor rates for a variety of contractors and locations

# CAAG Software Maintenance Estimating History

- Prior to 2009 rules of thumb for E2 SW maintenance used (Source Lines of Code (SLOC)/maintainer)
- Circa 2009 NRO CAAG team improved on rule of thumb by using data to create a curve of Full Time Equivalents (FTE)/KSLOC vs. age that shows decreased effort with age. Due to the limited data available this curve is more an analogy than CER.
- The current effort expands upon the 2009 work to include additional data points and explores additional cost drivers

# Potential Cost Drivers

- Size of code being maintained (logical SLOC)
- Years since first maintained (proxy for age of code and maturity of organizational processes)
- Age of code (not available in current dataset)
- Mission function (e.g. Command and Control ($C^{2)}$ vs. Mission Processing (MP))

# E2 SW Maintenance CER 2013 Update

- Added new data points from additional programs
- Program data used in total to prevent one program from dominating the dataset
- 15 different CER regressions attempted using Log Ordinary Least Squares (LOLS) and Zero-bias Minimum Percent Error (ZMPE)
- Developed CERs with SLOC scale variable and Years system is operational
- Developed stratified CER with Mission Processing dummy variable
- Developed CERs with continuous time and stratified into new / old (>2 years)

# Data Availability

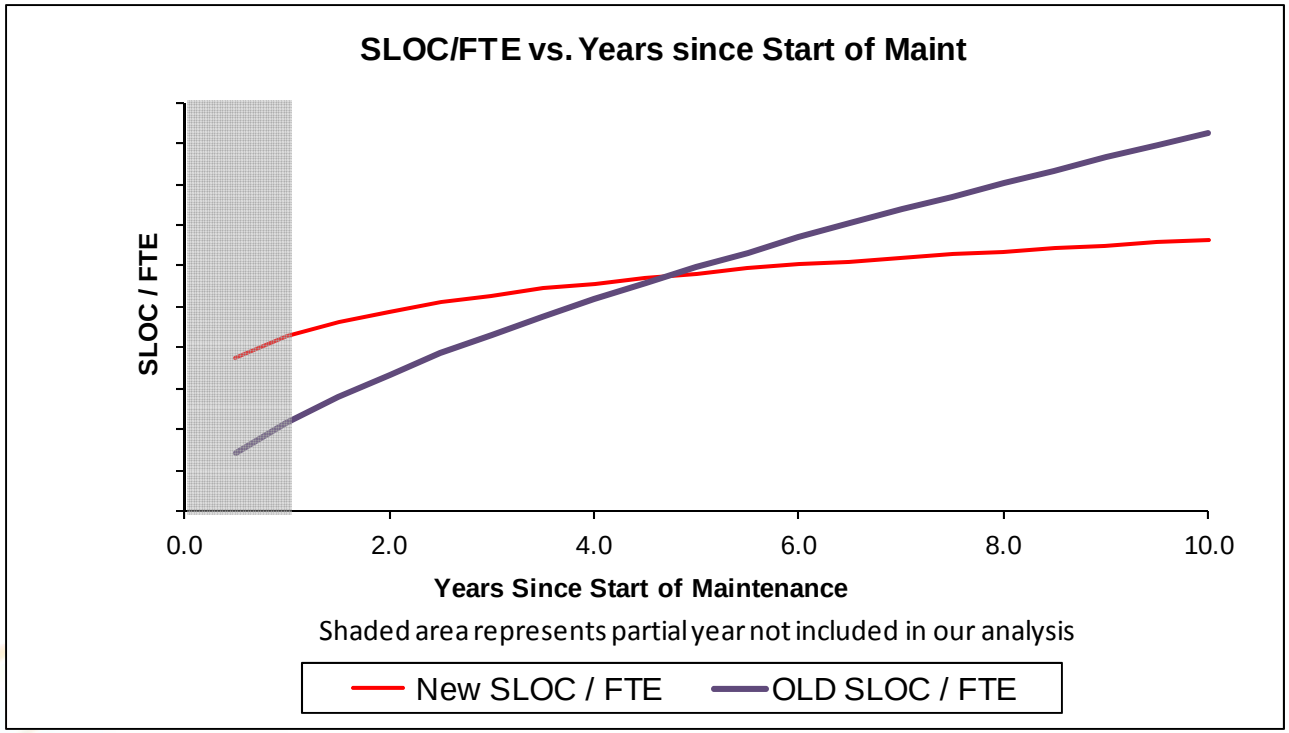
- NRO CAAG has made data collection a priority
- Data comes directly from contractors with data collection CDRLs on all new ground contracts
- NRO Standard O&M WBS established in 2011 assists data collection of E2 Software Maintenance Data
- Over the next 2-years NRO CAAG will more than double our E2 SW maintenance dataset

# Data Issues

- A valid data point must segregate E2 Maintenance effort from ongoing development and have a corresponding code count
- Mission Processing programs in our data set tend to be larger than C2 and may be creating an artificial economies of scale in the regression equation
- Age since start of maintenance is only a rough proxy for age of code as it fails to identify the mix of new code added to the baseline, old code deleted from the baseline, or the maturity of reused code
- Age of code is correlated with size and may cause harmful multicollinearity in the regression equation

# Prior E2 SW Maint CER

- Prior CER had functional form: FTE/kSLOC = aYears$^c$
- When new data is added to prior CER the average bias is -15% and $R^2$ declines to 57% while SPE remains low at 38%
- Additional data suggests old CER could be improved

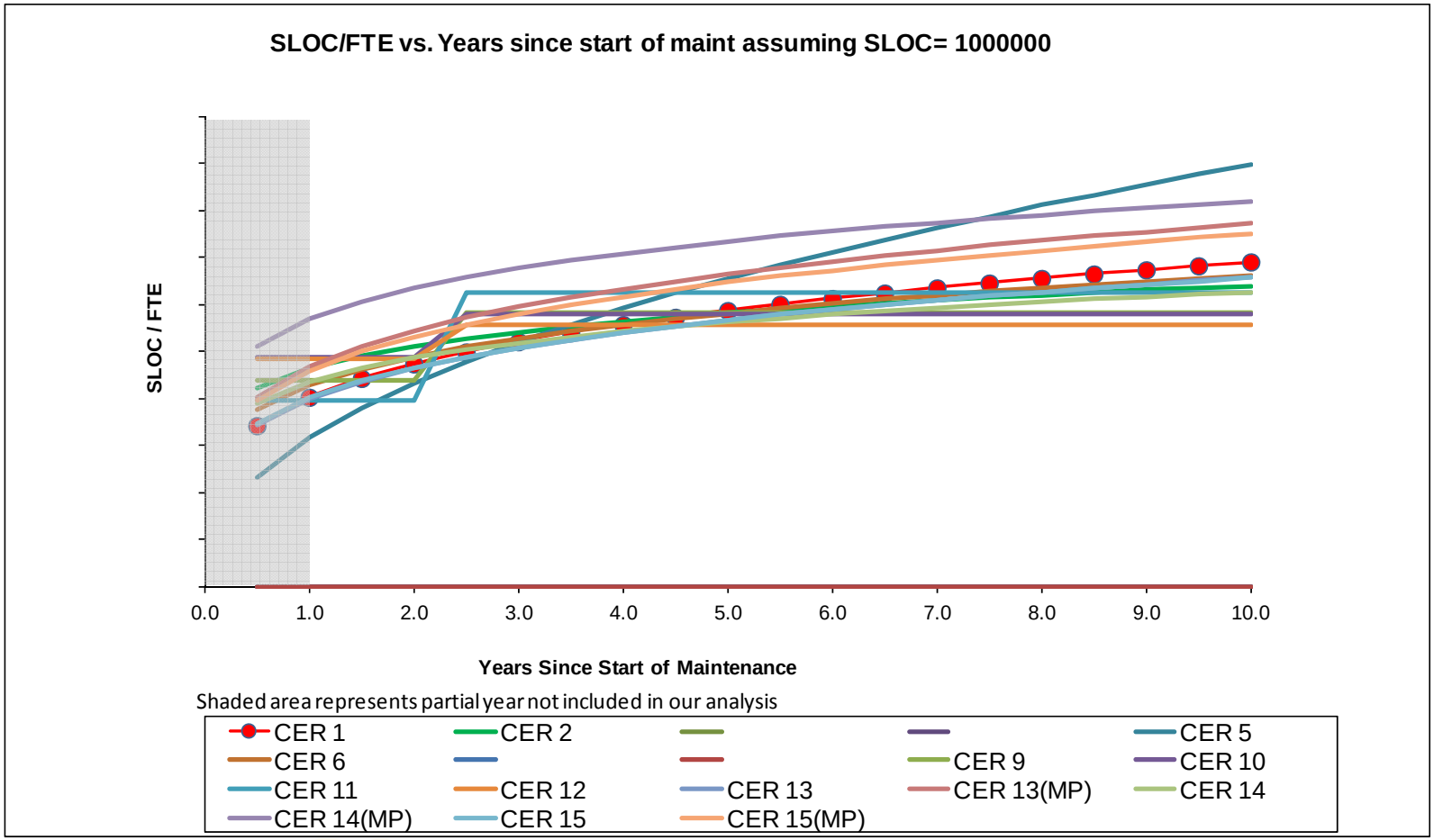


# E2 SW Maintenance Regressions

| # | CER Functional Form | Regression Method | $R^2$ | SPE | Bias | Comments |
|---|---|---|---|---|---|---|
| OLD | $FTE/kSLOC = aYears^c$ | LOLS | 56.5% | 37.7% | -14.9% | OLD CER with new Data |
| 1 | $FTE = aSLOC^bYears^c$ | ZMPE | 69.7% | 28.6% | 0.0% | Recommended - Lowest SPE, $R^2$ near .7 improves on old CER by allowing SLOC to have an exponent |
| 2 | $FTE = aSLOC^bYears^c$ | LOLS | 74.1% | 29.3% | -0.1% | Good Stats however model has large economies of scale and relatively little age impact |
| 3 | $FTE = aSLOC^b$ | ZMPE | 71.6% | 31.9% | 0.0% | Good Statistics, based only on size |
| 4 | $FTE = aSLOC^b$ | LOLS | 71.7% | 31.9% | 0.0% | Good Statistics, based only on size |
| 5 | $FTE = aSLOCYears^c$ | ZMPE | 68.3% | 38.5% | 0.0% | CER1 shows SPE improves when SLOC has an exponent |
| 6 | $FTE = aSLOCYears^c$ | LOLS | 78.6% | 41.9% | -0.1% | Update of OLD CER - Poor SPE relative to #1 |
| 7 | $FTE = a+bSLOC^cYears^d$ | ZMPE | 69.7% | 29.6% | 0.0% | A solves to zero making this form equivalent to #2 |
| 8 | $FTE= a+bSLOCYears^c$ | ZMPE | 72.9% | 49.2% | 0.0% | A solves to zero making this form equivalent to #5 |
| 9 | $FTE = aSLOC^bc^{new}$ | ZMPE | 65.7% | 29.1% | 0.0% | New Old stratifier does not improve goodness of fit compared to #1 |
| 10 | $FTE = aSLOC^bc^{new}$ | LOLS | 69.6% | 29.7% | -0.1% | New Old stratifier does not improve goodness of fit compared to #1 |
| 11 | $FTE = aSLOCc^{new}$ | ZMPE | 67.2% | 41.4% | 0.0% | SPE is worse when SLOC exponent is fixed at 1 |
| 12 | $FTE = aSLOCc^{new}$ | LOLS | 73.1% | 45.1% | -0.1% | SPE is worse when SLOC exponent is fixed at 1 |
| 13 | $FTE = aSLOC^bYears^c*d^{MP}$ | ZMPE | 68.1% | 30.1% | 0.0% | MP Stratifier comes out to a reasonable value but $R^2$ and SPE do not improve |
| 14 | $FTE = aSLOC^bYears^c*d^{MP}$ | LOLS | 68.5% | 30.8% | 0.0% | MP Stratifier comes out to a reasonable value but $R^2$ and SPE do not improve |
| 15 | $FTE = aSLOC^{(b+(MP*d))}Years^c$ | ZMPE | 67.9% | 30.2% | 0.0% | MP variable has little impact |

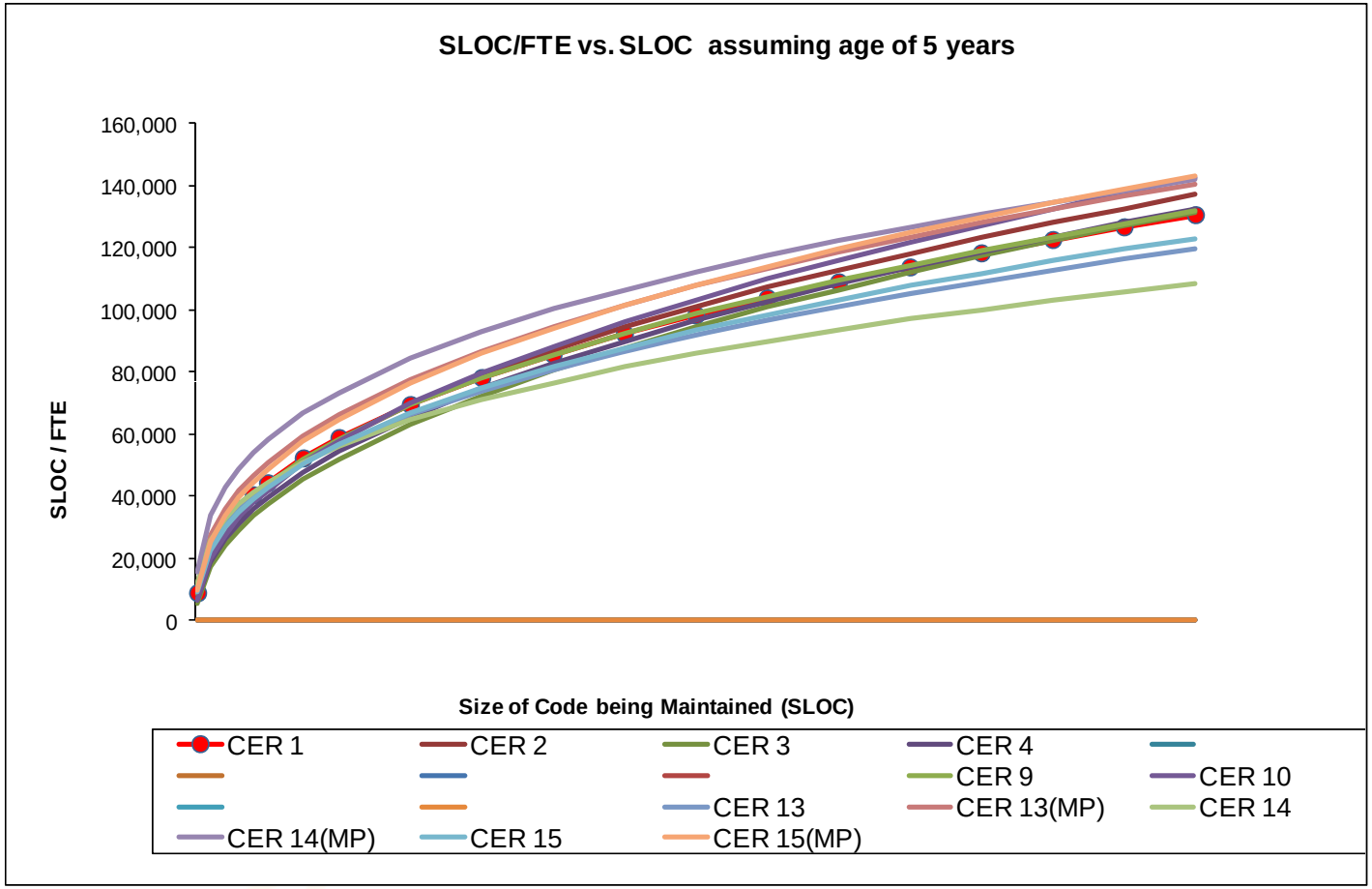
- #1 Overall best SPE and good $R^2$
- #9 Best candidate with new/old stratifier
- #13 Best candidate with MP stratifier
- #7&8 Solver drops additive term effectively eliminating this functional form

# E2 SW Maintenance Regressions vs. Time




- Regressions with similar statistics provide a wide range of results when graphed against time assuming 1M SLOC
- CERs without time variables are excluded

# E2 SW Maintenance Regressions vs. Size



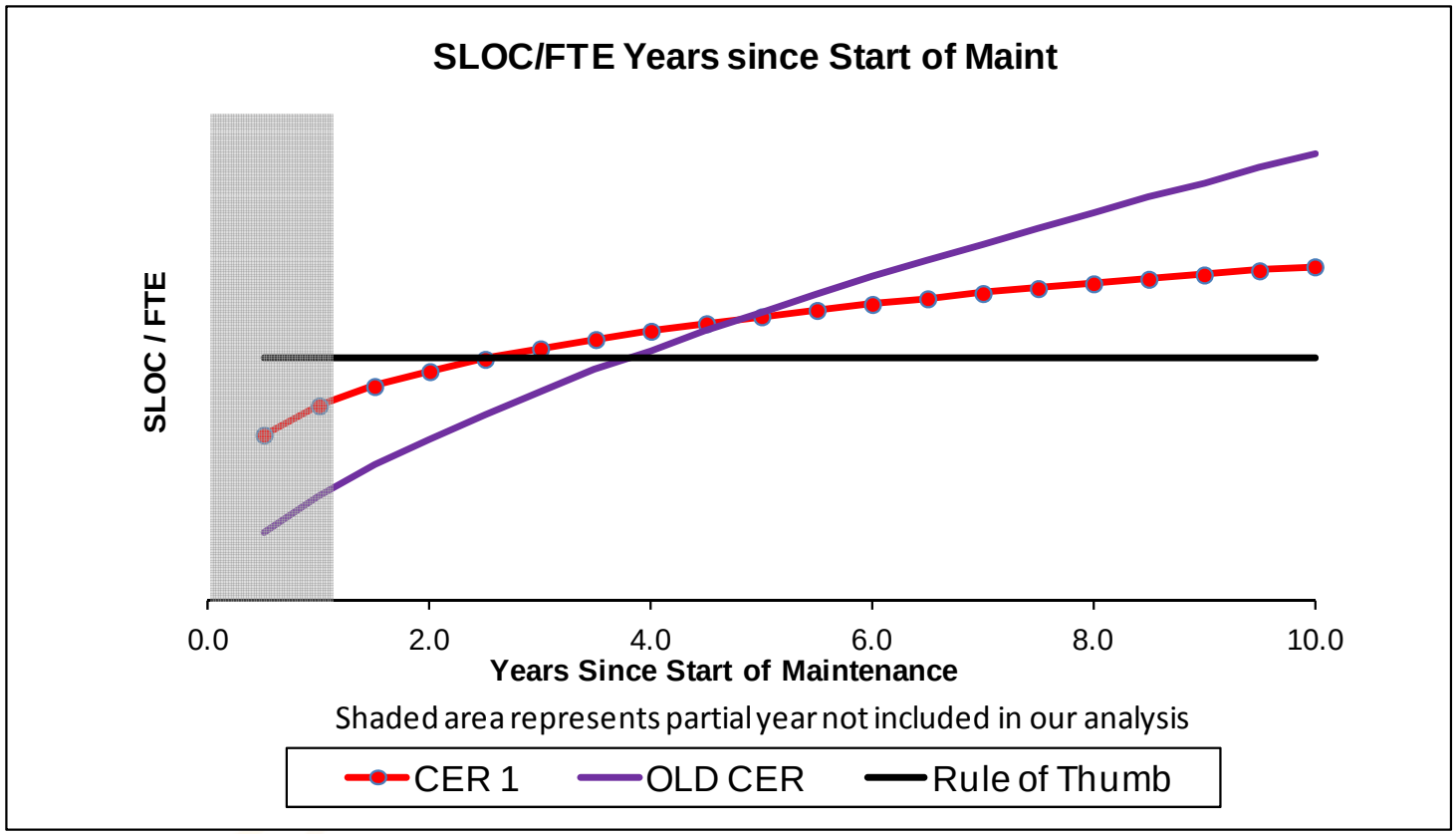

- Regressions with similar statistics provide a wide range of results when graphed against code size assuming 5 years age
- CERs not sensitive to size were excluded from the graph

# Candidate CER (#1)

| # | CER Functional Form | Regression Method | $R^2$ | SPE | Bias | Comments |
|---|---|---|---|---|---|---|
| 1 | $FTE = aSLOC^{b}Years^{c}$ | ZMPE | 69.7% | 28.6% | 0.0% | Recommended - Lowest SPE, $R^2$ near .7 improves on old CER by allowing SLOC to have an exponent |

- Reasonable economies of scale with size growth
- Reasonable maintenance efficiencies with age
- Regression statistics: SPE is the best of attempted CER functional forms. $R^2$ is among the best of attempted CER functional forms.

# Candidate CER (#1)




For CER 1 Assumed 1M Logical SLOC

# Candidate CER (#13)

| # | CER Functional Form | Regression Method | $R^2$ | SPE | Bias | Comments |
|---|---|---|---|---|---|---|
| 13 | $FTE = aSLOC^{b}Years^{c}*d^{MP}$ | ZMPE | 68.1% | 30.1% | 0.0% | MP Stratifier comes out to a reasonable value but $R^2$ and SPE do not improve |

- Segregates C2 from MP
- Reasonable economies of scale with size growth
- Reasonable maintenance efficiencies with age
- MP effort is modeled at ~85% of C2
- Regression statistics: $R^2$ and SPE among the best of attempted CER functional forms

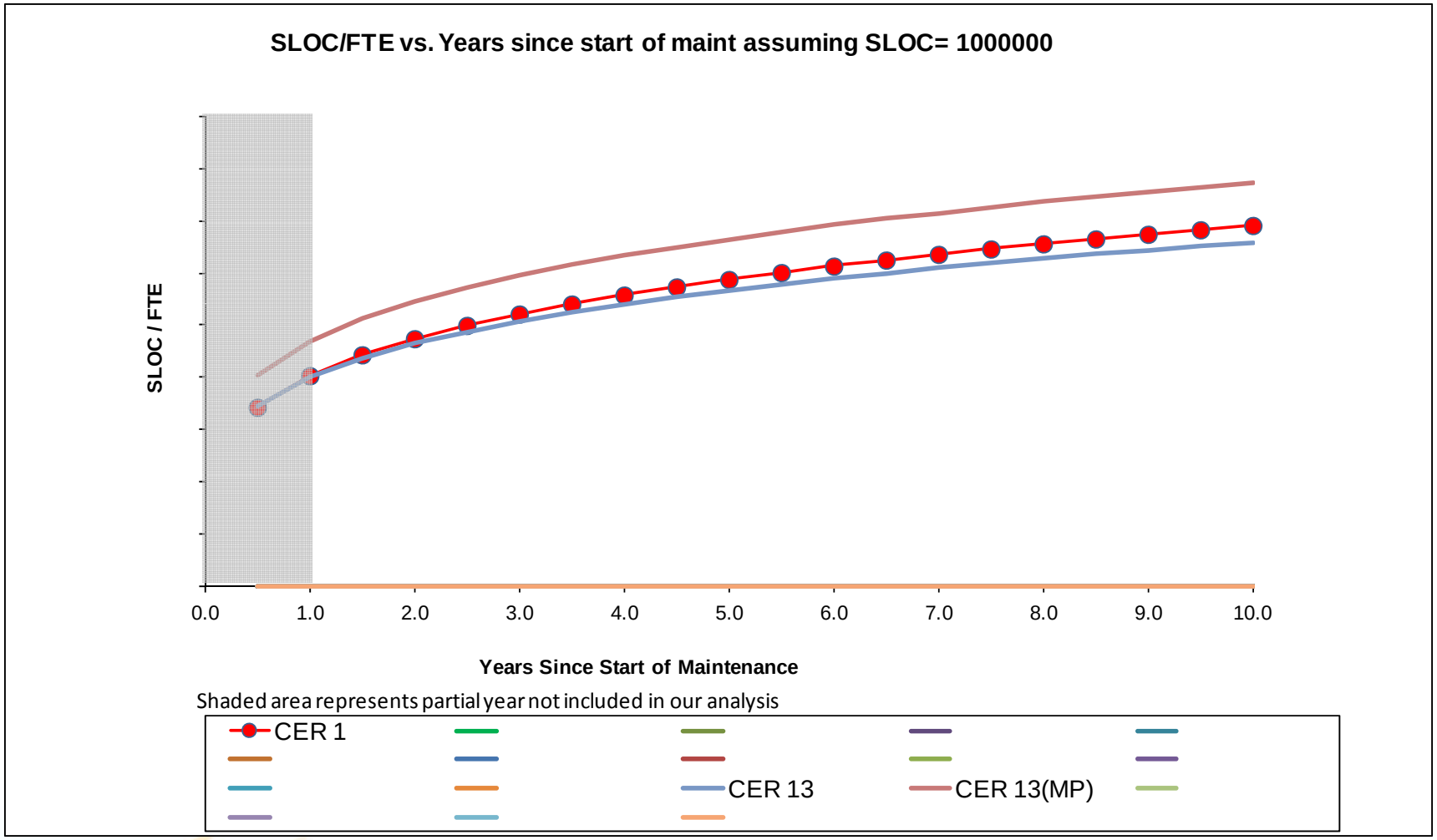
# To stratify or not to stratify? Pro

- Engineering judgment that C2 SW maintenance is less productive than other ground software functions such as MP
  - C2 has rigorous testing to ensure the health and safety of the vehicle is preserved
  - C2 requires precise and complex timing
- Our software development metrics indicate that C2 programs have lower productivity rates
- MP stratifier remained stable throughout Influential Data Point tool runs

# To stratify or not to stratify? Con

- Few data points for MP (adding another could dramatically influence results)
- MP points in CER #1 were not uniformly underestimated
- MP data points had average bias of -9.44% which is reasonable for the sample of data points (similar magnitude bias observed from randomly picking any data points)

# Candidate CERs (#1 and #13)



# Summary

- The recommended CER is a better model that the prior one due to:
  - Additional data points with much greater variety in both size and function
  - Economies of scale is revealed as an important driver of cost
  - Updated data reveals old CER reduces maintenance effort too much with age
- The NRO CAAG accepted CER #1 to be used to estimate E2 SW Maintenance for new future ground systems
- When estimating existing systems extrapolation from actuals should be considered a viable alternative to the CER on a case by case basis

# Recommended CER Implementation

- This CER provides an FTE estimate for a single year and should be modeled for each year of O&M (not a single estimate spread with a phasing model)
- Each year has its own distribution and can be modeled in @risk as such
  - Correlation between distributions should be extremely high
- Logical SLOC tends to increase with age due to DR fixes and enhancements
  - Analyst should not assume a static code value across all years

# Next Steps

- Continue to add data points as they become available and release CER to public once data from multiple sources can be included
- Explore stratification for function of code when enough data is available for each type (e.g. C2, Mission Processing)
- Consider making adjustments to account for SLOC from Autogenerated code or as part of COTS
- Test for harmful multicollinearity that may be present between age and size variables and MP and size variables
- Look into E2 productivity differences between contractors

# NATIONAL RECONNAISSANCE OFFICE

SUPRA ET ULTRA