# An Application of Change Point Analysis Techniques for the Characterization of the To Complete Indexes Threshold Criterion for Major Acquisition Category I Programs in the Department of Defense

by Gregory C. Stamp

B.S. in Aerospace Engineering, May 2003, University of Maryland, College Park
M.S. in Aerospace Engineering, May 2006, University of Maryland, College Park
M.S. in Aerospace Management, June 2012, Embry-Riddle Aeronautical University
M.S. in Aerospace Operations, June 2012, Embry-Riddle Aeronautical University

A Praxis submitted to

The Faculty of
The School of Engineering and Applied Science
of The George Washington University
in partial fulfillment of the requirements
of the degree of Doctor of Engineering

August 31, 2018

Praxis directed by

Amir H. Etemadi
Assistant Professor of Engineering and Applied Science

Ebrahim Malalla
Visiting Associate Professor of Engineering and Applied Science

The School of Engineering and Applied Science of The George Washington University certifies that Gregory C. Stamp has passed the Final Examination for the degree of Doctor of Engineering as of the date of praxis defense, July 24, 2018. This is the final and approved form of the praxis.

**An Application of Change Point Analysis Techniques for the Characterization of the To Complete Indexes Threshold Criterion for Major Acquisition Category I Programs in the Department of Defense**

Gregory C. Stamp

Praxis Research Committee:

Amir H. Etemadi, Assistant Professor of Engineering and Applied Science, Praxis Director

Ebrahim Malalla, Visiting Associate Professor of Engineering and Applied Science, Committee Member

Kevin Gessner, Program Management Consultant, NSWC IHEODTD Committee Member

© Copyright 2018 by Gregory C. Stamp.
All rights reserved

## Dedication

The author wishes to dedicate this Praxis Research to my parents, Mr. Clive Oliver Stamp and Ms. Pearl Ulysses-Agate as well as my immediate family members.

## Acknowledgments

The author wishes to acknowledge the following individuals. I have reached this point in my academic career with the help and support of family and friends. My lifelong desire to pursue the highest level of academic achievement and love of learning was inspired by father, Clive Oliver Stamp. His passion for learning and dedication academic pursuits were instilled in me from a very early age. My mother, Pearl Annie Agate provided the unyielding belief in my abilities and moral compass in life. Even in the most challenging moments, she refused to accept nothing but victory, ‘steadfastness in the midst of the storm' was expressed to me regularly throughout this journey. My late maternal grandmothers (Mikhalia Stamp and Sister Spencer) were also true spiritual inspirations in my life. I would also like to acknowledge the following immediate family and extended family members. First, my beautiful and loving wife, Fayruz Hassen Mustefa, who provides the rock-solid foundation for our four children, Hannan Degefe, Iyoube Degefe, Jasmine Fatum (Princess) Stamp, and Hassen (‘King') Clive Stamp. Acknowledgments go out my siblings, and special thanks go out to my brother Shomari Levy for many hours of late night conversation about the vicissitudes of life. Thanks go out to all the Stamp and Mustefa family members. What is life without friends that have been there when life seems dismal. Genuinely endearing thanks goes out to Amardip Gosh and Sidhant Krishnaswami for all the wonderful times at University of Maryland, College Park. Lance Campbell, Sharon Campbell, Cassandra Campbell, Nicole Williams, and Ayda Hassen for nonstop support throughout this journey. Also, an immense amount of gratitude is extended to my Navy family with the support from upper management

and supporting staff in Human Resource (HR). I also want to thank Dean Riffat, the faculty, and staff in the School of Engineering and Applied Science.

## Abstract of Praxis

### An Application of Change Point Analysis Techniques for the Characterization of the To Complete Indexes Threshold Criterion for Major Acquisition Category I Programs in the Department of Defense

Project Managers in the Department of Defense and private industry are interested in utilizing effective and accurate methods to monitor and track projects with useful Earned Value Management project indicators. This praxis research addresses the problem of non-effective means of detecting a significant change in the To Complete Indexes trend at the threshold value for project performance indicators in Department of Defense projects. This research was executed by conducting an empirical study of EVM data sets regarding several Major Acquisition Category I programs. In particular, the analysis focused on the examination of project performance indexes, Cost Performance Index and Schedule Performance Index with respect to time along with the To Complete Indexes to determine whether there is a threshold value that illustrates the possibility of an abrupt change in the relationship between the parameters. It is believed that a *To Complete Performance Index* and a *To Complete Schedule Performance Index* value greater than 1.10 indicate that projects have a low probability of obtaining the originally desired project performance objectives, in terms cost and schedule. The quantitative methodology uses the Bayesian Change Point analysis and Wild Binary Segmentation to examine individual projects for the determination of the statistical significance of the pre-threshold and post-threshold change points.

## Table of Contents

Dedication……...…….……...iv

Acknowledgments……...……..……...v

Abstract of Praxis…..…….…….……...vii

List of Figures……..……..…….…x

List of Tables……..…….…….xiii

List of Symbols/Nomenclature…….……..…….…..xiv

Glossary of Terms…….xvi

CHAPTER 1: Introduction…….…….1

1.1 Background……....1

1.2 Problem Statement…….....3

1.3 Thesis Statement……....3

1.4 Research Objective……....4

1.5 Research Questions and Hypotheses……4

1.6 Research Limitations……5

1.7 Organization of Praxis……...6

CHAPTER 2: Literature Review…….....8

CHAPTER 3: Methods……...……...16

3.1 Earned Schedule Model……...16

3.2 Project Performance Indicators Overview……...19

3.2.1 To Complete Performance Index……..19

3.2.2 To Complete Schedule Performance Index……...……22

3.3 Data Collection and Segmentation…...…..……...23

3.4 Change Point Analysis Methodology…...……..…….....24

3.4.1 Bayesian Change Point Analysis…....…….....27

3.4.2 Wild Binary Segmentation Analysis…….……....29

3.4.3 Statistical Hypothesis Testing……...32

CHAPTER 4: Results…….……..…….....36

4.1 Results of Earned Value Management Analysis…..……..….....36

4.2 Results of Bayesian Change Point Analysis and Wild Binary Segmentation…….……..…….....45

4.2.1 Simulated Data Test Results…..…...…….....46

4.2.2 United States Navy Projects…..…...……..…...49

4.2.3 United States Army Projects…..…...……...……..…...61

4.2.4 United States Air Force Projects…...……...…….....64

4.3 Hypothesis Testing Results……...…….....67

CHAPTER 5: Discussion and Conclusions…….....72

5.1 Contributions…,,.…….…….....73

5.2 Future Directions……..…….....74

References……..……...76

Appendices……..……..81

## List of Figures

Figure 1. Example EVM metrics graph which depicts Earned Schedule…….……..17

Figure 2. Illustration of Behavior of TCPI with respect to a constant CPI value……..…..21

Figure 3(a-c). Illustration of both project cost and schedule escalating out of control.....38

Figure 4 (a-c). Illustration of project schedule escalating out of control…….…....40

Figure 5(a-c). Illustration of cost escalating out of control…….....42

Figure 6(a-c). Illustration of non-escalating cost and schedule (in control)……..…….44

Figure 7. Histogram of simulated test data for Bayesian Change Point(BCP) analysis…...46

Figure 8. Bayesian Change Point(BCP) analysis of simulated test data……..……47

Figure 9. Wild Binary Segmentation (WBS) for simulated test data…….……..48

Figure 10. Illustration of the time series trend of TCPI values for Project B5…….…....49

Figure 11. Bayesian Change Point (BCP) Analysis of Project B5 TCPI Data……..….....50

Figure 12. Wild Binary Segmentation (WBS) Analysis for Project B5 TCPI Data…….....50

Figure 13. Illustration of the combine posterior probability and change point locations (Project B5 TCPI Values)……...51

Figure 14. Illustration of the time series trend of TSPI values for Project B5……..…..51

Figure 15. Bayesian Change Point (BCP) Analysis of Project B5 TSPI Data…….…..…..52

Figure 16. Wild Binary Segmentation (WBS) Analysis for Project B5 TSPI Data…….....52

Figure 17. Illustration of the combined posterior probability and change point locations (Project B5 TSPI Values)……..……..53

Figure 18. Illustration of the time series trend of TCPI values for Project H7…..……..54

Figure 19. Bayesian Change Point (BCP) Analysis of Project H7 TCPI Data…….……..54

Figure 20. Wild Binary Segmentation (WBS) Analysis for Project H7 TCPI Data..……..55

Figure 21. Illustration of the combined posterior probability and change point

locations (Project H7 TCPI Values)……..…….....55

Figure 22. Illustration of the time series trend of TSPI values for Project H7……..……..56

Figure 23. Bayesian Change Point (BCP) Analysis of Project H7 TSPI Data..…..…..…..56

Figure 24. Wild Binary Segmentation (WBS) Analysis for Project H7 TSPI Data…….....57

Figure 25. Illustration of the combined posterior probability and change point locations (Project H7 TSPI Values)…….…….....57

Figure 26. Illustration of the time series trend of TCPI values for Project E9…..…….....58

Figure 27. Bayesian Change Point (BCP) Analysis of Project E9 TCPI Data……..…..…58

Figure 28. Wild Binary Segmentation (WBS) Analysis for Project E9 TCPI Data……..…59

Figure 29. Illustration of the combined posterior probability and change point locations (Project E9 TCPI Values)…….……..59

Figure 30. Illustration of the time series trend of TSPI values for Project E9…….…..…..60

Figure 31. Bayesian Change Point (BCP) Analysis of Project E9 TSPI Data……..……..60

Figure 32. Wild Binary Segmentation (WBS) Analysis for Project E9 TSPI Data…..……...61

Figure 33. Illustration of the combined posterior probability and change point locations (Project E9 TSPI Values)….....…….....61

Figure 34. Illustration of the time series trend of TCPI values for Project FF1……..62

Figure 35. Illustration Bayesian Change Point analysis of TCPI values for Project FF1…….……...….62

Figure 36. Wild Binary Segmentation (WBS) Analysis for Project FF1 TCPI Data……..63

Figure 37. Illustration of the combined posterior probability and change point locations (Project FF1 TCPI Values)……..…..……..63

Figure 38. Illustration of the time series trend of TCPI values for Project GGG1…….….64

Figure 39. Illustration Bayesian Change Point analysis of TCPI values for Project GGG1…….……64

Figure 40. Wild Binary Segmentation (WBS) Analysis for Project GGG1 TCPI Data……...…….……...65

Figure 41. Illustration of the combined posterior probability and change point locations (Project GGG1 TCPI Values)……..……....65

Figure 42 (a-b). Illustration of the Q-Q plot for pre-threshold and post-threshold posterior probability data for TCPI…….....…….……...…..67

Figure 43 (a-b). Illustration of the Q-Q plot for pre-threshold and post-threshold posterior probability data for TSPI…….....…….……...…..68

## List of Tables

Table I: Project Outcome Results based on Threshold Criterion for TCPI and TSPI…...45

Table II. Shapiro-Wilk Test for normality (Pre-Threshold TCPI)……..…….....67

Table III. Shapiro-Wilk Test for normality (Post-Threshold TCPI)……...……...67

Table IV. Shapiro-Wilk Test for normality (Pre-Threshold TSPI)……...…….69

Table V. Shapiro-Wilk Test for normality (Post-Threshold TSPI)……..…....69

Table VI. Mann-Whitney Test (TCPI)……...…....70

Table VII. Mann-Whitney Test (TSPI)…….…...…..70

## List of Symbols/Nomenclature

b – Change Point

$B_i$ – Between Block Sum of Squares

p – Transition Probability

P – Conditional Probability

$\gamma$ – Tunning Parameter in BCP

$\lambda$ – Tunning Parameter in BCP

U – Change Point Position

X – data values

$W_i$ – Within Block Sum of Squares

AC – Actual Cost

ACAT – Acquisition Category

AT – Actual Time

BAC – Budget at Completion

BCP – Bayesian Change Point

BS – Binary Segmentation

CPI – Cost Performance Index

CUSUM – Cumulative Sum

CV – Cost Variance

CPA – Change Point Analysis

DoD – Department of Defense

DAMIR – Defense Acquisition Management Informational Retrieval

EAC – Estimate at Completion

EDM – Earned Duration Method

ES – Earned Schedule

ESM – Earned Schedule Method

EU – European Union

EV – Earned Value

EVM – Earned Value Management

IR – Iterative Robust

IEAC – Independent Estimate at Completion

MCMC – Markov Chain Monte Carlo

MDAP – Major Defense Acquisition Program

PD – Planned Duration

PM – Project Manager

PMBOK – Project Management Body of Knowledge

PMI – Project Management Institute

PV – Planned Value

PVM – Planned Value Method

SV – Schedule Variance

SPI – Schedule Performance Index

SIC – Schwartz Information Criterion

TC – Total Cost

TCPI – To Complete Performance Index

TD – Total Duration

TSPI – To Complete Schedule Performance Index

VAC – Variance at Completion

## Glossary of Terms

**Actual Cost (AC)** –: The realized cost incurred for the work performed on an activity during a specific time period.

**Actual Time (AT)** –: The actual time period of the project.

**Budget at Completion (BAC)** –: The sum of all budgets established for the work performed.

**Cost Performance Index (CPI)** –: This a measure of the cost efficiency of budgeted resources expressed as the ratio of Earned Value (EV) and Actual Cost (AC).

**Cost Variance (CV)** –: The amount of budget deficit or surplus at a given point in time. It is expressed as the difference between Earned Value (EV) and Actual Cost (AC).

**Estimate at Completion (EAC)** –: The expected total cost of completing the all work expressed as the sum of the Actual Cost (AC) to date and the estimate to complete.

**Earned Schedule (ES)** –: The time duration associated with the point on the Performance. Measurement Baseline (PMB) where the Earned Value (EV) is equal to Planned Value (PV).

**Earned Value (EV)** –: The measure of work performed expressed in terms of the budget authorized for that work.

**Independent Estimate at Completion (IEAC)** –: This is a forecasting method of the final cost through the use of a derived overall performance efficiency or performance factor.

**Planned Duration (PD)** –: The duration associated with the project plan.

**Planned Value (PV)** –: The authorized budget assigned to schedule work.

**Schedule Variance (SV)** –: This is the conventional method of a measurement of schedule performance expressed as the difference between the Earned Value (EV) and the Planned Value (PV) where the units are in terms of dollars.

**Schedule Variance ($SV_t$)** –: This the modern approach to the measurement of schedule performance expressed as the difference between the Earned Schedule (ES) and Actual Time (AT) where the units are in terms of time periods.

**Schedule Performance Index (SPI)** –: This is the conventional method of a measurement of schedule efficiency expressed as the ratio of the Earned Value (EV) and the Planned Value (PV).

**Schedule Performance Index ($SPI_t$)** –: This the modern approach to the measurement of schedule efficiency expressed as the ratio of the Earned Schedule (ES) and Actual Time (AT).

**Total Cost (TC)** –: This is the desired total cost of the project which can be either Budget at Completion (BAC) or the Estimate at Completion (EAC).

**To Complete Performance Index (TCPI)** –: This is a measure of the cost performance that is required to be achieved with the remaining resources in order to meet a specified management goal. It is expressed as the ratio of the cost to finish the remaining work to the remaining budget.

**Total Duration (TD)** –: This is the desired total time period which can be either the Planned Duration (PD), the Negotiated Duration (ND), or the Estimated Duration (ED).

**To Complete Schedule Performance Index (TSPI)** –: This is a measure of the schedule performance that is required to be achieved with the remaining resources in order to meet a specified management goal. It is expressed as the ratio of the time required to finish the remaining work to the remaining amount of time available or allocated for the project.

**Variance at Completion (VAC)** –: This term is a projection of the amount of budget deficit or surplus, expressed as the difference between the Budget at Completion (BAC) and the Estimate at Completion (EAC).

# Chapter 1 – Introduction

## 1.1 Background

High failure rates of large-scale and high-profile projects are significant concerns for industry and government stakeholders. The failure rate of these projects can be directly attributed to substantial cost and schedule overruns. According to D'Amico et al. (2018), several research studies and reports on numerous acquisition programs in the Department of Defense (DoD) have demonstrated how risky the various acquisition programs were and the susceptibility to cost and schedule overruns. High failure rates translate into losses of millions of dollars and the termination of projects before performance objectives are met. It is critical that individuals involve in the planning and budgeting need to be fully aware of the tendencies for the programs to have higher cost and longer time frames to complete than initially anticipated. One wonders why this is the case and are there any tools can better anticipate these potential costs and schedule growth (D'Amico, White, and Ritschel, 2018). Providing Project Managers (PMs) with better tools to predict project performance objectives will minimize cost and schedule overruns and reduce the high failure rates of projects.

Previous research efforts within the private sector using predictive tools in Earned Value Management (EVM) have focused on the *To Complete Performance Index (TCPI) and the To Complete Schedule Performance Index (TSPI)* as a means of enabling project managers to track and monitor the progress of a project effectively. Humphreys (2014) stated that one of the main uses of TCPI is to compare it with CPI because "this comparison provides an immediate indication" of a realistic estimated cost or an

achievable budget. Similarly, the TSPI is a representation of the projected performance of the project's schedule.

According to McCarthy (2016), forecasting techniques in Earned Value (EV) theory use existing performance to predict the completion of the project. This type of EV forecasting method concerning cost is TCPI (McCarthy, 2016). One may view the TCPI as a representation of "projected future performance" of the project's cost, which can be based on the project's budget (Humphreys, 2014). Humphreys (2014) stated that one might look at TCPI as a projection of the comparison between the allocated funds for remaining work with the estimate for the remaining amount of work.

Some researchers have argued that when a project's TCPI value exceeds a threshold value of 1.10, a project is predicted to be out of control and beyond the point of recoverability to achieve the initial desired project cost (Lipke, 2009a). At this point, Project Managers will more than likely need to re-baseline the project by either requesting more money or time, with a new budget requiring upper management or sponsor authorization (Lipke, 2009a). This potentially useful tipping point is still observed from a theoretical perspective Lipke (2016a). However, a strong validation has not taken place because of the limited application of this claim.

Within the TCPI and TSPI characteristic trends, there are possible multiple change points. What is the statistical significance of the change point for TCPI and TSPI values greater than or equal to 1.10. Does this indicate a significant threshold value that must be addressed? This praxis research addresses the problem of non-effective means of detecting statistically significant change in the To Complete Indexes trend at the threshold value for project performance indicators in a broad spectrum of Department of

Defense (DoD) projects (e.g. land, sea, and air mission environments), specifically, Major Defense Acquisition Programs (MDAPs), also known as Acquisition Category (ACAT) I programs. The quantitative methodology uses the Bayesian Change Point (BCP) analysis was implemented to examine individual projects to determine the statistical significance of the pre-threshold and post-threshold change points. The praxis research will show that Change Point Analysis (CPA) Techniques (e.g., Bayesian CPA and Wild Binary Segmentation) can be used to characterize the threshold criterion for the To Complete Indexes (TCPI and TSPI) for the determination of the point at which the project performance is considered to be out of control.

## 1.2 Problem Statement

This praxis research addresses the following problem:

*There is no clear or robust method for estimating the threshold criterion for the To Complete Indexes (TCPI and TSPI) to determine the point at which the project performance is considered to be out of control for Major ACAT I Programs which will more than likely lead to project schedule and cost overruns with little probability of recovery.*

## 1.3 Thesis Statement

*Change Point Analysis Techniques (e.g., Bayesian CPA and Wild Binary Segmentation) can be used to characterize the threshold criterion for the 'To Complete Indexes' (TCPI and TSPI) for the determination of the point at which the project performance is considered to be out of control*.

## 1.4 Research Objective

The objective of this praxis research is to determine if the threshold criterion for the 'To Complete Indexes' (TCPI and TSPI) be characterized for the determination of the point at which the project performance is considered to be 'out of control', or the probability of recovering the desired Total Cost (TC) and Total Duration (TD) is very low (Lipke, 2016a). The practical application of this research is the potentially significant impact on how the Project Manager (PM) perceives the immediate need for project intervention and management action.

## 1.5 Research Questions and Hypotheses

The following are the research questions:

- Can change point analysis methods presented in this praxis be used to characterize the threshold criterion for the To Complete Performance Index (TCPI)?
- Can change point analysis methods presented in this praxis be used to characterize the threshold criterion for the To Complete Schedule Performance Index (TSPI)?

From the problem statement presented above, the following hypotheses will either be accepted or rejected based on statistical analysis of the comparison of the probability distribution of pre-threshold and post-threshold change points:

- Alternate Hypothesis ($H1_a$): There is a statistically significant difference between the post-threshold and pre-threshold change points for the To

Complete Performance Index (TCPI) regardless of military branch of service (Navy, Army, Air Force).

- Alternate Hypothesis ($H2_a$): There is a statistically significant difference between the post-threshold and pre-threshold change points for the To Complete Schedule Performance Index (TSPI) regardless of military branch of service (Navy, Army, Air Force).

## 1.6 Research Limitations

One of the limitations of this praxis research is having a larger pool of project data to analyze. The author was limited to data that was available via the Federal Government database known as the Defense Acquisition Management Information Retrieval (DAMIR) which is a repository of programmatic data for numerous DoD programs. Most of the projects that were used were active projects, so access to previously completed projects and EVM data associated with those projects were not available at the time of data acquisition and analysis conducted for this research. Access to this information would have been useful in providing a basis to develop a more robust model and validation process in characterizing the threshold criterion.

Another limitation of this research is that a detailed analysis was not conducted to determine the probability of recovery for the individual projects after it was determined that either TCPI or TSPI had surpassed the threshold criterion value of 1.10. Also, once the likelihood of the project recovery was established, a follow-on analysis of the actual project status was not conducted to validate the recoverability of the project based on whether the project To Complete Indexes exceeded the threshold criterion.

## 1.7 Organization of Praxis

The remaining sections of the praxis are organized in the following manner. In Chapter 2, a literature review is presented that highlights some of the prior research that has been done regarding Earned Value Management (EVM) and other aspects of Project Management. This chapter highlights distinct elements of government projects versus private sector projects and how the different culture could have an impact on the project outcomes. The literature review goes in-depth into the foundation of Earned Schedule (ES) which is fundamental to the performance index, TSPI. This chapter also provides some fundamental prior research work behind the concept of Change Point Analysis (CPA), the methods used in this research effort and its application to time series data.

In Chapter 3, the quantitative methods that were used in this research are presented. This chapter provides a detail look at Earned Value Management (EVM) methods with specific formulations of EVM metrics that were used in this research, such as the equations for Earned Schedule (ES), TCPI, and TSPI. Furthermore, the Change Point Analysis (CPA) methods used in this research are explained in detail as well as fundamental aspects of statistical testing methods that were used in this research. The results of the analysis are presented in Chapter 4 starting with several illustrations of EVM parameters, specifically the trends for TCPI and TSPI for the various projects considered in this research. The results were organized to illustrate different cases where the behavior of trends showcases some interesting aspects of the To Complete Indexes (TCPI and TSPI) with respect to project completion. The results also demonstrate the application of the CPA methods as well as the tabulation results of the hypothesis testing. Finally, Chapter 5 provides a discussion of the results and summarizes the findings and

offers perspective on the author's contributions with this research in the conclusion section.

# Chapter 2 – Literature Review

The Project Management Institute (PMI) defines Earned Value Management as “a management methodology for integrating scope, schedule, and resources, and for objectively measuring project performance (PMI, 2008)”. According to Kim et al. (2003), it is a means of allowing managers to accurately track the status of their projects by looking through the lenses of both the cost and schedule of the project. The Department of Defense (DoD) was at the forefront of the development of this unique management tool (Kim et al., 2003). Kim et al. (2003) stated that at its inception the private sector was slower to adopt EVM as compare to DoD, and in a broader sense, numerous Federal Government Agencies. According to Kim et al. (2003), the private sector is paying closer attention to this invaluable management technique to control their projects better and compete on the global scale. Lipke et al. (2009b) and Fleming and Koppelman (2000) are useful sources for readers to become familiar with basic terminologies that are currently being used in the lexicon of EVM.

According to Larson and Gray (2014), practitioners in the field of Project Management, more specifically EVM, like to use project performance indexes to monitor the progress of the project. The cost and schedule indexes are typically more preferred over the absolute values obtained from Schedule Variance (SV) and Cost Variance (CV) by practitioners in the field of Project Management (Larson and Gray, 2014). One of the critical reasons for this is that the indexes are efficiency ratios (Larson and Gray, 2014).

Larson and Gray (2014) stated that when the indexes are graphed over the project life cycle, this information can be very "illuminating and useful." Indexes are typically used

at the cost account level and above. However, in practice, the database is used to develop indexes that allow the project manager and customer to view the project's progress from various perspectives (Larson and Gray, 2014). Therefore, one can see the practical value of using indexes as an objective tool to effectively communicate with the customers as well as upper-level management regarding the decision-making process. PMs can use EVM to investigate longitudinal pattern over time which in turn can lead to a discovery of a range of factors that are linked to the cost growth over various time frames (D'Amico et al., 2018). One could argue that even though D'Amico et al. (2018) focus was cost growth with EVM metrics, an analogy could be made for growth factors regarding schedule.

Lipke (2009a) was one of first to introduce the concept of the Earned Schedule (ES) factor that was applied to EVM data analysis. ES is critical because it is based on actual time increments versus conventional analysis of dollar values. ES was created to study the challenges encountered in project management. ES is also used to generate the TSPI, which is similar to the TCPI (Lipke, 2009a). The main difference is that the TCPI is a performance indicator for cost while the TSPI is a performance indicator for schedule. Both indicators are used to inform Project Managers (PMs) about the necessary cost and schedule performance efficiency that will be needed throughout the remainder of the project to achieve desired final project performance objectives (Lipke, 2016a). The threshold for both indexes can be a driving force behind a PM's decision to intervene and execute necessary actions to achieve project performance objectives (Lipke, 2016a). Once a project exceeds this threshold value, the project is considered to be out of control and the probability of the project's recovering is very low (Lipke, 2009a). Hence, the

ability to return to the originally desired project's cost and schedule is unlikely, and the Project Manager will need to re-baseline the project (Lipke, 2009a).

Lipke (2016b) theorized that the value of 1.10 could be considered as a credible threshold value for performance indicators, TCPI and TSPI. If this baseline is empirically sound, it could become a tool for anticipating a change at the threshold value for TCPI and TSPI in EVM to prevent a higher failure rate of ACAT I projects. It was observed from a thorough literature review that the performance indicator threshold value for both TCPI and TSPI has never been applied to a broad spectrum of ACAT I projects within the DoD (e.g., Army, Air Force, Navy, and Marine Corps).

However, Lipke's (2016a) research examined the threshold value for both TCPI and TSPI with the use of data from 25 private sector projects. These findings were promising but were preliminary, and a test for ACAT I projects is a worthy application. Research has shown that DoD project management is culturally and abruptly distinct from that of the private sector (Sommer, 2016). The DoD business model is safety-driven rather than profit-driven, which is the primary business model in the private sector. According to Goure (2015), the primary objective of these companies is to make money, which is a driving factor in how they conduct their business. Goure (2015) argues that money is very relevant to any private company because it allows these companies to provide a paycheck to the employees, fabricate and sell their product line, execute R&D investments, paying out dividends to shareholders, etc. Some of these aspects of a private company are not relevant for DoD, so the motivation and mission are different.

The PMBOK Guide (2006), asserts that the unique characteristics of the government illuminate some of the critical differences between the public sector and the

private sector. The PMBOK Guide (2006) further elaborated on the need for project teams to recognize the vital and unique characteristics of government projects that will allow them to efficiently and effectively manage their projects. According to PMBOK Guide (2006), a few unique features associated with government projects that project teams and Project Managers should be aware of are the legal constraints on government projects, public accountability, and practical use of public resources. In general, government projects are subjected to specific legal constraints that private sector projects may not be obligated to meet. This could have a significant impact on how the PM execute the project to achieve project objectives (PMBOK Guide, 2006). Accountability is an extremely crucial factor in the public sector because it extends beyond the agency's immediate stakeholder, some external stakeholders are involved.

In contrast, PMs for the private sector projects are mainly accountable to their immediate client and very few stakeholders (PMBOK Guide, 2006). One could argue that the decision-making process is different from the PMs in the private sector and the public sector. How quickly a project is executed can be affected by the accountability process for the PM. The acquisition of public sector project budget is funded through taxes, bonds, and other fees which must be authorized for release by the appropriate budget authority (PMBOK Guide, 2006). The delay of the release of the proper funds to execute a project could affect the effective, adequate execution of the project. According to PMBOK Guide (2006), the project manager in the public sector must use public funds to satisfy the objectives set out by the government agency.

Arena et al. (2005) look at this comparison from another perspective through real cases associated with shipbuilders. Arena et al. (2005) interviewed major shipbuilders in

the United States, United Kingdom, and the European Union (EU) and observed that private sector shipbuilding has different schedule performance objectives from those of the DoD. One of the significant differences in schedule performance could be attributed to the unique needs of the private sector and the DoD (Arena et al., 2005). Other differences cited were the approach to the management of change orders by each entity in addition to the contractor's incentives for on-time completion of projects. These differences spawn different organizational cultures in how business is done. Their data showed that change orders and late production definition overwhelmingly caused the significant schedule slippage (as reported by shipbuilders). Arena et al. (2005) said that it takes up to four times longer for the DoD projects to resolve change orders, which extends schedules and consequently cost overruns. One can see the motivating factor to understand EVM metrics better to provide project managers with better tools to improve their completion of projects and achieve significant cost savings.

Therefore, the question arises, could there be a difference in the threshold values that impact the behavior of the TCPI and TSPI as well as the rates of change of both the TCPI and the TSPI? This is one of the motivating factors behind this praxis research effort. This praxis research examines Lipke's (2016) theory about the claim of the significant impact of the threshold value for both the TCPI and the TSPI. Specifically, the analysis is focused on the examination of project performance objectives (the CPI and the SPI(t)) and the To Complete Indexes (TCPI and TSPI) to determine if a threshold value illustrates the possibility of an abrupt change in the trends of the performance indicators. The objective is to determine if the threshold value of 1.10 for both TCPI and TSPI is a realistic value that can be applied to DoD Projects and if it is not, what threshold value

for TCPI and TSPI would be more applicable to DoD Projects. Crumrine's (2013) research effort focused on the application of ES and Schedule Performance Index (SPI(t)) concerning time. It didn't establish a definitive threshold value but did raise questions about what that threshold value should be for various Air Force platforms.

With regards to analytical methods used in this research, time series analysis approach has been implemented. The application of time series analysis has been playing a critical role in diverse fields of study including medicine, aerospace, finance, business, and meteorology (Aminikhanghahi and Cook, 2017). According to Aminikhanghahi and Cook (2017), one can see time series data as repeated measurements over some time frame that describes the behavior of the system. These behaviors can change over time due to various external events and internal systematic changes in dynamics or distribution (Aminikhanghahi and Cook, 2017). Expounding on the application of times series analysis, Zhang et al. (2017) stated that in academia and industry this type of analysis has become more frequent.

One of those approaches to time series analysis is the method of Change Point Analysis, which can be considered as the examination of changes in the behavior of the system over time. According to Sharma et al. (2016), the detection of these change points can be considered as the identification of abrupt variation in the process behavior due to distributional or structural changes in the trend. Brodsky and Darkhovsky (1993) stated that one of the central problems of statistical inference is that of the change point problem. It can be viewed as linking together various statistical methods such as the theory of estimation, hypothesis testing, statistical control theory, frequency analysis and Bayesian approaches, as well as segmented analysis (Brodsky and Darkhovsky, 1993).

The incorporation of Change Point Analysis (CPA) in Earned Value Management (EVM), according to Hossein-Mortaji et al. (2015), facilitates the means of examining dynamics of project performance over some time frame. Hossein-Mortaji et al. (2015) article is one of the first, to the best of my knowledge, to use Change Point Analysis techniques to examine EVM metrics, in particular, project performance such final cost and duration. The authors focus on the use of the cumulative sum (CUSUM) estimator method to analyze how a project performance changes over time. My research uses the same concept for examining how a project changes over time and where the significant changes are occurring with the goal of characterizing the threshold criterion. This research examines where the significant change points are occurring by using more robust change point methods such as Bayesian Change Point and Wild Binary Segmentation.

In this research, the Binary Segmentation (BS) technique and more specifically, the Wild Binary Segmentation (WBS) that was developed by Fryzlewicz (2014) was implemented in the time series analysis. Sharma et al. (2016) have examined the WBS method along with three other change point methods such as the Bayesian Change Point (BCP) Analysis, E-Agglomerative(E-Agglo.) algorithm, and Iterative Robust Detection (IR) method. Sharma et al. (2016) primary objective was to conduct a comparison of those four methods with a goal of determining which one is the most capable of finding the exact number and locations of change points. According to Sharma et al. (2016), the developers of the WBS change point method proclaims that with no prior assumption on the distribution of the data this method can detect the number of critical change points as well as their locations in the trend.

In addition, Sharma et al. (2016) stated that one of the primary reason for the effectiveness of the WBS method, according to developers of this change point method, is that it solves the problems associated with the Binary Segmentation (BS) method as well as the improvement in computation. The WBS method uses the concept of computing the Cumulative Sum (CUSUM) of intervals that are drawn randomly, and the first change point candidate is considered as the largest CUSUMs that will be tested against some stopping criteria (Sharma et al., 2016). Sharma et al. (2016) stated that the Bayesian Change Point (BCP) Analysis algorithm is implemented based on a product partition model which computes the change point based on data and current partition. For more accuracy in detecting the change points tuning parameters values are implemented (Sharma et al., 2016). For the comparison of the BCP method with the other change point methods, Sharma et al. (2016) implemented the 'bcp' R package that was originally developed by Erdman and Emerson (2007).

# Chapter 3 – Methods

## 3.1 Earned Schedule Model

By definition, Schedule Variance (SV) is the difference between Earned Value (EV) and Planned Value (PV) as shown in equation (1), while SPI is the ratio between EV and PV as shown in equation (2). Both EV and PV are EVM metrics, which are provided in project status reports in dollar amounts. One can calculate the SV and SPI if EV and PV are known (Lipke, 2009a).

$$SV = (EV - PV) \quad (1)$$

$$SPI = \left(\frac{EV}{PV}\right) \quad (2)$$

The Schedule Variance (SV) in a traditional sense tells Project Managers how much a project is ahead or behind in dollar amounts but doesn't provide a clear understanding of how much time a project is behind schedule. It becomes a challenge to accurately assess the actions needed to make adjustments when a project moves out of control. Another common tool, the SPI, is an index value in dollar amounts for EV and PV (Lipke, 2009a). The SPI is an efficiency indicator that tells Project managers how well a project is performing in the "planned schedule" (Lipke, 2009a). Lipke (2009a) created the concept of Earned Schedule to remedy the issues associated with the conventional analysis of SV and SPI.

According to Batselier and Vanhoucke (2015), three well-known methods for project duration were explored over the years in the field of EVM. The three methods for

forecasting actual time periods are the Planned Value Method (PVM) which was created by Anbari (2003), the Earned Schedule Method (ESM), which is Lipke's concept, (2003), and the third was created by Jacob and Kane, known as the Earned Duration Method (EDM). Also, according to (Vandevoorde and Vanhoucke, 2006), these three forecasting methods were, to the best of their knowledge, "the only three project duration forecasting methods that were presented in literature" at that time. Through empirical research, it was shown that of the three schedule forecasting methods, Earned Schedule Method (ESM) "produced more accurate results" (Batselier and Vanhoucke, 2015). This praxis research effort uses ESM explained by Lipke (2009a) to execute this empirical analysis. Figure 1 illustrates a model for calculating Earned Schedule (Lipke, 2009a). Which was reproduced here with permission by Lipke.

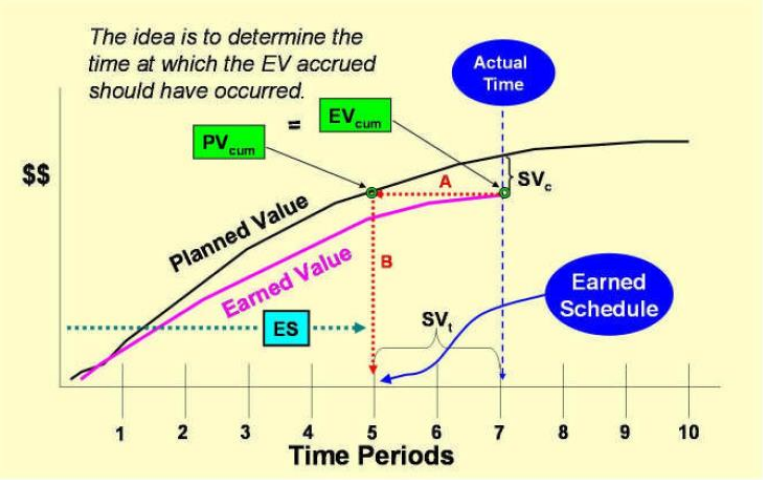



**Figure 1. Example EVM metrics graph which depicts Earned Schedule (reproduced with permission granted from Lipke (2009))**

The definition of the equation for TSPI is dependent upon Lipke's concept of Earned Schedule (ES) (Lipke, 2009a). Lipke (2009a) formulated ES to provide Project Managers with practical information about their project's schedule performance. This terminology extends the subject area of Earned Value Management (EVM) in a more reliable way that contributes to the Engineering Management Body of Knowledge. Research in this specific area of EVM is expanding globally with research efforts from Henderson (2004). According to Lipke (2009a), one of the main reasons ES provides a more reliable analysis of schedule performance is because it is time-based as opposed to traditional schedule performance parameters such as, Schedule Variance (SV) and Schedule Performance Index (SPI), which are based on project cost. Schedule Variance (SV) was used for many years but produce inaccurate results Lipke (2009a).

According to Lipke (2009a), ES is the determination of the amount of accrued Earned Value (EV) that should have occurred. The formulation defined by Lipke (2009a) is shown in equation (3):

$$ES = C + \left(\frac{EV - PV_C}{PV_{C+1} - PV_C}\right) \quad (3)$$

Where the C term is "the number of whole time increments of the Performance Measurement Baseline (PMB) such that EV ≥ PV" (Lipke, 2009a). Another way of describing the C term is in- regards to the identification of the "planned value period" where the cumulative Earned Value (EV) is greater than or equal to the aggregate Planned Value (PV) (Lipke, 2009a). This equation for Earned Schedule (ES) provides a definitive time factor that can be assessed by Project Managers.

## 3.2 Project Performance Indicators Overview

With a known ES, one can calculate the projected performance based on actual time performance by examining the To Complete Performance Index and To Complete Schedule Performance Index. It is known that TCPI is an efficiency measurement of the work remaining as compared to the amount of funds remaining. Fleming and Koppelman (2000) acknowledged the potential usefulness of this predictive indicator for PMs and its ability to provide a means of

ascertaining performance factors needed to accomplish the remaining effort to achieve a specific management objective. Traditionally, this index is calculated to evaluate the Estimate at Completion (EAC). However, it was not used in the past to assess the potential impact once TCPI value crosses a specific threshold criterion.

### 3.2.1 To Complete Performance Index

Research has shown that the EVM community has accepted that there are primarily two ways to calculate TCPI based on Earned Value metrics (Lipke, 2009). One of the accepted formulations is in terms of a planned cost or the Budget at Completion (BAC), which is the project manager's baseline budget without any Management Reserves (MR) involved, Earned Value (EV), and Actual Cost (AC), shown in equation (4) (Lipke, 2009). Another way to calculate TCPI is with respect to EAC, which is shown in equation (5) (Lipke, 2009). Formulations of TCPI are as follows:

$$TCPI_P = \left(\frac{BAC-EV}{BAC-AC}\right) \quad (4)$$

$$TCPI_E = \left(\frac{BAC-EV}{EAC-AC}\right) \quad (5)$$

In this praxis research, the planned TCPI, equation (4) formulation was used to calculate the TCPI values. The TCPI analysis based on EAC, equation (5) was not used in this analysis because the goal of the investigation was to determine if the project would achieve its proposed budget. According to Lipke (2009a), equation (5) is used to determine if the Project Manager's specifically desired estimate cost at the completion of the project is achievable. The estimated TCPI show in equation (5) can be used in "what if" scenarios for desired cost. This value at times can be dramatically different from the original project planned budget. One of the unique characteristics of this project performance indicator is that it provides Project Managers with an effective way to monitor and track projects (Lipke, 2009a). In his examination of the TCPI indicator, Lipke discovered some unique aspects of the indicator when it was plotted with respect to the completion time or fraction of completion, see Figure 2. The primary purpose of Figure 2 is to illustrate the behavior of TCPI trend as the project progresses towards completion (reproduced here with Lipke's permission). This image was used to demonstrate what happens to the TCPI trend with a constant CPI value. Mainly, it shows that the TCPI trend will surpass the threshold criterion of 1.10 at approximately 35 percent (%) completion with a constant CPI value of 0.85. According to Lipke (2009a), the "planned final cost" or BAC will not be achieved if this situation remains the same. It should be noted that the BAC is assumed to be the Total Cost (TC) in this scenario.

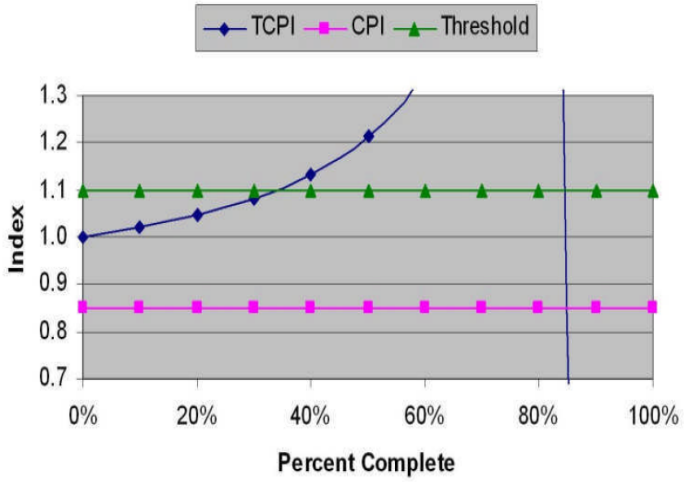



**Figure 2. Illustration of Behavior of TCPI with respect to a constant CPI value (reproduced with permission granted from Lipke (2009a))**

According to Lipke (2016a), the trend for TCPI crosses a unique criterion, known as the threshold value. With further examination from a theoretical aspect of TCPI, he explored what happens when the TCPI trend crosses this threshold value and then escalates "out of control" and the impact on the project performance objectives (Lipke, 2009a). Lipke (2009a and 2016b) also determined that the TCPI could be used to determine the probability of recovery once the index value crosses the threshold criteria of 1.10. This factor was not explored in this research. He further examines this behavior by studying the rate of increase of the TCPI with respect to Earned Value percent (%) or fraction of completion (Lipke, 2009a). The concept of the rate of increase for both indexes was not explored in this paper. Fleming and Koppelman (2000) also alluded to predictive nature of the TCPI by stating that this indicator not only places emphasis on the remaining project tasks but also showcases possibility for a project to adequately recover from a cost performance that is negative.

### 3.2.2 To Complete Schedule Performance Index (TSPI)

Before the creation of the Earned Schedule (ES) concept, TSPI was unknown to the EVM community (Lipke, 2014). Lipke (2014) mentioned that the analysis and formulation that was demonstrated for TCPI could also be executed for TSPI. Lipke (2009a) defined TSPI in two ways similar to the TCPI. The first definition shown in equation (6) is with respect to Planned Duration (PD), Total Duration, which is the Planned Duration (PD) and any schedule contingencies Lipke (2009a). The second method is regarding Estimate at Completion (EAC) given Actual Time (AT) of the project, shown in equation (7):

$$TSPI_P = \left(\frac{PD-ES}{TD-AT}\right) \quad (6)$$

$$TSPI_E = \left(\frac{PD-ES}{EAC(t)-AT}\right) \quad (7)$$

Like the planned TCPI formulation, in this research, the planned TSPI formulation, equation (6) was used to calculate the TSPI values. It should be noted that the Total Duration can be equal to the Planned Duration (PD) if there is no known additional reserve time or "schedule risk reserve" (Lipke, 2016b). An understanding of TCPI and TSPI with the corresponding rate of increase can be invaluable to the PM because it can be used to effectively assess the cost and schedule objectives (Lipke, 2014). This allows PMs the ability to determine the point at which a project is considered unrecoverable and necessary available time for management intervention (Lipke, 2014).

## 3.3 Data Collection and Segmentation

Data was collected to ensure that the sample used in this analysis encompasses a broad spectrum of DoD platforms that represented land, sea, and air mission environments. This approach was to address Crumrine's (2013) recommendations, recognizing a need to validate Lipke's Earned Schedule (ES) methodology across Major ACAT I DoD Programs, rather than the limited program reach of his 2013 study. The concern was that there could be potentially significant differences between the platforms that could produce inconsistent results. The primary source of the data is from a Federal Government database known as the Defense Acquisition Management Information Retrieval (DAMIR). The Office of Performance Assessment manages this database and Root Cause Analyses (PARCA) Directorate, which is one of the Directorates in the Office of Assistant Secretary of Defense for Acquisition. PARCA was established to execute performance assessments and root cause analyses of Major Defense Acquisition Programs (MDAP) within DoD (PARCA website, 2017). The EVM data is periodically checked for accuracy as a part of the agency's roles and responsibilities. In general, this entity is held accountable for the EVM performance, oversight, and governance across the entire Department of Defense (DoD) (PARCA website, 2017).

The DAMIR database provides a streamlined process for acquisition management and oversight. DAMIR is an authoritative source for data acquisition, data collection, as well as a repository for a variety of acquisition program information (e.g., EVM data source). Various data sources used by the DoD Acquisition community for the effective management of Major Defense Acquisition Programs (MDAP), also known as Acquisition Category (ACAT) programs as well as Major Automated Information

Systems (MAIS) programs are showcased in the DAMIR database. The data in this praxis research also differentiates itself from the private sector data Lipke (2016a) evaluated in his work. Lipke (2016a) examined project data from three broad areas in the private sector. The data sources identified in Lipke (2016a) stems from four information technology projects, 12 high technology product development projects), and nine construction type projects.

Since these primary sources of data in Lipke's (2016) are from primarily private industry, it was inferred that a similar analysis could be done extensively on military platforms to see if this threshold criterion could be applied to Major ACAT I Programs in DoD. This study is the first to examine both TCPI and TSPI with use of Bayesian Change Point (BCP) analysis method focusing on a considerable number projects representing a broad spectrum of military platforms. The projects investigated in this research had a project size, concerning dollars, ranges from approximately several millions of dollars ($M) to several billions of dollars ($B). Also, these projects were mostly from the production and development phases of the project cycle.

## 3.4 Change Point Analysis Methodology

The methodology used in this praxis research is fundamentally a quantitative analysis with the use of a statistical algorithm for conducting Bayesian Change Point Bayesian Change Point (BCP) analysis and Wild Binary Segmentation (WBS) analyses. The investigation was performed on individual projects to determine the statistical significance of the pre-threshold and post-threshold change points. The data from each project was selected based on two main criteria. First, the projects must have a fraction of

completion of at least 20 percent (%) and no major re-baselining where significant funds or time is infused in the project. Combining the data investigated if a general conclusion can be made about the behavior of the relationship between the parameters. Therefore, an analysis was conducted to detect if a consistent and significant abrupt change was occurring in the data sets.

From a historical perspective, the concept of change point detection has been observed to have a connection with the industrial quality control and continuous inspection with the publication from E.S. Page (1954) as well as Girshick and Rubin (1952) (Brodsky and Darkhovsky, 1993). A.N. Kolmogorov developed the theoretical formulation of the change point problem in the late 1950's. (Brodsky and Darkhovsky, 1993). According to Brodsky and Darkhovsky (1993), "the problem of optimal sequential change-point detection" was solved by A.N. Shiryaev (1963). The application of Bayesian and the min-max solutions for both discrete and continuous time (Brodsky and Darkhovsky, 1993). Brodsky and Darkhovsky (1993) stated that current researchers in this field of statistical research are exploring a different algorithm that tries to address this idea of the existence of homogeneity in the sample of observations.

Four fundamental classifications of change point problems are identified as the off-line (*a posterior* method) or on-line (sequential method) observations of the data, *a priori* statistical information (parametric, semi-parametric, and non-parametric), data characteristics (random process versus random field), and disorder type (abrupt versus gradual changes) (Brodsky and Darkhovsky, 1993). Most of the literature concerning the change point problem is focused on "the abrupt changes in probabilistic characteristics of observations" while minimal research effort has been put forth for the examination of

“gradual changes” in probabilistic characteristics. Other aspect of the change point problem examines the number of change points that exist in the observations, either single or multiple change points (Brodsky and Darkhovsky, 1993).

More current applications of change point analysis are showcased by Chen and Gupta (2012), in which, several detail analyses of practical problems are were addressed. The authors, Chen and Gupta (2012), have explored the daily fluctuations of stock prices from multiple companies and possible abnormal shifts in the trends. In addition, Chen and Gupta (2012) examined research data from various samples acquired from different geological sites. From the biomedical discipline, gene expression measurements for a "cell line" may contain significant changes that may elude some critical dynamics of "gene regulation." According to Chen and Gupta (2012), this information may be of significant interest to biomedical researchers. In cancer research, the variations in the "DNA copy number" are being examined with the use of the change point model (Chen and Gupta, 2012).

In the book edited by Barber et al. (2011), the authors stated many time series are characterized by structural changes in the trend, which can be manifested as sudden jumps in level or volatility. Typically, in practice, the number of change points will be unknown (Barber et al., 2011). According to Barber et al. (2011), from a methodological and practical standpoint, the ability to effectively detect change points carry a prominent level of significance. The authors in (Barber et al., 2011), declared that for multiple change point applications one of the most reliable CPA methods is the Bayesian Change Point method. Barber et al. (2011) emphasized that a variety of disciplines have been increasingly showing a keen interest in the ability to quickly and accurately detect change

points. In agreement with Chen and Gupta (2012), the authors indicate the rapid increase in the application of change point methods in areas such as bioinformatics, network traffic analysis, finance, and climatology (Barber et al., 2011). This praxis implements two change point analysis methods, Bayesian Change Point (BCP) and the Wild Binary Segmentation (WBS). The next section describes the overall theory of each process.

### 3.4.1 Bayesian Change Point Analysis

Barry and Hartigan extensively explored the theoretical Bayesian analysis for change point problems in 1993 (Erdman and Emerson, 2007). Erdman and Matteson (2007) created an R package called *"bcp"* that was based on the theoretical approach that was proposed by Barry and Hartigan (1993). According to Erdman and Emerson (2007), the Bayesian method distinguishes itself from the frequentist procedures by generating a probability distribution that illustrates the location of the change points in the trend and their associated probability values. Hence, the practical aspect of this method is that it provides the analyst a "more informative summary that reflects the degree of uncertainty in the change points" (Erdman and Emerson, 2007). This is one of the critical differences in comparison to the frequentist methods, which primarily focus on the estimation of the number and specific locations of the change points (Erdman and Emerson, 2007).

The authors in Barber et al. (2011) also stated that one of the critical advantages of the BCP method is that, in comparison to alternative change point methods, this method can more easily quantify the uncertainty in the change point positions. The ability to adequately summarize the posterior distribution is a disadvantage as compared to other change point methods that indicate the specific statistically significant changepoints

(Barber et al., 2011). In this research, an attempt is made to overcome this disadvantage of the BCP method. Also, this research uses the BCP analysis method by applying the *"bcp"* R package to the TCPI and TSPI time series data to determine the probability distribution of statistically significant change points. This method can be used to determine if the threshold value greater than or equal to 1.10 is a statistically significant change point for each of the Major ACAT I programs explored in this research.

According to Erdman and Emerson (2007), the algorithm uses "a partition $\rho = (U_1, U_2, \ldots U_n)$, where the $U_1 = 1$ indicates a change point at position $i + 1$. Barry and Hartigan (1993) formulated the "transition probability, *p,* for the conditional probability of a change point at the position $i + 1$, can be determined the following formulation, where **X** is the data:

$$\frac{p_i}{(1-p_i)} = \frac{P(U_i=1|\mathbf{X},U_j,j\neq i)}{P(U_i=0|\mathbf{X},U_j,j\neq i)} \qquad (14)$$

According to Erdman and Emerson (2007), equation 14 represented a simplified expression of the ratio between the transitional probability and the conditional probability. This was further expanded by Barry and Hartigan (1993) to include the tuning parameters, γ, and λ, as well as "within and between block sum of squares" terms, $W_0, B_0, W_1, B_0$ Erdman and Emerson (2007). Equation 15, which was formulated by Erdman and Emerson (2007), is like Barry and Hartigan (1993) theoretical formulation with the primary difference being stability in the terms when implemented with Markov Chain Monte Carlo (MCMC) simulation (Erdman and Emerson, 2007). In Equation 15,

the variable ‘b’ is the number of segments obtained while ‘n’ is number of observations (Erdman and Emerson, 2007).

$$\frac{p_i}{(1-p_i)} = \left(\frac{W_0}{W_1}\right)^{\frac{n-b-2}{2}} * \left(\frac{B_0}{B_1}\right)^{\frac{b+1}{2}} * \left(\sqrt{\frac{W_1}{B_1}}\right) * \left( \frac{\int_0^{1+\frac{\frac{B_1\lambda}{W_1}}{\frac{B_1\lambda}{W_1}}} p^{\frac{(b+2)}{2}}(1-p)^{\frac{(n-b-3)}{2}}\,dp}{\int_0^{1+\frac{\frac{B_0\lambda}{W_0}}{\frac{B_0\lambda}{W_0}}} p^{\frac{(b+1)}{2}}(1-p)^{\frac{(n-b-2)}{2}}\,dp} \right) * \left( \frac{\int_0^{\gamma} p^{b}(1-p)^{n-b-1}\,dp}{\int_0^{\gamma} p^{b-1}(1-p)^{n-b}\,dp} \right) \tag{15}$$

This allows the algorithm to more tractable and produces more stable solutions for long sequences and therefore making it more robust for various time series observation (Erdman and Emerson, 2007). The Markov Chain Monte Carlo method was used to implement the estimates for the posterior distributions of the change points as well as the means (Erdman and Emerson, 2007). One is referred to Erdman and Emerson (2007) as well as Barry and Hartigan (1993) for more detail formulation of the algorithm used in the ‘bcp’ analysis R package.

### 3.4.2 Wild Binary Segmentation Analysis

According to Fryzlewicz (2014), the interesting problems associated with change point detection have been of keen interest for statisticians for many years. All the various methods primary objectives to detect the number of change points and the location of changes in the evolving time series data (Fryzlewicz, 2014). A couple of noteworthy methods are the "piecewise-stationary modeling" finds appeal in the idea of estimating the number of change points and their location which can be shown to link to the behavior of the trend. For the construction of more complex models, Fryzlewicz (2014) stated that statisticians had used a posterior segmentation model, which considers the

piecewise constants as being random and evolving variables with the use of Markovian type mechanism. Fryzlewicz (2014) elaborated on the fact that the computational complexity has been a major drawback of change-point estimators that have been formulated as multivariate optimization problems. However, over the years, efforts have been made to reduce the computational complexity with algorithms such as “pruned dynamic programming” from Rigaill (2010) and minimization of description length with the use of a genetic algorithm proposed by Davis, Lee, and Rodriguez-Yam (2006).

According to Fryzlewicz (2014), one of the multiple change point detection methods that have been producing promising results is the generic technique known as Binary Segmentation (BS). This method first analyzed the entire dataset for a single change point using the CUSUM technique. Next, if a change point is detected then the data is split into two sub-segments by the algorithm, hence the name binary segmentation (Fryzlewicz, 2014). A similar search routine is implemented on either subsegment, possibly resulting in the further splitting of the subsegments. This “recursive action on a given segment continues until a certain criterion is satisfied" (Fryzlewicz, 2014). In terms of the general Binary Segmentation (BS) method, according to Fryzlewicz (2014), it is believed that L. Vostrokva made the first analytical formulation that proposed this method in a stochastic process environment in 1981.

Fryzlewicz (2014) have observed some of the advantages of the BS method such as simple concept, minimal computational complexity, and very manageable code generation algorithm. In addition, other researchers in the change point analysis discipline have noted that the BS method is "arguably the most widely used change point search method" (Fryzlewicz, 2014). Despite its popularity, the BS method has observed

some of its flaws that limit its accuracy in detecting multiple change points. Most real case random process will more than likely to have multiple change points instead of a single change point in a process.

Therefore, Fryzlewicz (2014) proposes an enhancement of the BS method by capitalizing on its strengths and expanding its limitation by implementing a multiple change point detection algorithm. This newly proposed method is called the Wild Binary Segmentation method. According to Fryzlewicz (2014), conceptually, the idea is simple. Fryzlewicz (2014) describes the process as such, instead of using a global CUSUM statistic on the entire data sample ($X_1$, $X_2$, …. $X_T$), the first stage involves numerous random samples that drawn to create a vector subsample. Then the subsamples from the CUSUM statistics are maximized, where the largest being chosen from the entire collection of CUSUMs. These values are then tested against a certain threshold. A statistical significance test is executed, and if it has been determined that they are in fact statistically significant, then the same procedure is repeated on each side of the data (Fryzlewicz, 2014).

The hope, according to Fryzlewicz (2014), is that at least one change point and its location can be detected even for a relatively small number of random draws. As well as establishing a lower bound for the number of draws that provides favorable draw with high probability (Fryzlewicz, 2014). A stopping criterion based on what is known as the strengthened Schwarz Information Criterion (SIC) is also implemented in the WBS change point algorithm (Fryzlewicz, 2014).

In terms of formulation, According to Fryzlewicz (2014), both the BS and WBS change point algorithm is essentially based on the CUSUM statistic that is defined by the

"inner product between the vectors and specific vector" of contrast weights given in the following equation:

$$\tilde{X}_{s,e}^{b} = \left(\sqrt{\frac{e-b}{n(b-s+1)}}\ \sum_{t=s}^{b} X_t\ \right) - \left(\sqrt{\frac{b-s+1}{n(e-b)}}\ \sum_{t=b+1}^{e} X_t\ \right) \qquad (16)$$

where $s \leq b < e$ and $n = e\ -\ s\ +\ 1$

According to Fryzlewicz (2014), the general BS algorithm computes $\tilde{X}_{1,T}^{b}$ and the first prospective change point is obtained by using equation 17 as follows:

$$b_{1,1} = arg\ max_{b:1\leq b\leq T}\left|\tilde{X}_{1,T}^{b}\right| \qquad (17)$$

The significance of the change is compared against a criterion. The general BS algorithm is extended to examine data for multiple change point using the Wild Binary Segmentation.

### 3.4.3 Statistical Hypothesis Testing

From a theoretical standpoint, the concept of hypothesis testing of which tests are derived as solutions of optimum problems was developed by Jerzy Neyman and Egon Sharpe Pearson in the 1930s (Lehman, 1959). Neyman and Pearson were credited as being the first to emphasize that the hypothesis testing must consider both the hypothesis and its associated alternative (Lehman, 1959). Their mathematical proof of hypothesis testing is not provided in this research; however, the reader is recommended to examine theoretical formulation in the book written by Lehmann (1959). From a scientific

perspective, one could argue that at the cornerstone of scientific inquiry and analysis is hypothesis testing (Dytham, 2003). Conceptually, the idea is that if statistical testing is conducted one is examining the probability that a hypothesis is correct (Dytham, 2003). If the probability is small, then the hypothesis is not valid and therefore is rejected for the alternative (Dytham, 2003). At the heart of hypothesis testing is the test for or against the null hypothesis, which is often labeled ($H_0$). Dytham (2003) describes a statistical test as an examination of the likelihood of the null hypothesis being true. Therefore, if the null hypothesis is rejected on must accept the alternative hypothesis ($H_1$).

According to Dytham (2003), in statistical testing, this is frequently done with the use of the p-value, which is considered as the "bottom-line of statistical tests." The critical value of accepting or rejecting the null hypothesis is the value of 0.05, which has been designated as alpha or the level of statistical significance (Dytham, 2003). For example, a p-value of 0.07 means that the hypothesis has a 7% chance of being true (Dytham, 2003). Because the fact that the null hypothesis is frequently the primary hypothesis being tested, it is customary to look for a low p-value to decide whether to reject the null hypothesis and accept the alternative (Dytham, 2003). A p-value of 0.05 is considered the threshold point where the critical decision is made for or against the null hypothesis.

Hypothesis testing can be broken down into two main categories, parametric and non-parametric. Most statistical testing involves the assumption that the data samples have a normal distribution, or some other known distribution (Walker and Lev, 1953). If the data does conform to some known distribution, then the parametric tests are more relevant hypothesis testing methods to use. A few well know parametric hypothesis testing are t-test, F-test, and Analysis of Variance (ANOVA). If, however, the distribution of the data

is non-normal, or unknown, it is more appropriate to use the non-parametric testing methods, also known as "distribution-free" (Walker and Lev, 1953).

Before implementing any of the hypothesis testing methods, it is recommended that the data must be analyzed for normal distribution. The underlying distribution in numerous inferential procedures is based on the assumption that they are Gaussian (Hanagal, 2009). Some of the most well-known methods are the Shapiro-Wilk's test, Kolmogorov-Smirnov, and Bartlett test (Huck, 2012). Whenever these tests are used, the null hypothesis is tested to determine acceptance or rejection. If the p-value is less than the level of significance, 'known as alpha', then one must reject the null hypothesis. Hence the data has not met the normality requirements (Huck, 2012). According to Hanagal (2009), the use of a quantile plot is also an effective means of checking the normality of the data. The analysis of the quantile plots is centered around how the data is distributed around a linear pattern (Hanagal, 2009). Therefore, if the data in question is of normal distribution, then the data points will fall close to a straight line. Otherwise, if the real distribution is not normal or Gaussian, then one should see the data points deviate away from the line.

Some well-known non-parametric testing methods are Wilcoxon signed rank tests, Mann-Whitney U test, and the Kruskal-Wallis test. The Wilcoxon signed rank test is the non-parametric equivalence to the paired t-test. Unlike the Wilcoxon signed rank test, the Mann-Whitney U test does not require the samples to be paired. This test examines samples from two independent groups (Dytham, 2003). According to Dytham (2003), the Mann-Whitney U test is considered to be the non-parametric equivalence to the independent samples t-test and one-way ANOVA, however, the critical difference is that

this test doesn't make any prior assumption about "homogeneity of the variances" and distribution of the data (Dytham, 2003).

In terms of the Kruskal-Wallis test, this non-parametric statistic test compares three or more independent groups. One could consider this test an extension of the Mann-Whitney U test. Therefore, if an analysis is needed to compare three or more groups with either non-normal or unknown distribution data then it is recommended to use the Kruskal-Wallis test. In this research, the Mann-Whitney U test was chosen because this analysis focuses on the comparison of the pre-threshold and post-threshold data with different variances and the unknown distribution of the two independent samples. It is the most appropriate of the three non-parametric statistical testing methods. This method will be used to determine if there is a statistical difference between the pre-threshold and post-threshold data.

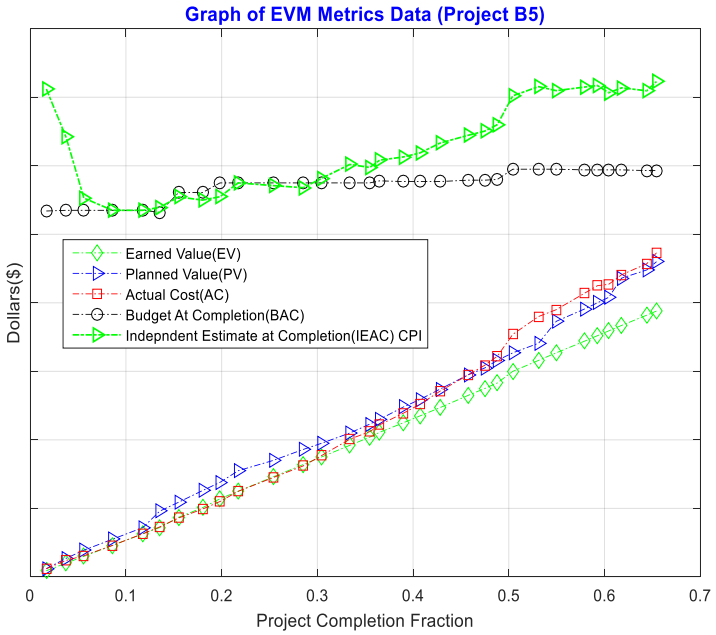
# Chapter 4 – Results

## 4.1 Results of Earned Value Management Analysis

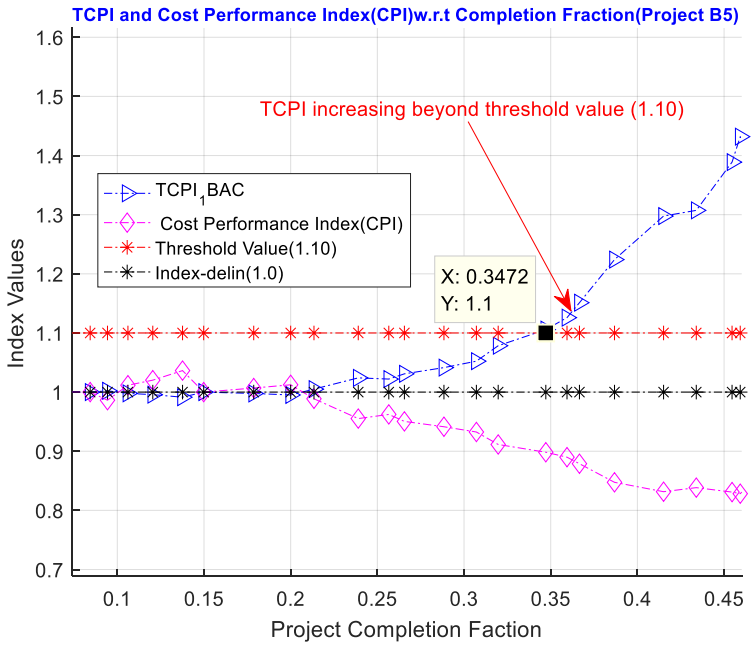
Several examples of the trends for individual projects are shown below in Figures 3f through 6. Values on the y-axis were removed from Figures 3a, 4a, and 5a due to the sensitivity of the data. These figures were selected to illustrate the different cases for EVM metrics data, the TCPI and TSPI trends with respect to project completion. Recall, that the TCPI and TSPI are future performance indicators while CPI and SPI (t) are current performance indicators. TCPI and TSPI address the question about a project's status given the amount of remaining funds and time available. The first case demonstrates how both cost and schedule are out of control. The second scenario depicts what happens when a project is primarily showing a schedule that is escalating out of control. The third case is when the cost increases significantly as compared to schedule.

The final scenario looks at when both TCPI and TSPI are below the threshold value of 1.10. It should be noted that most of the projects used in this research are active projects; therefore, the data will indicate projects that are not at 100% completion, or project completion fraction less than one. Since TCPI and TSPI are future performance indicators, the project doesn't have to be completed. Based on the project's allocated budget and schedule, the Project Manager is interested in determining if the project will meet its objectives given the remaining amount of Funds and Schedule. In a sense, it is a predictor tool that Project Managers (PMs) can use to monitor the progress of the project and try to prevent crossing the critical threshold. This provides the PM the ability to allocate the necessary resources that can potentially affect the future estimates.

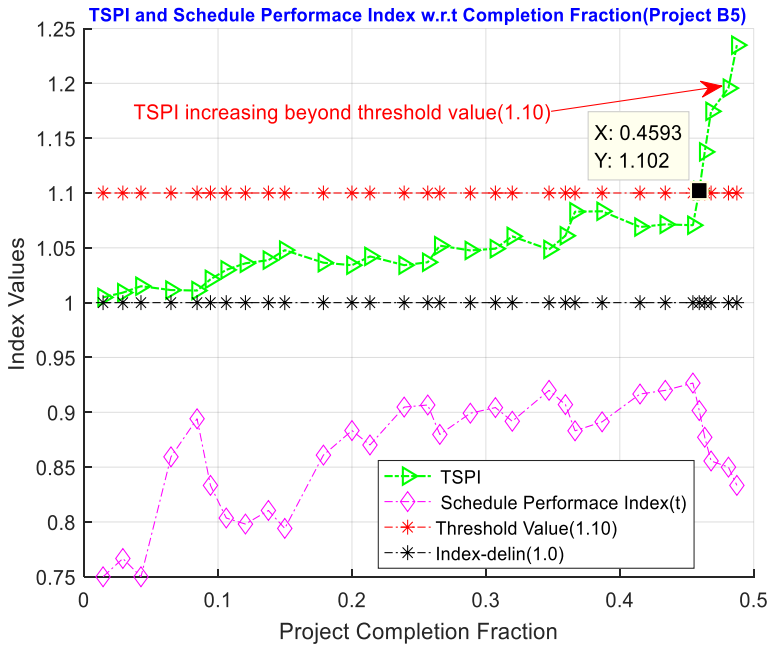
Figures 3(a-c) illustrates the standard EVM metric values, a comparison between the TCPI and CPI as well as the TSPI and SPI(t) for Project B5, which exhibits the behavior of a project in a severe predicament regarding cost and schedule overruns. It was observed that both the TCPI and TSPI graphs follow a similar response to the graphs produced by Lipke (2009) in Figure 2. It becomes quite clear that the values for both TCPI and TSPI are escalating exponentially while the CPI and SPI values are below the benchmark value for project success set equal to 1.0 (Fleming and Koppelman, 2000). Both the threshold value of 1.10 and the benchmark value for project success (1.0) are also superimposed on the graph for relative comparison.




3 (a)



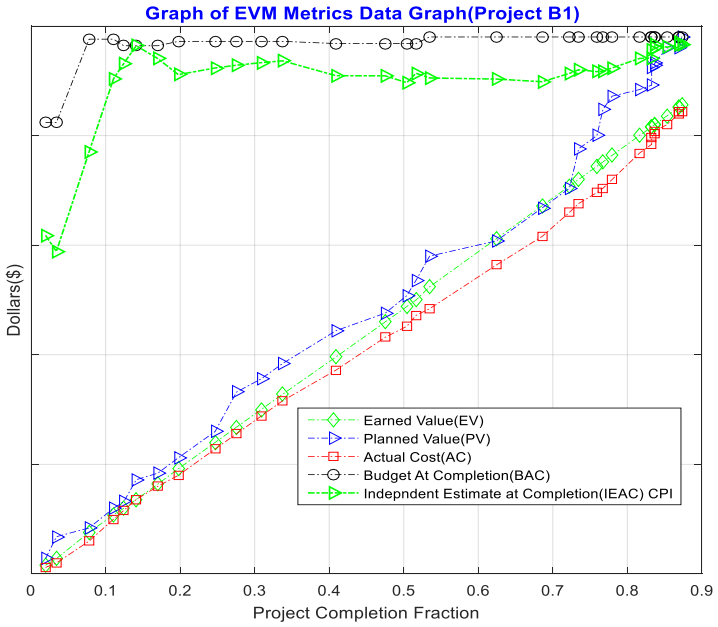

3 (b)



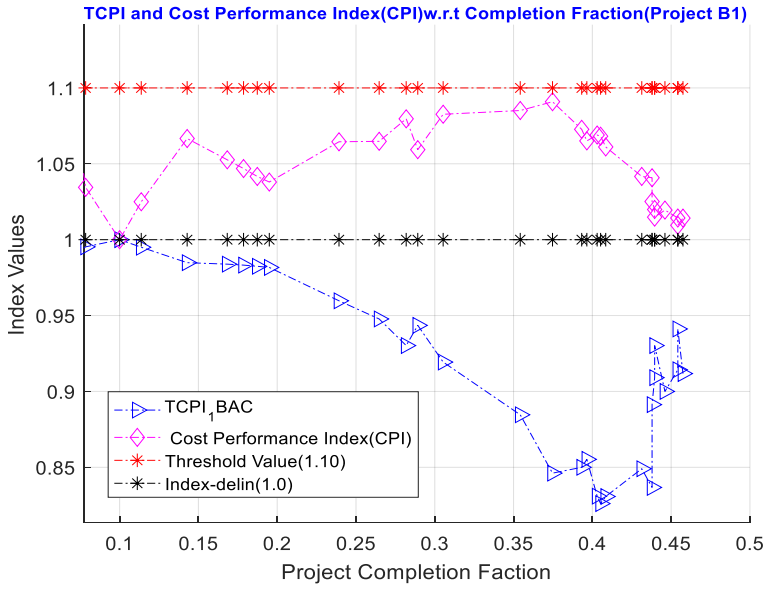

3 (c)

**Figure 3(a-c). Illustration of both project cost and schedule escalating out of control**

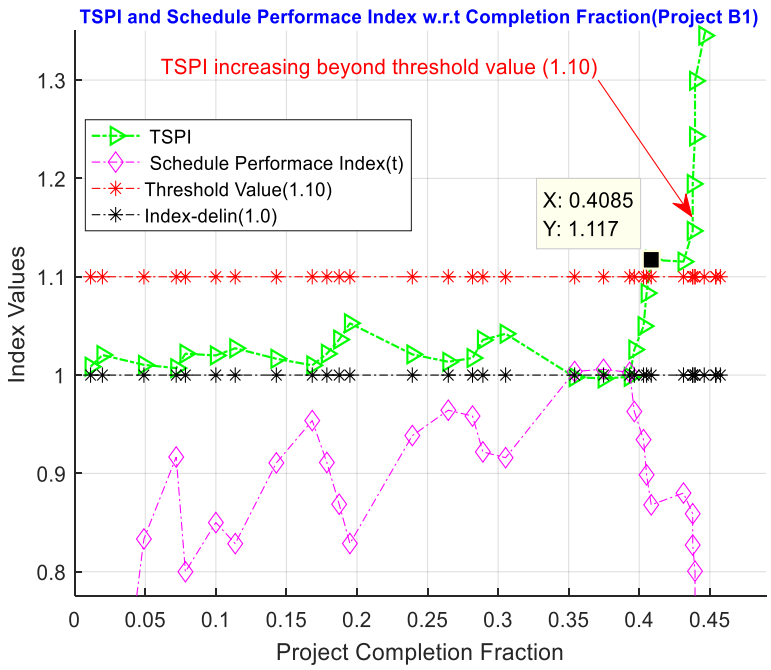
Comparably, Figures 4(a-c) illustrate how the schedule is escalating out of control but, the cost is still within budget as demonstrated by the TCPI graph with respect to CPI (Figure 4b) and TSPI graph with respect to SPI(t) (Figure 4c) for Project B1.




4 (a)



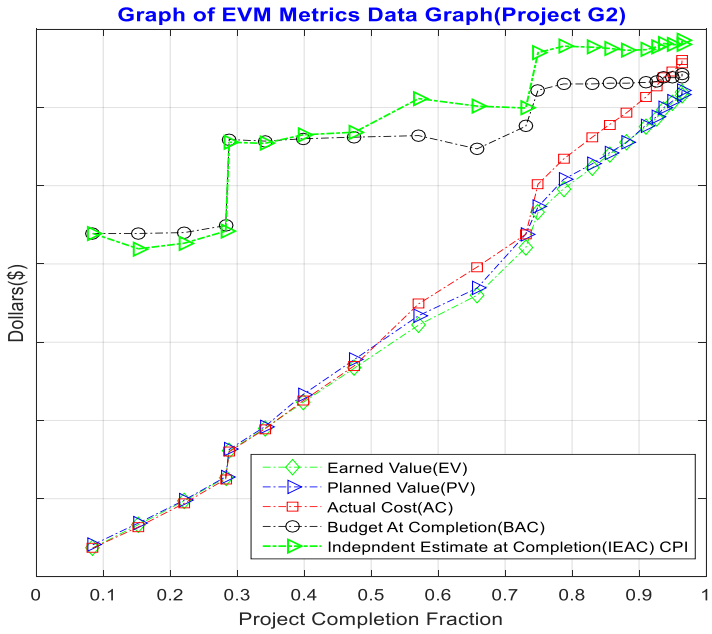

4 (b)



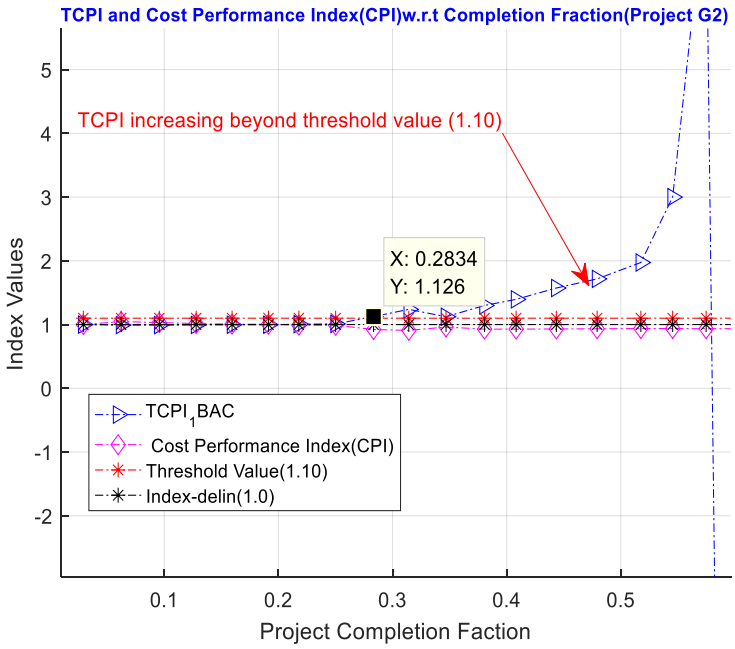

4 (c)

**Figure 4 (a-c). Illustration of project schedule escalating out of control**

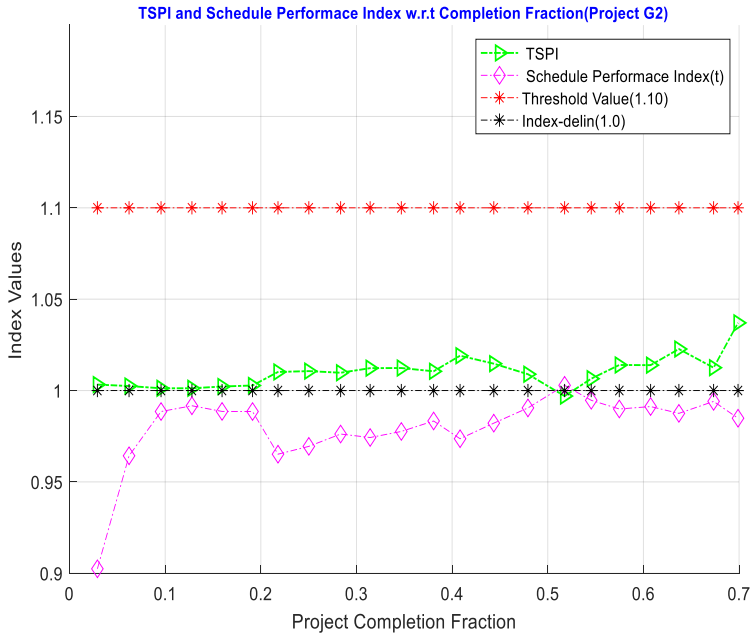
Figures 5(a-c) illustrate how the cost is escalating out of control, but the schedule is still within time duration as demonstrated by the TCPI graph with respect to CPI (Figure 5b) and the TSPI graph with respect to SPI(t) (Figure 5c) for Project G2. In Figure 5b, the decreasing trend of the CPI values is not apparent in this graph because the values for the increasing trends of TCPI overwhelm the values CPI when observed as a whole.


Graph of EVM Metrics Data Graph(Project G2)
Dollars($)
Project Completion Fraction
Earned Value(EV)
Planned Value(PV)
Actual Cost(AC)
Budget At Completion(BAC)
Indepndent Estimate at Completion(IEAC) CPI
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

5 (a)


TCPI and Cost Performance Index(CPI)w.r.t Completion Fraction(Project G2)
TCPI increasing beyond threshold value (1.10)
X: 0.2834
Y: 1.126
Index Values
Project Completion Faction
TCPI1BAC
Cost Performance Index(CPI)
Threshold Value(1.10)
Index-delin(1.0)
5
4
3
2
1
0
-1
-2
0.1
0.2
0.3
0.4
0.5

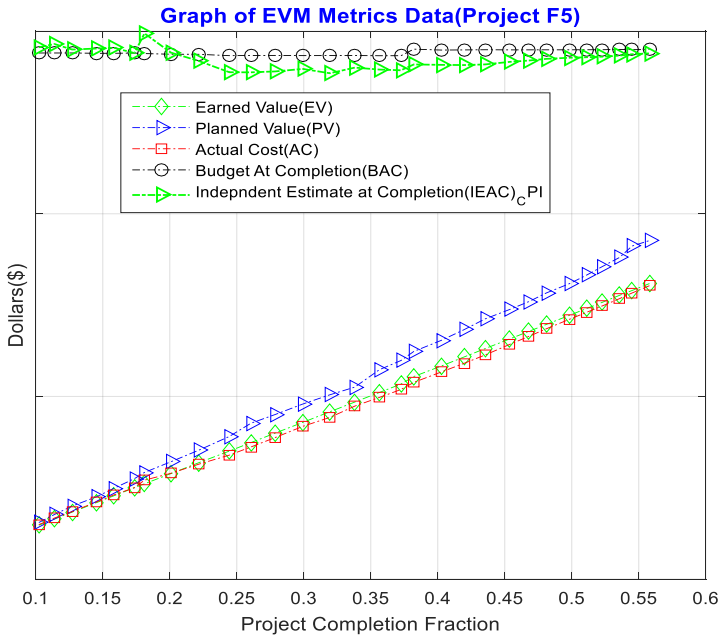
5 (b)



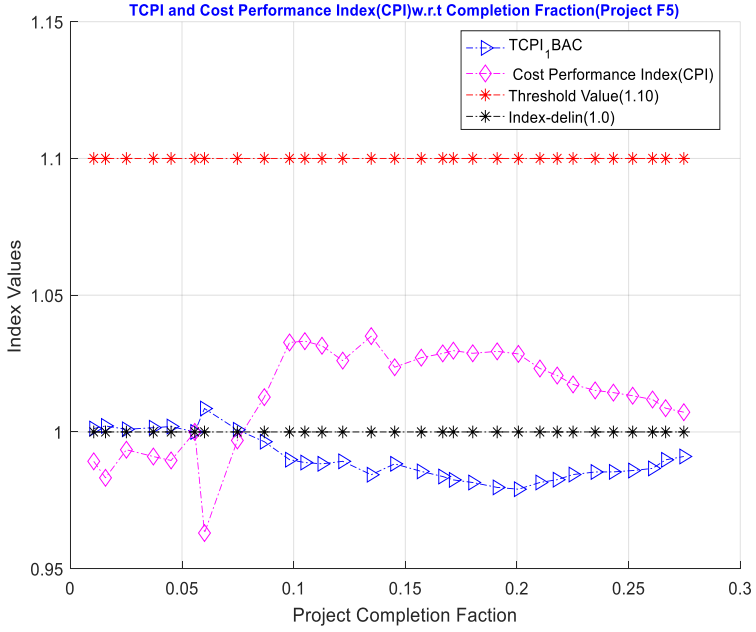

5 (c)

**Figure 5(a-c). Illustration of cost escalating out of control**

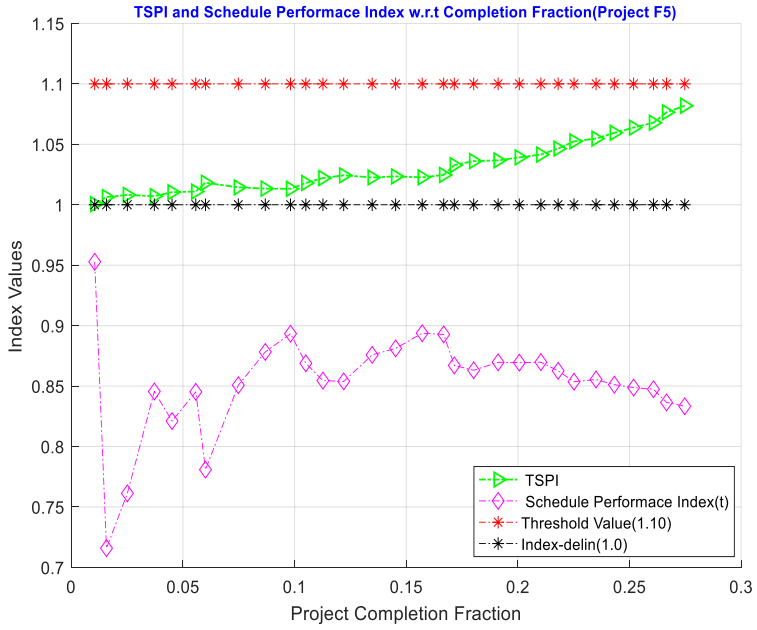
Figure 6(a-c) illustrates the scenario when neither the TCPI nor the TSPI trend is above the threshold value of 1.10. The figures illustrate where they are at or below 1.00. Whenever this occurs, the project is in good standings regarding cost and schedule, as shown in 6b and 6c, respectively.


Graph of EVM Metrics Data(Project F5)
Earned Value(EV)
Planned Value(PV)
Actual Cost(AC)
Budget At Completion(BAC)
Indepndent Estimate at Completion(IEAC)cPI
Dollars($)
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45
0.5
0.55
0.6
Project Completion Fraction

6 (a)


TCPI and Cost Performance Index(CPI)w.r.t Completion Fraction(Project F5)
TCPI1BAC
Cost Performance Index(CPI)
Threshold Value(1.10)
Index-delin(1.0)
Index Values
1.15
1.1
1.05
1
0.95
0
0.05
0.1
0.15
0.2
0.25
0.3
Project Completion Faction

6 (b)




6 (c)

**Figure 6(a-c). Illustration of non-escalating cost and schedule ("in control")**

The graphs produced above represent a small sample of numerous similar illustrations that were generated for each of the projects listed in Table 1. For the sake of brevity, details of TCPI and TSPI trends for the individual projects are summarized in Table 1. It illustrates the list of all the Major ACAT I DoD Projects that were chosen for this research effort. Because of the sensitive information, contractor's names and associated programs were not listed. The data shows that when the TCPI and TSPI value surpasses the threshold value 1.10 as time progresses and nothing is done to remedy the situation (e.g., additional resources such as money or increased time) the project begins to escalate "out of control" as shown in figures above. This indicates that the threshold value of 1.10 could potentially be applied to any Major ACAT I Project. Also, the results of the Variance at Completion (VAC) for cost and schedule are also shown.

**Table I. Project Outcome Results based on Threshold Criterion for TCPI and TSPI**

Table 1

Results of Project Outcomes based on Threshold Value Criterion

| Item# | Project # | TCPI >1.10 (yes/no) | IEAC_cost > BAC (yes/no) | Variance at Completion(VAC) cost (positive/negative) | Project Cost Out of Control (TCPI Values) (yes/no) | TSPI >1.10 (yes/no) | IEAC_schedule > PD (yes/no) | Variance at Completion(VAC) schedule (positive/negative) | Project Schedule Out of Control (TSPI Values) (yes/no) | Fraction of Completion | Branch of Service |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | A1 | No | No | positive | No | Yes | Yes | Negative | Yes | 0.411 | Navy |
| 2 | A4 | Yes | Yes | Negative | Yes | Yes | Yes | Negative | Yes | 0.414 | Joint |
| 3 | B1 | No | No | positive | No | Yes | Yes | Negative | Yes | 0.873 | Navy |
| 4 | B5 | Yes | Yes | Negative | Yes | Yes | Yes | Negative | Yes | 0.654 | Navy |
| 5 | B6 | Yes | Yes | Negative | Yes | Yes | Yes | Negative | Yes | 0.981 | Navy |
| 6 | C6 | Yes | Yes | Negative | Yes | Yes | Yes | Negative | Yes | 0.792 | Navy |
| 7 | E8 | Yes | Yes | Negative | Yes | Yes | Yes | Negative | No | 0.969 | Navy |
| 8 | E9 | Yes | Yes | Negative | Yes | Yes | Yes | Negative | Yes | 0.558 | Navy |
| 9 | E10 | No | No | positive | No | Yes | Yes | Negative | Yes | 0.585 | Navy |
| 10 | F5 | No | No | positive | No | No | Yes | Negative | No | 0.558 | Navy |
| 11 | G1 | Yes | Yes | Negative | Yes | No | Yes | Negative | No | 0.947 | Joint |
| 12 | G2 | Yes | Yes | Negative | Yes | No | Yes | Negative | No | 0.926 | Joint |
| 13 | G3 | Yes | Yes | Negative | Yes | No | Yes | Negative | No | 0.754 | Joint |
| 14 | H6 | Yes | Yes | Negative | Yes | Yes | Yes | Negative | Yes | 0.936 | Navy |
| 15 | H7 | Yes | Yes | Negative | Yes | Yes | Yes | Negative | Yes | 0.920 | Navy |
| 16 | H8 | No | No | positive | No | Yes | Yes | Negative | Yes | 0.877 | Navy |
| 17 | H9 | No | No | positive | No | Yes | Yes | Negative | Yes | 0.532 | Navy |
| 18 | H10 | Yes | Yes | Negative | Yes | Yes | Yes | Negative | Yes | 0.664 | Navy |
| 19 | M1 | Yes | Yes | Negative | Yes | Yes | Yes | Negative | Yes | 0.390 | Navy |
| 20 | BB4 | Yes | Yes | Negative | Yes | Yes | Yes | Negative | Yes | 0.925 | Army |
| 21 | BB5 | No | No | positive | No | Yes | Yes | Negative | Yes | 0.975 | Army |
| 22 | BB6 | No | No | positive | No | Yes | Yes | Negative | Yes | 0.898 | Army |
| 23 | DD1 | Yes | Yes | Negative | Yes | Yes | Yes | Negative | Yes | 0.941 | Army |
| 24 | EE1 | Yes | Yes | Negative | Yes | No | Yes | Negative | No | 0.516 | Army |
| 25 | FF1 | Yes | Yes | Negative | Yes | Yes | Yes | Negative | Yes | 0.809 | Army |
| 26 | BBB2 | Yes | Yes | Negative | Yes | Yes | Yes | Negative | Yes | 0.629 | Air Force |
| 27 | EEE2 | Yes | Yes | Negative | Yes | No | Yes | Negative | No | 0.901 | Air Force |
| 28 | GGG1 | Yes | Yes | Negative | Yes | No | Yes | Negative | No | 0.816 | Air Force |
| 29 | GGG3 | Yes | Yes | Negative | Yes | No | Yes | Negative | No | 0.961 | Air Force |
| 30 | GGG4 | Yes | Yes | Negative | Yes | Yes | Yes | Negative | Yes | 0.952 | Air Force |
| 31 | GGG5 | Yes | Yes | Negative | Yes | Yes | Yes | Negative | Yes | 0.602 | Air Force |

## 4.2 Results of the Bayesian Change Point Analysis and Wild Binary Segmentation

Change Point analysis technique and the statistical significance of those three change points. This BCP model was used on the real datasets for the projects examined in this research.

### 4.2.1 Simulated Data Test Results

Testing the ability to detect change points in a sequential dataset accurately, a simple case was created by generating four successive random normally distributed data set. A total of 900 observations were formed with the first data set having 300 observations with the mean and standard deviation of zero and one, respectively. Similarly, the second data set has 200 observations with a mean and standard deviation of five and one, respectively. Also, the third data set has 200 observations with the mean and standard deviation of 10 and one, respectively. The final data set had an additional 200 observations with a mean and standard deviation of 15 and one, respectively. The histogram of the entire simulated dataset is shown in Figure 7.

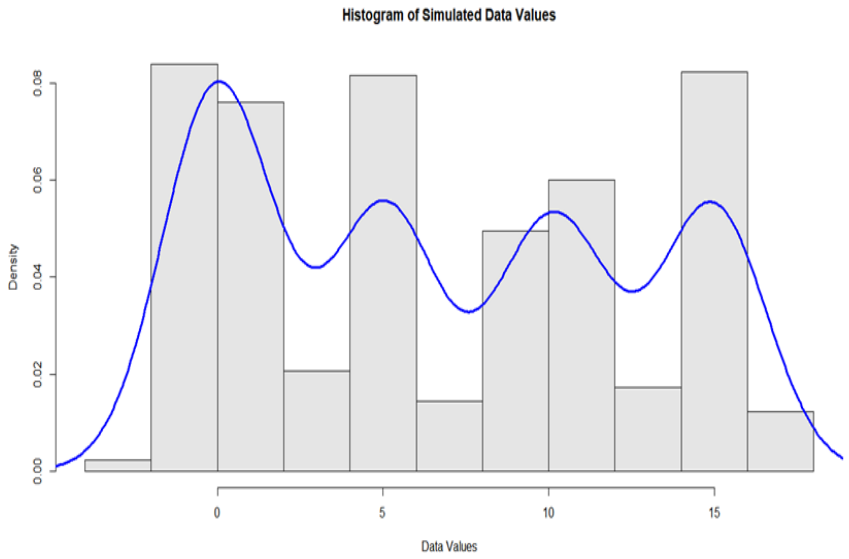



**Figure 7. Histogram of simulated test data for Bayesian Change Point (BCP) analysis**

As predicted, the illustration of the histogram of simulated test data shows the four distinct distributions. The BCP algorithm in the 'bcp' R package that was created by

Erdman and Emerson (2007) was used to detect these change points that were intentionally designed in the simulated dataset.

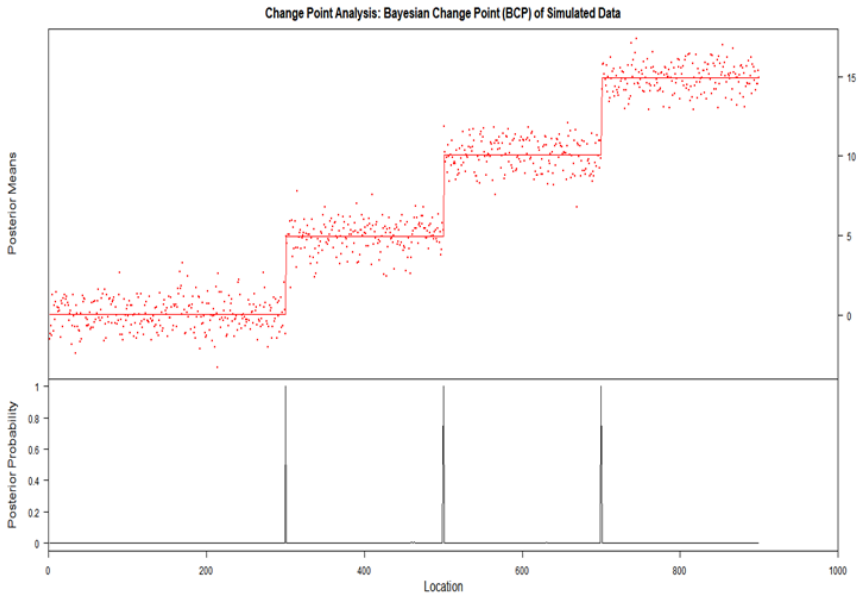



**Figure 8. Bayesian Change Point (BCP) analysis of simulated test data**

A plot of the BCP analysis of the test data is shown in Figure 8. For the 'bcp' plots the graphs are separated into two panels. The upper panel shows the original data and the posterior means. The lower panel shows the location of detected change points and the respective posterior probability. The BCP analysis detected the three change points indicating the shift occurring at locations 300, 500, and 700. As expected, these were the successive changes in the mean that were created in the simulated data. As illustrated by the three sharp spikes, the bcp analysis indicated that these three change points were statistically significant with a posterior probability of approximately 1.00.

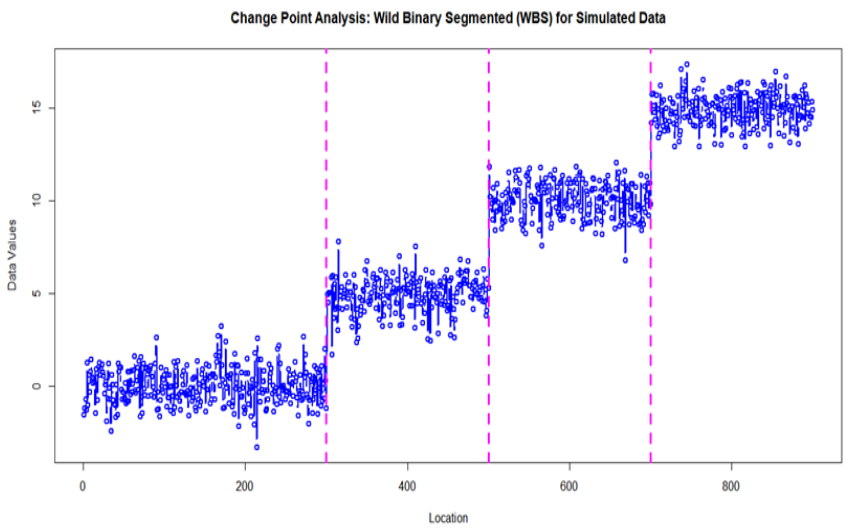



**Figure 9. Wild Binary Segmentation (WBS) for simulated test data**

The Wild Binary Segmentation (WBS) was also used to analyze the simulated data to capture the change points in the data. In a similar manner to the BCP method, the WBS method was able to detect the three change points at locations 300, 500, and 700. This is shown in Figure 9 with the three vertical dashed lines with the four distinct distributions. The WBS shows a complementary effect to the BCP method because it isolates the location of those statistically significant change points while the BCP method provides the posterior probability of isolated change points illustrated by the WBS method. The next section discusses the results of this tandem analysis (BCP and WBS methods) that was applied to the Navy, Army, Air Force projects' TCPI and TSPI time series data to determine the statistically significant change points.

### 4.2.2 United States Navy Projects

Figures 10 - 33 illustrates the Change Point Analysis (CPA) results of a sample of the Navy projects that were analyzed in this research. The figures demonstrate the BCP and WBS Change Point Analysis techniques being applied to the individual project performance trends concerning TCPI and TSPI. Figure 10 shows a graph of calculated TCPI data with respect to the project time periods. One can observe the significant growth of TCPI trend over time. The general trend profile was captured for every project before the Change Point Analysis methods were implemented.

From project to project, the BCP shows a split graph, as illustrated in Figures 11 and 15, with posterior mean on the top half and posterior probability distribution of the potential change points are shown in the bottom half. Similarly, the WBS results show the statistically significant change points, depicted as vertical magenta lines, as illustrate in Figures 12 and 16, superimposed on the TCPI and TSPI trends.

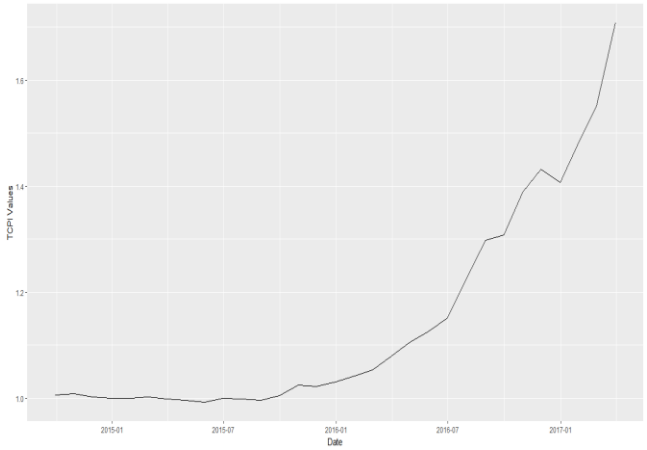



**Figure 10. Illustration of the time series trend of TCPI values for Project B5**

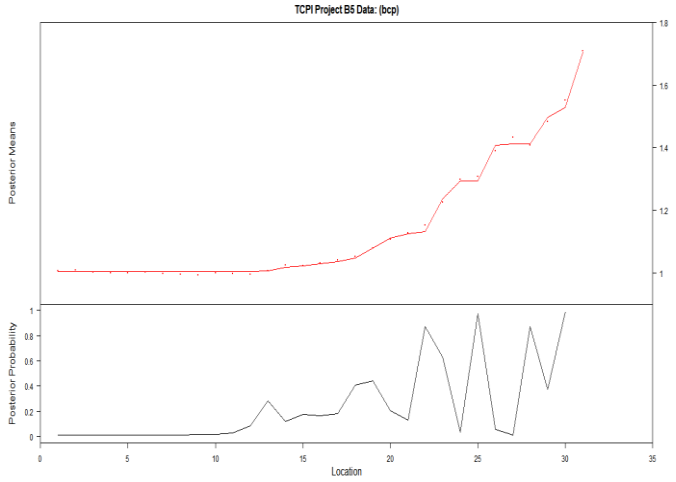



**Figure 11. Bayesian Change Point (BCP) Analysis of Project B5 TCPI Data**

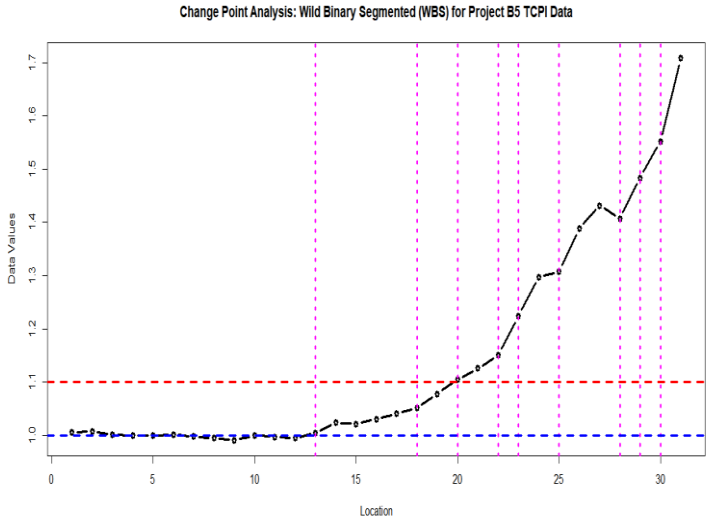



**Figure 12. Wild Binary Segmentation (WBS) Analysis for Project B5 TCPI Data**

Figures 13 and 17, depicts an example of the combination of the BCP and WBS techniques. Both methods demonstrate that as each of the projects progresses over time and the TCPI and TSPI trends exceed the threshold criterion, the numerous change points are detected with high posterior probability values. Comparable results are observed for several projects as shown in Figures 18-33.

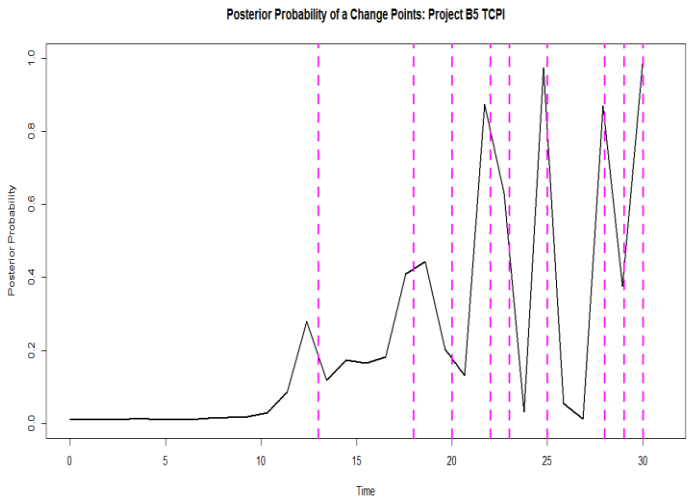



**Figure 13. Illustration of the combined posterior probability and change point locations (Project B5 TCPI Values)**

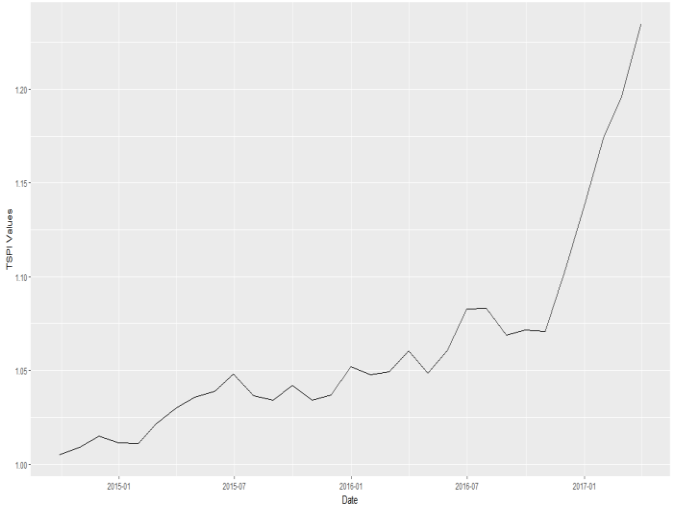



**Figure 14. Illustration of the time series trend of TSPI values for Project B5**

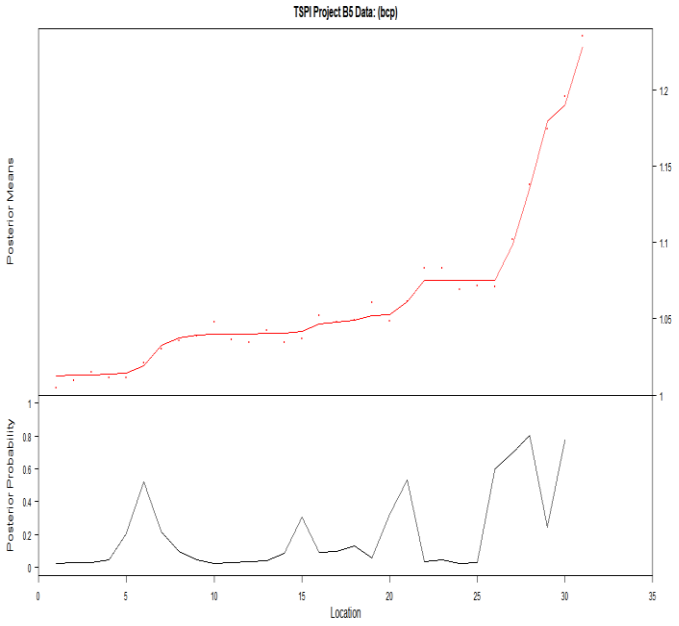



**Figure 15. Bayesian Change Point (BCP) Analysis of Project B5 TSPI Data**

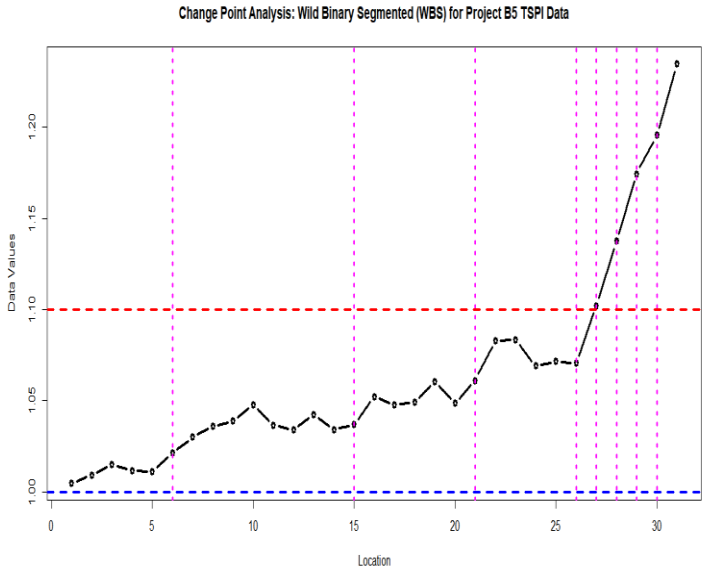



**Figure 16. Wild Binary Segmentation (WBS) Analysis for Project B5 TSPI Data**

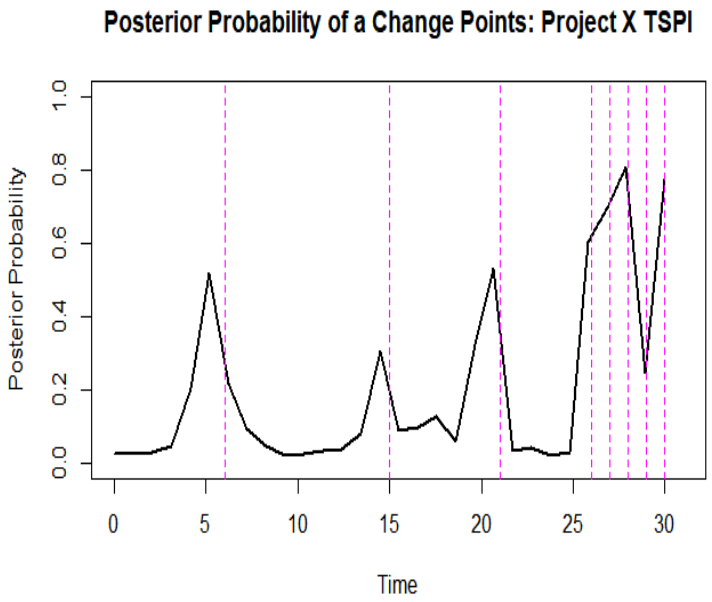



**Figure 17. Illustration of the combined posterior probability and change point locations (Project B5 TSPI Values)**

Like Project B5, the TCPI trend for Project H7 displays an interesting growth profile. As the project progresses over time, it the cost index is relatively flat with TCPI values close to 1.0. Then after some time frame, the index escalates gradually until it reaches orders of magnitude higher than initial values, see Figure 18. With the implementation of the BCP analysis method for the TCPI trend for Project H7, the results are shown in Figure 19. Some interesting observations can be observed from the posterior probability values. At the beginning of the project, on can see no significant changes have occurred as demonstrated by the TCPI profile in Figure 18, so the probability values are very low. These values start to increase dramatically as the TCPI trend starts increasing between 2015 and 2016.

The WBS method also shows a significant amount of change points being observed later in the project where the dramatic increase in the TCPI trend, see Figure 20.

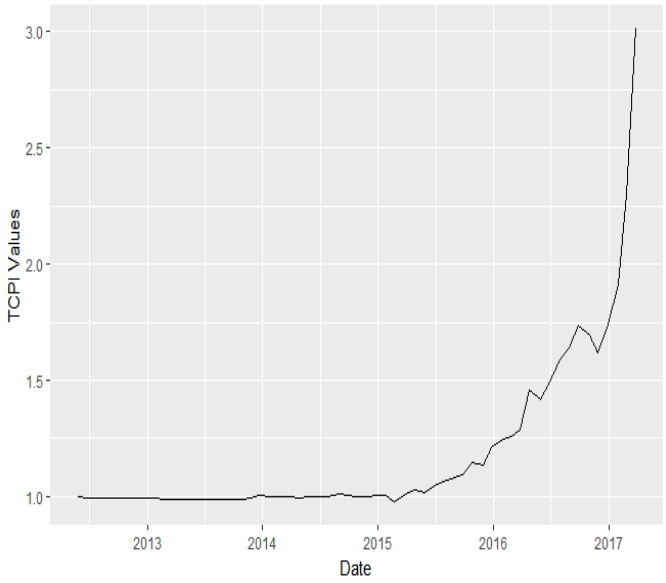



**Figure 18. Illustration of the time series trend of TCPI values for Project H7**

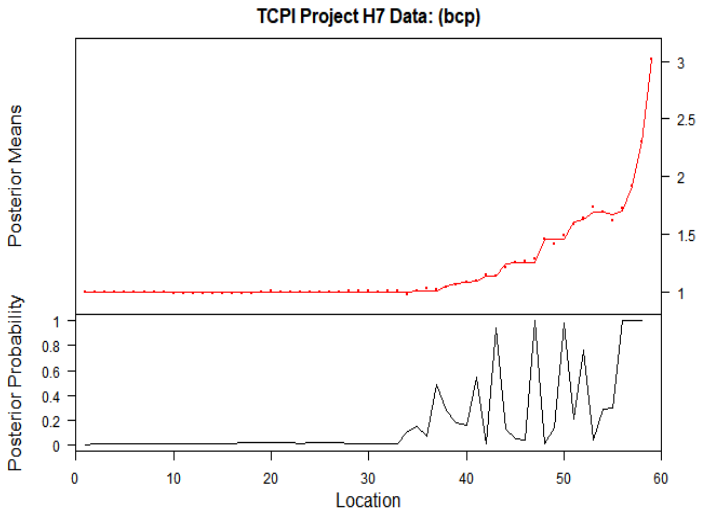



**Figure 19. Bayesian Change Point (BCP) Analysis of Project H7 TCPI Data**

When the WBS change points, depicted by the vertical magenta lines, were superimposed on the BCP posterior probability graph, see Figure 21, one can see that the

higher posterior probability values coincide with the numerous change points detected by the WBS change point analysis method.

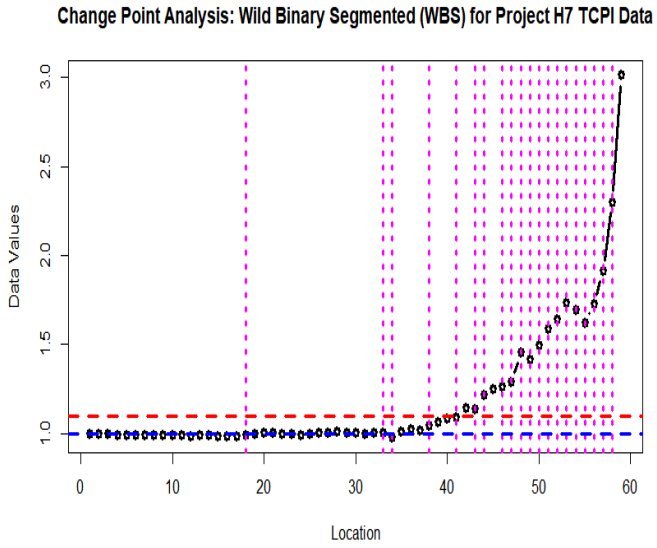



**Figure 20. Wild Binary Segmentation (WBS) Analysis for Project H7 TCPI Data**

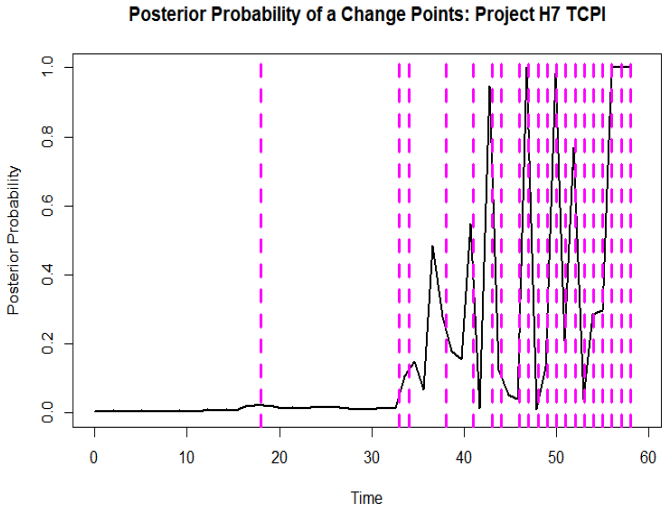



**Figure 21. Illustration of the combined posterior probability and change point locations (Project H7 TCPI Values)**

The TSPI profile for Project H7 also demonstrates this gradual increase towards escalating out of control, where the project will more than likely not meet the original project objectives. The results of BCP and WBS change point analysis techniques shown in Figures 23 and 24, respectively, indicates significant changes occurring beyond the proposed threshold criterion value of 1.10, shown as a dashed red line.

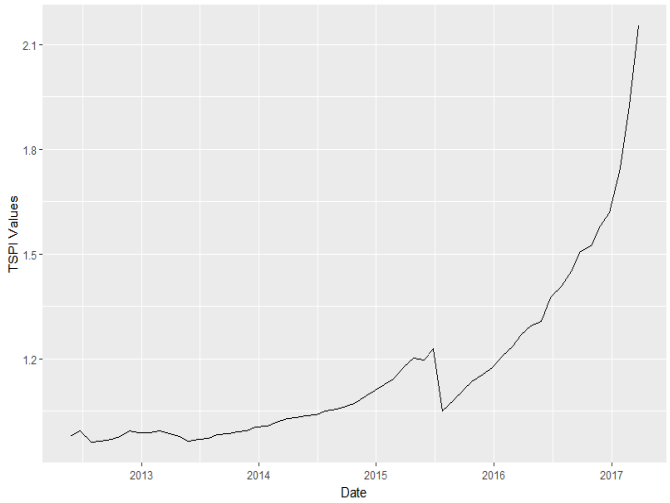



**Figure 22. Illustration of the time series trend of TSPI values for Project H7**

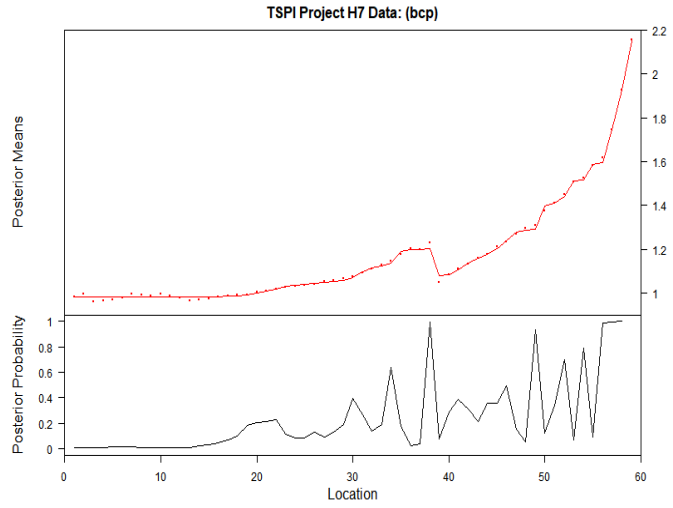



**Figure 23. Bayesian Change Point (BCP) Analysis of Project H7 TSPI Data**

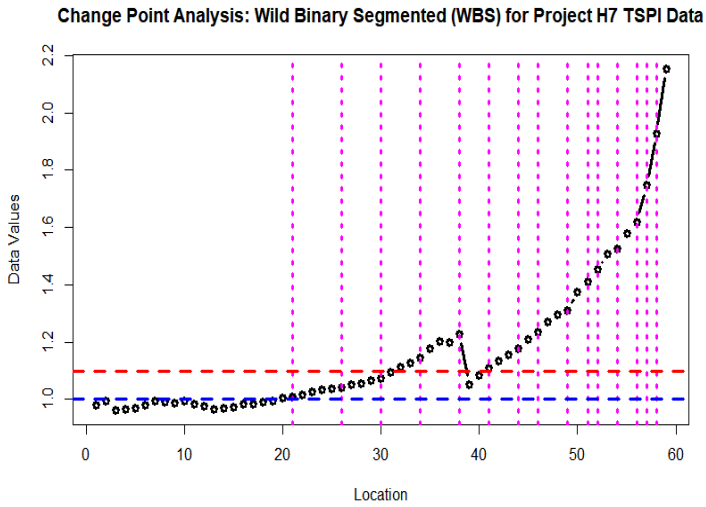



**Figure 24. Wild Binary Segmentation (WBS) Analysis for Project H7 TSPI Data**

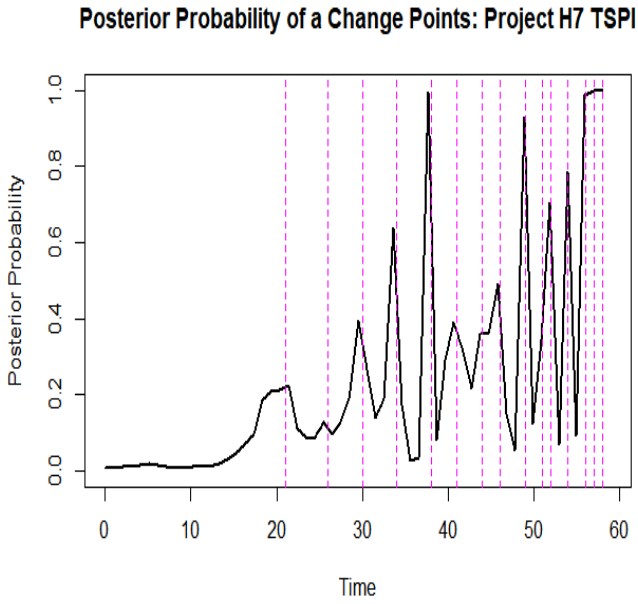



**Figure 25. Illustration of the combined posterior probability and change point locations (Project H7 TSPI Values)**

Project E9 TCPI trend profile shown see Figure 26 also illustrates a similar type profile as project B5 and H7. Again, the cost index starts out relatively flat and then at

some point starts to increase dramatically. As depicted in Figures 27 - 29, the BCP and WBS methods detect significant amount change points with high posterior probability values occurring when the project increases rapidly.

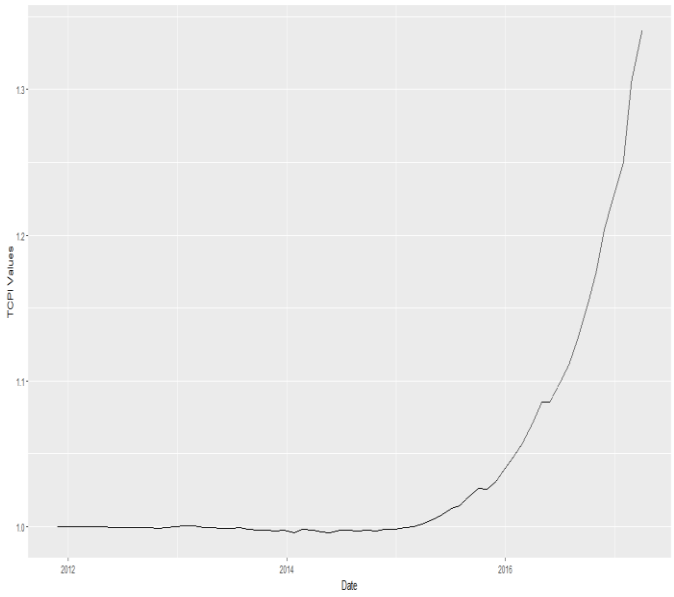



**Figure 26. Illustration of the time series trend of TCPI values for Project E9**

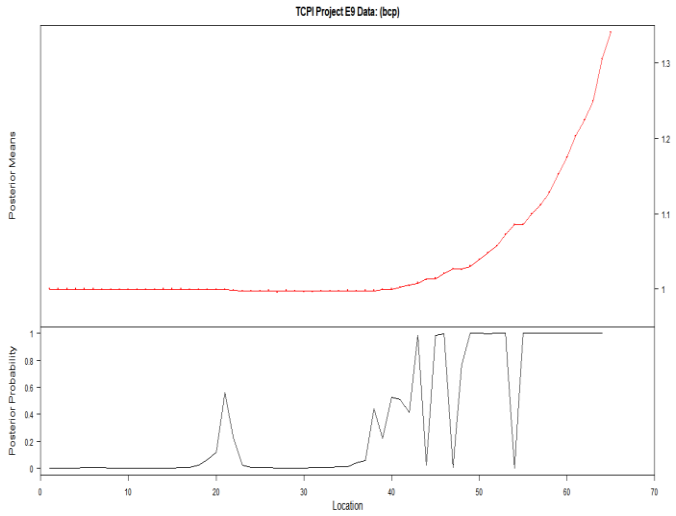



**Figure 27. Bayesian Change Point (BCP) Analysis of Project E9 TCPI Data**

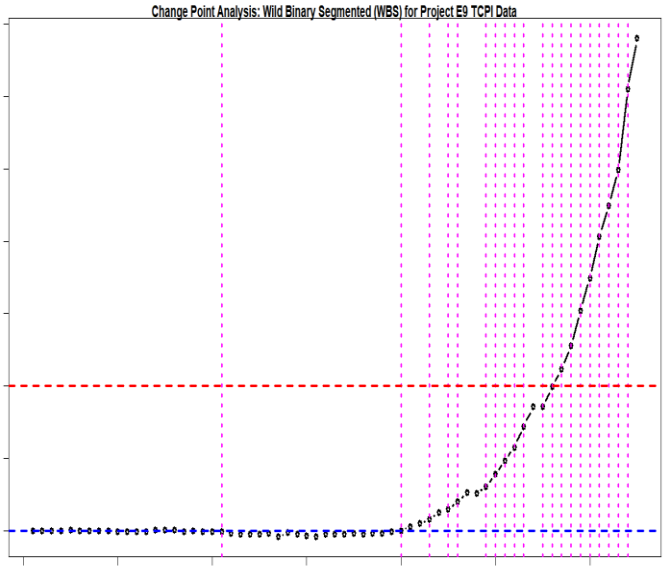



**Figure 28. Wild Binary Segmentation (WBS) Analysis for Project E9 TCPI Data**

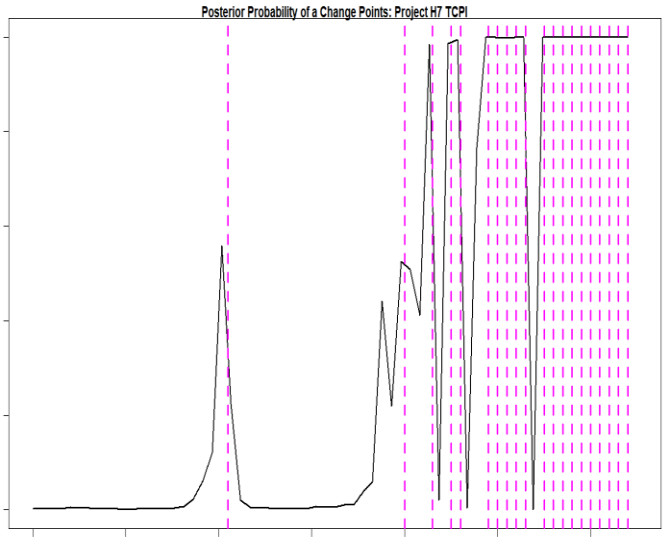



**Figure 29. Illustration of the combined posterior probability and change point locations (Project E9 TCPI Values)**

The TSPI trend for E9 also shows the growth profile over time. The change point techniques detect statistically higher change point values later in the project as observed in Figures 30-33.

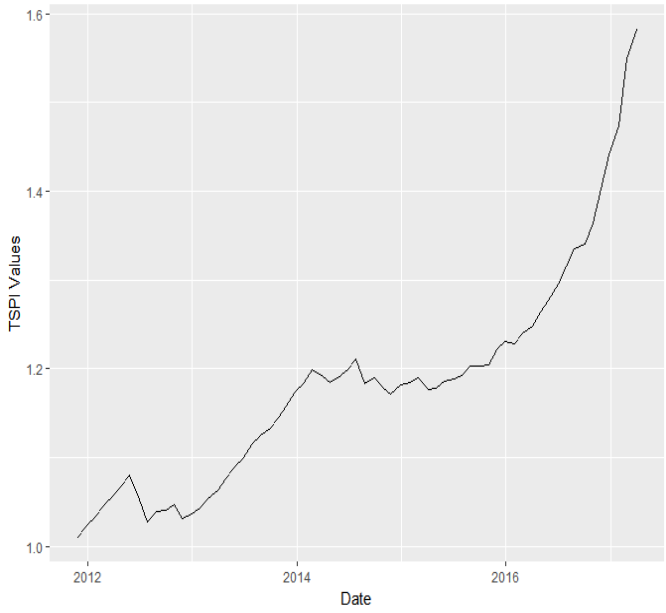



**Figure 30. Illustration of the time series trend of TSPI values for Project E9**

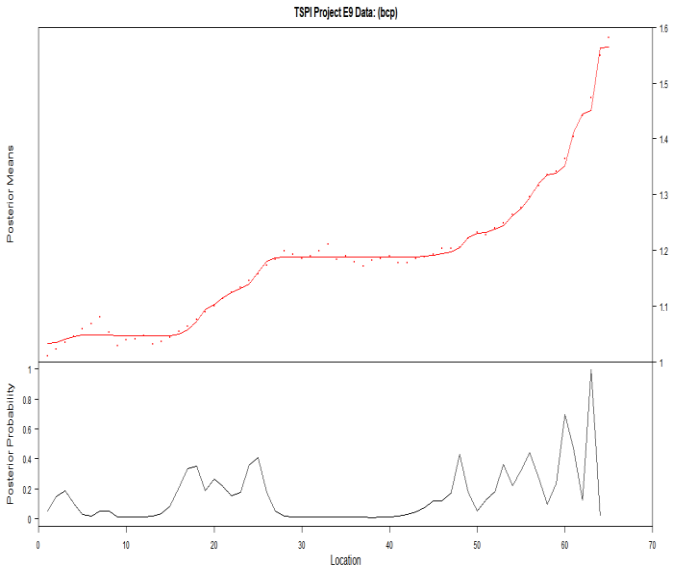



**Figure 31. Bayesian Change Point (BCP) Analysis of Project E9 TSPI Data**

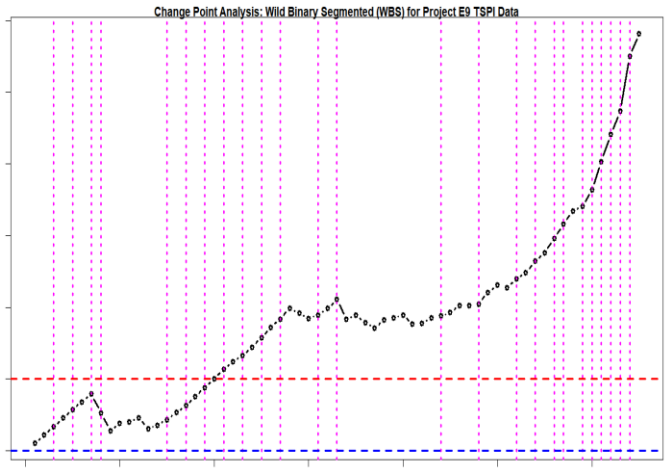



**Figure 32. Wild Binary Segmentation (WBS) Analysis for Project E9 TSPI Data**

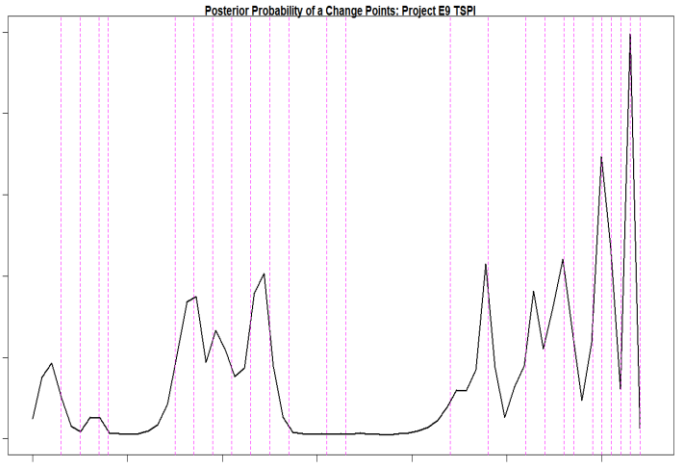



**Figure 33. Illustration of the combined posterior probability and change point locations (Project E9 TSPI Values)**

### 4.2.3 United States Army Projects

Figures 34- 37 illustrates the Change Point Analysis results of an Army project. The TCPI trend for Project FF1 also exhibits similar growth trends as the Navy projects. The

statistically significant change points with higher posterior probability values occur at or beyond the proposed threshold criterion of 1.10. No significant differences observed between the Army projects and Navy Projects.

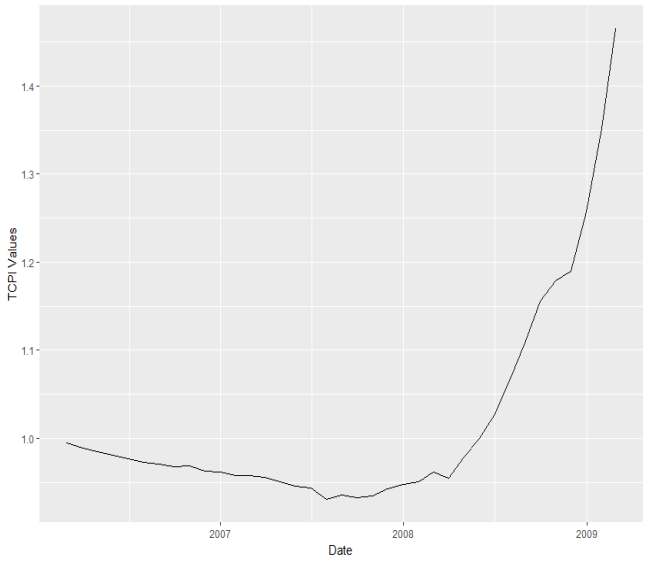



**Figure 34. Illustration of the time series trend of TCPI values for Project FF1**

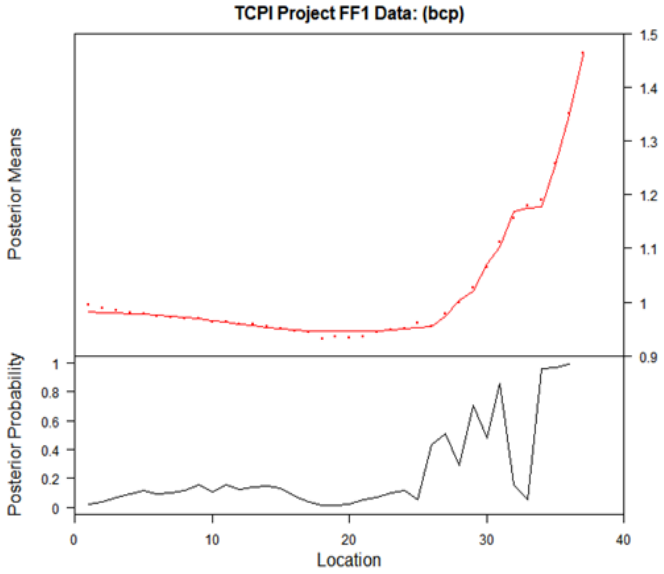



**Figure 35. Illustration Bayesian Change Point analysis of TCPI values for Project FF1**

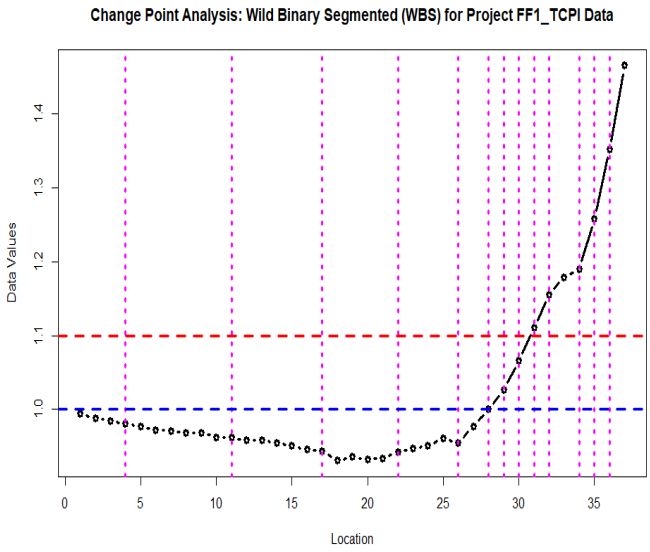



**Figure 36. Wild Binary Segmentation (WBS) Analysis for Project FF1 TCPI Data**

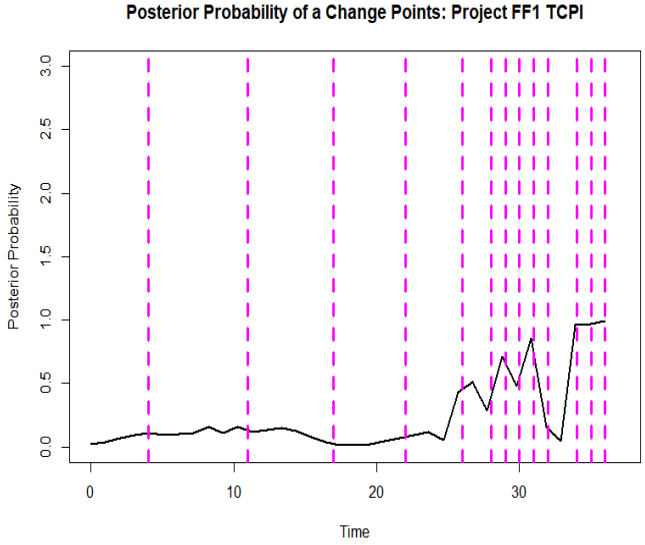



**Figure 37. Illustration of the combined posterior probability and change point locations (Project FF1 TCPI Values)**

### 4.2.4 United States Air Force Projects

Figures 38 - 41 illustrates the Change Point Analysis results of an Air Force project. As in the Navy and Army projects, one can observe similar TCPI growth trend for Project GGG1. As in the Navy and Army projects, the statistically significant change points with higher posterior probability values occur at or beyond the proposed threshold criterion of 1.10. No significant differences observed between the major branches of DoD (Navy, Army, and Air Force)

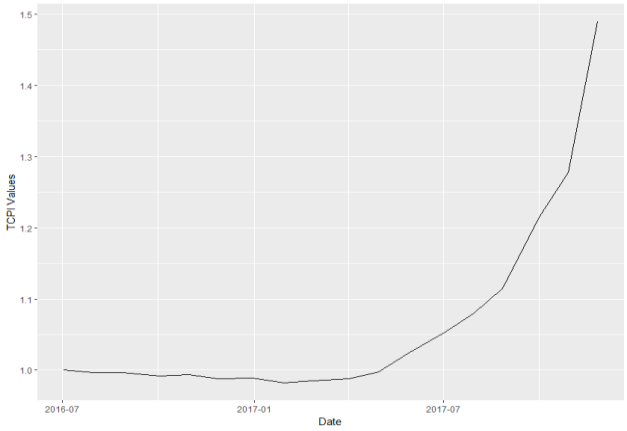



**Figure 38. Illustration of the time series trend of TCPI values for Project GGG1**

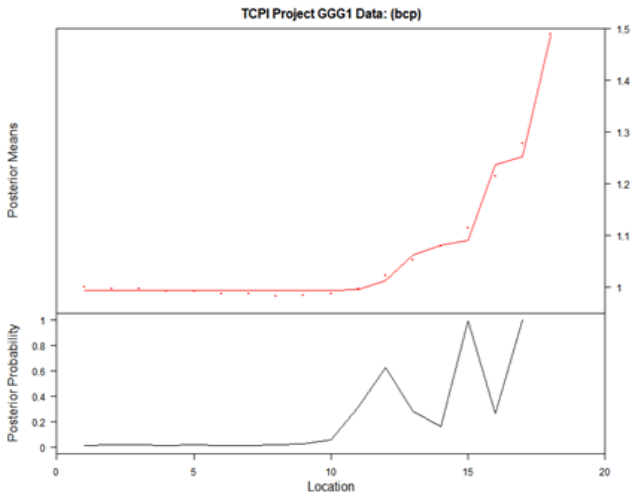



**Figure 39. Illustration Bayesian Change Point analysis of TCPI values for Project GGG1**

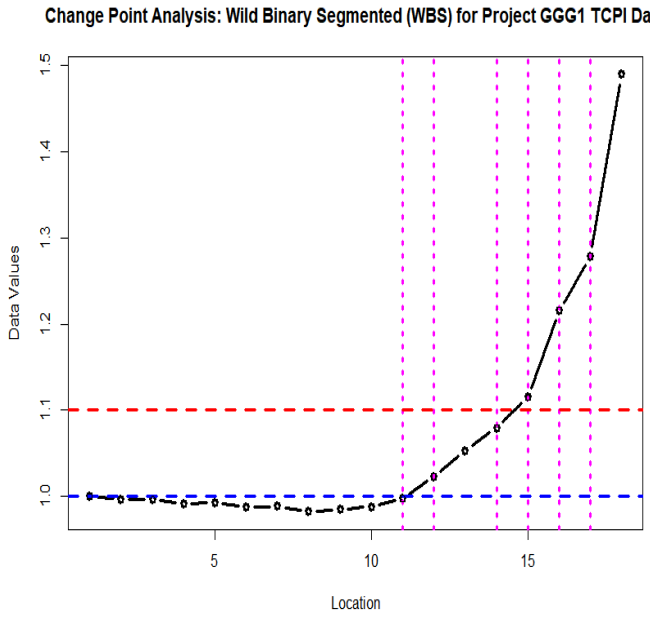



**Figure 40. Wild Binary Segmentation (WBS) Analysis for Project GGG1 TCPI Data**

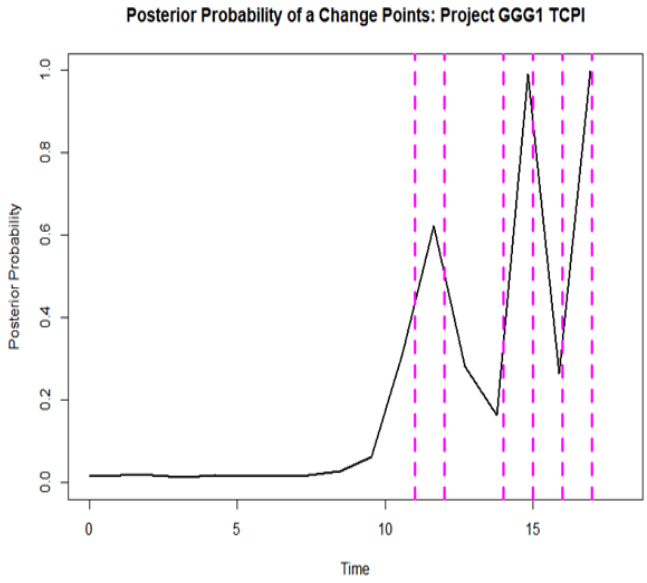
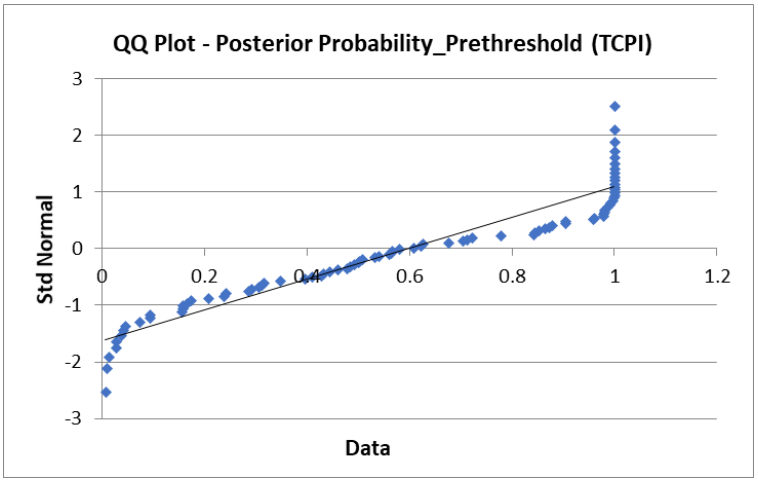


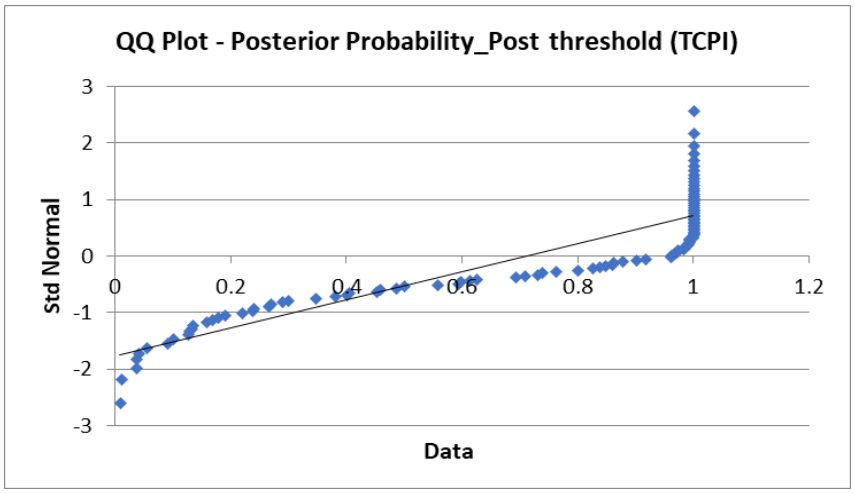

**Figure 41. Illustration of the combined posterior probability and change point locations (Project GGG1 TCPI Values)**

## 4.3 Hypothesis Testing Results

A normality test was conducted to determine what type of hypothesis testing will be more appropriate the examine the statistical difference between the pre-threshold and post-threshold posterior probability data. If the data is normally distributed then parametric hypothesis testing such as the t-test, F-test, and Analysis of Variance (ANOVA) (Huck, 2012). Otherwise, non-parametric test such as the Mann-Whitney U test, Wilcoxon-Whitney test, and Kruskal-Wallis should be used to avoid errors in the analysis (Huck, 2012). For this research, the Shapiro-Wilk test, see Tables II-V, was used to conduct normality testing. Also, a quantile plot, see Figures 42(a-b) and 43(a-b), was used to illustrate whether the data followed a normal distribution.




a)




b)

**Figure 42 (a-b). Illustration of the Q-Q plot for pre-threshold and post-threshold posterior probability data for TCPI**

Table II and Table III show the results of the Shapiro-Wilk Test for normality, with regards to the pre-threshold and post-threshold posterior probability data for TCPI.

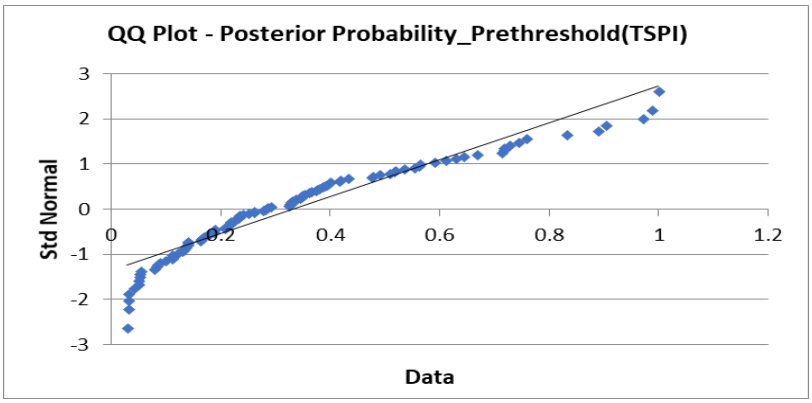
**Table II. Shapiro-Wilk Test for normality (Pre-Threshold TCPI)**

| Shapiro-Wilk Test | | |
|---|---|---|
| | | |
| *Posterior Probability_Prethreshold (TCPI)* | | |
| W-stat | 0.88526 | |
| p-value | 1.68392E-06 | |
| alpha | 0.05 | |
| normal | no | |

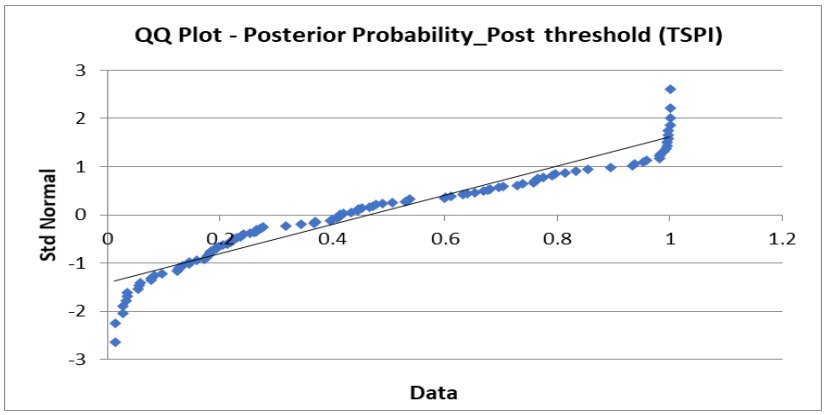
**Table III. Shapiro-Wilk Test for normality (Post-Threshold TCPI)**

| Shapiro-Wilk Test | | |
|---|---|---|
| | | |
| *Posterior Probability_Post threshold (TCPI)* | | |
| W-stat | 0.768730 | |
| p-value | 2.2486E-11 | |
| alpha | 0.05 | |
| normal | no | |

As shown in Figures 42(a-b), the quantile plots for the pre-threshold and post-threshold with regards to TCPI illustrate a non-normal distribution of the data. The data points exhibit a deviation from the straight line. The deviation from the linear pattern seems more prominent as the data increases towards the value of 1.0. As stated by Hanagal (2009), the further the data points move away from the straight line one can conclude that the data in question does not follow a normal distribution. To further verify this conclusion, the Shapiro-Wilk's test shows that the pre-threshold and post-threshold data with regards to TCPI were not normally distributed.




a)




b)

**Figure 43(a-b). Illustration of the Q-Q plot for pre-threshold and post-threshold posterior probability data for TSPI**

The normality test on the pre-threshold and post-threshold with regards to TSPI data was executed with the results shown in Figures 43(a-b). One can observe that these data points demonstrate a deviation from the straight line. Similar to the TCPI data, the pre-threshold and post-threshold with regards to TSPI data seem more prominent as the data increases towards the value of 1.0. Similarly, Table IV and Table V show the results of the Shapiro-Wilk Test for normality, with regards to the pre-threshold and post-threshold posterior probability data for TSPI. The Shapiro-Wilks test results show that the pre-threshold and post-threshold data with regards to TSPI were not normally distributed.

**Table IV. Shapiro-Wilk Test for normality (Pre-Threshold TSPI)**

| Shapiro-Wilk Test | | |
|---|---|---|
| | | |
| *Posterior Probability_Prethreshold(TSPI)* | | |
| W-stat | 0.91185 | |
| p-value | 1.88849E-06 | |
| alpha | 0.05 | |
| normal | no | |

**Table V. Shapiro-Wilk Test for normality (Post-Threshold TSPI)**

| Shapiro-Wilk Test | | |
|---|---|---|
| | | |
| *Posterior Probability_Post threshold (TSPI)* | | |
| W-stat | 0.91718 | |
| p-value | 2.34838E-06 | |
| alpha | 0.05 | |
| normal | no | |
| | | |

As shown in the test for normality, both the pre-threshold and post-threshold probability distribution data exhibit non-normality; therefore, a non-parametric test must be executed. For this dataset, the Mann-Whitney U test was used because it satisfies the

criteria for un-paired or independent samples. In addition, the Mann-Whitney also "doesn't make the assumptions about the homogeneity of the variances or any specific distribution of the data" (Dytham, 2003). Table VI and Table VII show the results of the statistical test for significance between the pre-threshold and post-threshold criterion change points for TCPI and TSPI, respectively. The Mann-Whitney Test was used to test the Null Hypothesis stated above.

**Table VI. Mann-Whitney Test (TCPI)**

| Mann-Whitney Test for Two Independent Samples | | |
|---|---|---|
| | **Pre-threshold (TCPI)** | **Post-threshold (TCPI)** |
| count | 85 | 102 |
| median | 0.5778 | 0.9592 |
| rank sum | 6996.5 | 10581.5 |
| U | 5328.5 | 3341.5 |
| | | |
| | **one tail** | |
| alpha | 0.05 | |
| mean | 4335 | |
| std dev | 366.87 | |
| z-score | 2.708 | |
| effect r | 0.198 | |
| U-crit | 3731.54 | |
| p-value | 0.0034 | |
| sig | yes | |

**Table VII. Mann-Whitney Test (TSPI)**

| Mann-Whitney Test for Two Independent Samples | | |
|---|---|---|
| | **Pre-threshold (TSPI)** | **Post-threshold (TSPI)** |
| count | 111 | 116 |
| median | 0.281 | 0.4109 |
| rank sum | 11175.5 | 14702.5 |
| U | 7916.5 | 4959.5 |
| | | |
| | **one tail** | |
| alpha | 0.05 | |
| mean | 6438 | |
| std dev | 494.61 | |
| z-score | 2.99 | |
| effect r | 0.1984 | |
| U-crit | 5624.43 | |
| p-value | 0.0014 | |
| sig | yes | |

The results indicate that there is a statistical difference between the posterior probability distribution values of the pre-threshold versus the post-threshold criterion. This suggests that the claim that when either the TCPI or TSPI trends exceed the 1.10 value, there is strong possibility that the project is out of control and there is a minimal probability that the project will recover and achieves its desired project performance objectives. The BCP and WBS tools illustrate as potential tools that can be applied to project performance trends to be analyzed as time series to capture significant changes in the trends.

# Chapter 5 – Discussion and Conclusion

This praxis research demonstrates that the theory regarding the project performance objectives when values of To Complete Performance Index (TCPI) and the To Complete Schedule Performance Index (TSPI) exceeded the threshold value of 1.10 also applies to DoD ACAT I programs. The figures above illustrate that unless immediate actions from Project Managers are taken, the project could escalate out of control beyond a point of no return, where the project may not be recoverable (Lipke, 2009a). It was shown from this research effort that TCPI and TSPI are performance predictors in EVM and has proven to be a possible bell weather for Major ACAT I regardless of the mission operations, service branch, or type of platform. With the understanding of the threshold value for both TCPI and TSPI, this will help Project managers save significant cost and avoiding considerable project delays by reducing the probability that the projects will escalate out of control.

This praxis is important because it provides empirical analysis on a broad spectrum of Major ACAT I DoD Projects with the examination of both TCPI and TSPI to determine the abrupt change in the behavior of project performance indicator trends to identify the statistical significance that a threshold criterion exists in the military environment. In addition, this research fits into the world of EVM by providing an analytical method that helps PMs monitor, track and detect potential issues with the project management metrics. The two metrics of importance are TSPI and TCPI. The central contribution is the characterization of TCPI and TSPI threshold values for Major ACAT I Programs. The Bayesian Change Point (BCP) and Wild Binary Segmentation (WBS) Analysis method extend this analysis by examining the probability distribution of

the change points of individual projects TCPI and TSPI trends. The method determines the change points with the highest level of statistical significance and their respective locations. The Welch-test for ANOVA results gives positive results regarding pre-threshold and post-threshold posterior distribution of change points detected.

In terms of the TCPI pre-threshold and post-threshold posterior probability values, the p-value was less than alpha = 0.05; therefore, one must reject the null hypothesis and accept the alternative. Similarly, the TSPI pre-threshold and post-threshold posterior probability values have a p-value that was less than alpha = 0.05. Which means that one must reject the null hypothesis and accept the alternative. With these results, one could conjecture that the claim that when either the TCPI or TSPI trends exceed the 1.10 value, there is strong possibility that the project is out of control and the probability that the project will recover to meet its desired project performance objectives is low. In addition, the change point analysis techniques, BCP and WBS, that were used in this praxis research illustrate as potential tools that can be applied to project performance trends to be analyzed as time series to capture significant changes in the trends.

## 5.1 Contributions

Previous research regarding this subject matter focused primarily on a small sample of private sector projects. This research was focused on major acquisition programs across a broad spectrum of DoD programs. Based on the literature review, it was determined that there were noticeable cultural differences between the military environment versus the private sector. Therefore, it was deemed to be value added for

conducting this research to examine the threshold criterion applied to military projects with a broad mission focus. The author's contribution stems from the fact that this research was the first to apply Change Point Analysis (CPA) technique in the analysis of performance indexes such TCPI and TSPI in the context of Major ACAT I program across DoD. This method has some level of novelty to its approach because, to best of the author's knowledge, no CPA methods have been applied to the To Complete Indexes (e.g. TCPI ad TSPI) with the idea of characterizing a threshold value that Engineering Managers or other practitioners in the field of Project Management could use to determine when a project will more than likely not meet the project objectives and may not recover. This potentially could be the demise of the project at worse and at best significant loss of time and money.

The results seem very promising that the CPA methods indicate a significant difference in the behavior of the TCPI and TSPI trends once the threshold criterion of 1.10 has been exceeded. The author does recommend researching other methods that could be used to characterize these trends. The next section offers additional suggestions that could be used for future development and validation of these performance indexes.

## 5.2 Future Directions

A future effort for this research could be to gain access to EVM data for inactive DoD programs for the possibility of developing a more robust predictive model and validation process in characterizing the threshold criterion. This could be tied to the analysis of the probability of project recovery which could be confirmed with projects that survived or canceled due to significant cost and schedule overruns. Another future effort for this

research could be to look at ACAT II programs that may be more or less complex as ACAT I programs. Could the behavior the TCPI and TSPI trends reveal any nuances between the various levels of acquisition programs in DoD?

# References


Aminikhanghahi, S. and Cook, D.J. (2017). A survey of methods for time series change point detection. *Knwl Inf. Syst., 51*(2), 339-367.

Anbari, F. (2003). Earned value project management method and extensions. *Project Management Journal*, 34(4), 12–23.

Arena, V. M., et al. (2005). Monitoring the progress of shipbuilding programmes: How can the Defense Procurement Agency more accurately monitor progress? Santa Monica, CA. RAND Corporation.

Barber, D., Cemgil, A.T., & Chiappa, S. (2011). *Bayesian time series models.* New York, NY: Cambridge University Press.

Barry, D., & Hartigan, J.A. (1993). A Bayesian analysis for change point problems. *Journal of The American Statistical Association*, 35(3), 309-319.

Brodsky, B.E., & Darkhovsky, B.S. (1993). *Non-parametric methods in change-point problem.* Dordrecht, Netherlands: Kluwer Academic Publishers.

Batselier, J., & Vanhoucke, M. (2015). Evaluation of deterministic state-of-the-art forecasting approaches for project duration based on Earned Value Management. *International Journal of Project Management*, 33, 1588–1596.

Crumrine, K.T. (2013). A comparison of earned value management and earned schedule as schedule predictors on DoD ACAT I programs (Master's Thesis). Air Force Institute of Technology, Wright-Patterson Air Force Base, OH.

D'Amico, C.N., White, E.D., & Ritschel, J.D. (2018). Unmasking cost growth behavior: A longitudinal study. *Defense Acquisition Research Journal*, 25(1), 30-51.

DAMIR. (2015). (Website). Retrieved from http://www.acq.osd.mil/damir/

Dytham, C. (2003). *Choosing and using statistics: a biologist's guide.* Malden, MA: Blackwell Publishing Company.

Erdman, C. and Emerson, J.W. (2003). bcp: an R package for performing a Bayesian analysis of change point problems. *Journal of Statistical Software*, 23(3), 1-11.

ES website: Retrieved from http://www.earnedschedule.com/Home.shtml.

Fleming, Q.W., & Koppelman, J.M. (2000). *Earned value: project management*. Newton Square, PA: *Project Management Institute.*

Fryzlewicz, P. (2014). Wild binary segmentation for multiple change-point detection. *The Annals of Statistics*, 42(6), 2243-2281.

Goure, D. (2015). Retrieved from http://www.lexingtoninstitute.org/wp-content/uploads/2015/10/Incentivizing-a-New-Defense-Industrial-Base.pdf

Hanagal, D.D. (2009). *Introduction to Applied Statistics: A Non-Calculus Based Approach.* Oxford, U.K. Alpha Science International, LTD.

Henderson, K. (2004). Further developments in earned schedule. *The Measurable News* (Spring), 15(6), 20–22.

---------- (2005). Earned schedule in action. *The Measurable News* (Spring), 30, 23–28.

Humphreys, G.C. (2016). *Project management using earned value.* Irvine, CA: Humphreys & Associates, Inc.

Huck, S.W. (2012). *Reading statistics and research.* Boston, MA: Pearson Education Inc.

Jacob, D. (2003). Forecasting project schedule completion with earned value

metrics. *The Measurable News* (March), 1, 7-9.

Jacob, D., & Kane, M. (2004). Forecasting schedule completion using earned value metrics? revisited. *The Measurable News* (March), 1, 11-17.

Kim, E., Wells, W. G., & Duffey, M. R. (2003). A model for effective implementation of earned value management methodology. *International Journal of Project Management*, 21, 2375–82.

Lam, K.C., Lee, D., & Hu, T. (2001). Understanding the effect of the learning-forgetting phenomenon to duration of projects. *International Journal of Project Management* 19, 411–20.

Larson, E.W., & Gray, C.F. (2014). *Project management: the managerial process.* New York, NY: McGraw-Hill Education.

Lehman, E.L. (1959). *Testing statistical hypotheses.* New York, NY: John Wiley & Sons, Inc.

Lipke, W. (2009a). *Earned schedule: An extension of earned value management*. Lulu Publishing.

Lipke, W., Zwikael, O., & Anbari, F. (2009b). Prediction of project outcome: the application of statistical methods to earned value management and earned schedule performance indexes. *International Journal of Project Management*, 27, 400–7.

Lipke, W. (2012). Earned schedule: contribution to project management. *Project Management World Journal*, 1(2), 1-19.

---------- (2014). Earned schedule—ten years after. *Project Management World*

*Journal*, 3(1), 1-11.

---------- (2016a). Examination of the threshold for the To Complete Indexes. *Project Management World Journal*, 5(3), 1–12.

---------- (2016b). The probability of project recovery. *Project Management World Journal*, 5(6), 1–11.

McCarthy, C.C. (2016). *Program management in defense and high-tech environments.* Boca Raton, FL: CRC Press.

Hossein-Mortaji, S.T., Noorossana, R., & Bagherpour, M. (2015). Project completion time and cost prediction using Change Point Analysis. *The Journal of Management Engineering*, *31*(5), 1–11**.**

PARCA. (2017). (Website). Retrieved from http://www.acq.osd.mil/parca/

PMBOK Guide. (2006). *Government extension to the project management body of knowledge (PMBOK guide).3rd Ed.* Newtown Square, Pa: Project Management Institute, Inc.

R Development Core Team, R: A language and environment for statistical computing, Vienna: R foundation for statistical computing, 1020, URL http://CRAN.R-project.org

Ramanathan, R. (1993). *Statistical methods in econometrics.* San Diego, CA: Academic Press, Inc.

Sharma, S., Swayne, D.A., & Obimbo, C. (2016). Trend analysis and change techniques: a survey. Energy Ecology Environment, 1(3), 123-130.

Shiryaev, A.N. (1963). On optimal methods in the quickest detection problems. Teor. Veroyatn. Primen, 8(1), 26-51.

Sommer, R. (2016). Changing public sector organizations from the middle. *International Journal of Business & Management*, 4(4), 321-325.

Vandevoorde, S. and Vanhoucke, M. (2006). A comparison of different project duration forecasting methods using earned value metrics. *International Journal of Project Management*, 24, 289–302.

Vanhoucke, M. (2012). Measuring the efficiency of project control using fictitious and empirical project data. *International Journal of Project Management*, 30, 252–63.

Walker, H.M. and Lev, J. (1953). *Statistical inference.* New York, NY: Holt, Rinehart and Winston, Inc.

Zhang, W., James, N.A., and Matteson, D.S. (2017). Pruning and nonparametric multiple change point detection. doi: arXiv-eprint arXiv:1709.06421

# Appendices

# APPENDIX – Praxis Research MATLAB Code

```
%%
@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@
@@@@@%
%   Name: Gregory Stamp
%
%Praxis Research EVM Analysis-To Complete % Performance
Indexes
%
%
@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@
@@@@@

%###################  PROGRAM H
###############################
%
%  Example of Earned Value Management(EVM)calculation using
MATLAB        %
%  This Calculation includes Earned Schedule (ES) values
calculated using %
%  Walt Lipke(2009)'s calculators from Earned Schedule
website.         %
%
%  Raw data are entered in the Lipke Calculator to
determine the values  %
%
%  for cumulative Earned Schedule(ES) and Probability of
Recovery       %
%
%@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@
@@@@@

clc;
clear;

%% #################### Start reading data from MS Excel
file
% Read in the Excel data; this put the numeric data
directly in xls_data,
% and puts all other text data in xls_text. Ensure to
include ES column
% from Lipke's Calculator

%######################### Program Data Extraction
#############
%
```

```
%############## Program H Contract Data:

 % filename_Type1 = 'ProgramH1_Rev1.xlsx';

 % filename_Type1 = 'ProgramH2_Rev1.xlsx';

 % filename_Type1 = 'ProgramH3_Rev1.xlsx';

 % filename_Type1 = 'ProgramH4_Rev1.xlsx';

%   filename_Type1 = 'ProgramH5_Rev1.xlsx';

%   filename_Type1 = 'ProgramH6_Rev1.xlsx';

%   filename_Type1 = 'ProgramH7_Rev1.xlsx';

%   filename_Type1 = 'ProgramH8_Rev1.xlsx';

%   filename_Type1 = 'ProgramH9_Rev1.xlsx';

%   filename_Type1 = 'ProgramH10_Rev1.xlsx';

%   filename_Type1 = 'ProgramH11_Rev1.xlsx';

%   filename_Type1 = 'ProgramH9_Rev2.xlsx';

% Determine which Excel worksheet
 sheet = 'Cummulative_noOverrun';


 %% ##################### Set Data Range File Type 1
######################
% xlRange =  'A2:S7'; %Data Range Program H1  Cummulative
noOverrun sheet
% xlRange =  'A3:S55'; %Data Range Program H2  Cummulative
noOverrun sheet

%  xlRange = 'A2:S57'; %Data Range Program H3  Cummulative
noOverrun sheet
%  xlRange = 'A2:S64'; %Data Range Program H4  Cummulative
noOverrun sheet
%    xlRange = 'A3:S55'; %Data Range Program H5
Cummulative noOverrun sheet


         xlRange = 'A2:S60'; %Data Range Program H6
Cummulative noOverrun sheet
```

```
%        xlRange = 'A2:S60'; %Data Range Program H7
Cummulative noOverrun sheet

%         xlRange = 'A2:S49'; %Data Range Program H8
Cummulative noOverrun sheet

%          xlRange = 'A2:S37'; %Data Range Program H9
Cummulative noOverrun sheet

%          xlRange = 'A3:S34'; %Data Range Program H10
Cummulative noOverrun sheet
% xlRange = 'A3:S23'; %Data Range Program H11  Cumulative
noOverrun sheet


%% ###################### Read in Data Matrix
############################

EVM_dataH1 = xlsread(filename_Type1,sheet,xlRange)
%
display('EVM Data Type 1 has been Acquired with no problem
for Program H!!')

%% ##Extract from EVM Data Matrix- parameter values from
the column data line

 AT=  EVM_dataH1(:,1); %Actual time
 PV = EVM_dataH1(:,2); % Planned Value
 EV = EVM_dataH1(:,3); %Earned Value
 AC = EVM_dataH1(:,4); %Actual Cost
 BAC = EVM_dataH1(:,5); %Budget at Completion
 ES =  EVM_dataH1(:,6); %Earned Schedule
 PD =  EVM_dataH1(:,9); %Planned Duration
 TD =  EVM_dataH1(:,10); %Total Duration Desired

 %################### NEW SEGMENT OF CODE
##############################

%
% % Project Completion Fraction(Based on Planned Schedule)
%  CompletionFrac_BAC = ES./PD;


%Fraction complete for the project defined by EV divided by
BAC.
 CompletionFrac_BAC = EV./BAC;
```

```
 %% ######### Start reading data from MS Excel file for
File Type 2 #######

%############## Program H Contract Data:

%  filename_Type2 = 'ProgramH1_Rev1.xlsx';

%  filename_Type2 = 'ProgramH2_Rev1.xlsx';

%  filename_Type2 = 'ProgramH3_Rev1.xlsx';

%  filename_Type2 = 'ProgramH4_Rev1.xlsx';

%      filename_Type2 = 'ProgramH5_Rev1.xlsx';

%      filename_Type2 = 'ProgramH6_Rev1.xlsx';

%      filename_Type2 = 'ProgramH7_Rev1.xlsx';

%      filename_Type2 = 'ProgramH8_Rev1.xlsx';

%      filename_Type2 = 'ProgramH9_Rev1.xlsx';

%      filename_Type2 = 'ProgramH10_Rev1.xlsx';

 %     filename_Type2 = 'ProgramH11_Rev1.xlsx';

%      filename_Type2 = 'ProgramH9_Rev2.xlsx';


%Establish Worksheet
sheet = 'ContractSummaryData';


%% ##########Set Data Range for FileType 2 Program H
######################

% xlRange = 'A4:K9'; %Data Range Program H1 Contractor Data
Summary sheet
%  xlRange = 'A5:K57'; %Data Range Program H2 Contractor
Data Summary sheet

% xlRange = 'A4:K59'; %Data Range Program H3 Contractor
Data Summary sheet
```

```
% xlRange = 'A4:K66'; %Data Range Program H4 Contractor
Data Summary sheet
%  xlRange =  'A5:K57'; %Data Range Program H5 Contractor
Data Summary sheet

   xlRange =  'A4:K62'; %Data Range Program H6 Contractor
Data Summary sheet

%   xlRange =  'A4:K62'; %Data Range Program H7 Contractor
Data Summary sheet

%      xlRange =  'A4:K51'; %Data Range Program H8
Contractor Data Summary sheet

%     xlRange =  'A4:K39'; %Data Range Program H9
Contractor Data Summary sheet

%      xlRange =  'A5:K36'; %Data Range Program H10
Contractor Data Summary sheet
%    xlRange =  'A5:K25'; %Data Range Program H11
Contractor Data Summary sheet



%% ######################### Read in Data Matrix FileType 2
#####

EVM_dataH2 = xlsread(filename_Type2,sheet,xlRange);

display('EVM Data Type 2 has been Acquired with no problem
for Program H!!')



 display('EVM Parameters Data Extraction Completed for
Prorgam H!')

%% ###Calculation of Project Performance Metrics in terms
of dollars($$)##

% j = 0

 for j = 1 : length(EVM_dataH1(:,1))

CV(j) = EV(j)-AC(j); %Cost Variance($$Dollars)
```

```
SV(j) = EV(j)-PV(j); %Schedule Variance based on $dollar
amount and not time

CPI(j) = EV(j)./AC(j); %Cost Performance Index
SPI(j) = EV(j)./PV(j); %Schedule Performance Index based on
$$Dollars

ratio_CPSP(j) = CPI(j)./SPI(j); % ratio of CPI TO
SPI;controlling downward trend; maintain upward trend

EAC(j) = AC(j)+((BAC(j)-EV(j))./CPI(j)); %Estimate at
Completion($$Dollars)


%@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@
@@@@@@@@@@@@
% Clculating Cost Forescating: Simple Method

 IEAC_CPI(j) = BAC(j)./CPI(j); %Forecast cost with
Performance Factor = CPI

 %Genreal Method for Cost Forecasting
%IEAC_CPI(j) = AC(j) + ((BAC(j)- EV(j))./PF(j)); %General
Forecast


 VAC_cost(j) = (BAC(j)-IEAC_CPI(j)); %Variance at
Completion for Cost

%@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@
@@@@@@@@@@@@@@



SCI_criticalratio(j) = CPI(j).*SPI(j); % Critical Ratio
CI(j) = 0.8.* CPI(j)+ 0.2.*SPI(j);

%% ##### Calculating the To Complete Perfomance Index(TCPI)
w.r.t Cost ####

% TCPI = Work Remaining/Funds Remaining
% The these methods differ based on calculation of Funds
Remmaining due
% what budget that management had authorized originally or
make budget (EAC)
```

```
%% ##################### Begin TCPI Calculations
#########################

EAC_calc = EAC; % Calculated based on Basic Raw EVM Data

EAC_table = EVM_dataH2(:,11);

CostRatio(j) = EAC(j)./BAC(j);% Cost Ratio



 ThreshValue = EVM_dataH1(:,7);
 TCPI_1_BAC(j) = (BAC(j) - EV(j))/(BAC(j)-AC(j));% TCPI
Calculation Method 1

 log_TCPI(j) = log(TCPI_1_BAC(j)); %log of TCPI values

 TCPI_2_EACcalc(j) = (BAC(j) - EV(j))/(EAC_calc(j)-AC(j));%
TCPI Calculation Method 2

 TCPI_3_EACtable(j) = (BAC(j) - EV(j))/(EAC_table(j)-
AC(j));% TCPI Calculation Method 3

  %################### NEW SEGMENT OF CODE
##############################

 if TCPI_1_BAC(j) > ThreshValue ;

        TCPI_ExceedCapture(j)= nonzeros(TCPI_1_BAC(j));

       CPI_capture(j) = CPI(j)

         display('Found where TCPI greater than
ThreshValue(1.10) with CPI')

%      PercentTCPI_ExceedCapture(j) = PercentComplete(j) ;

 else

      TCPI_ExceedCapture(j) = 0 ;
      CPI_capture(j)= 0;

%
%      PercentTCPI_ExceedCapture(j) = PercentComplete(j) ;
%
%      PercentTCPI_BelowCapture(j) = PercentComplete(j);
```

```
     TCPI_BelowCapture(j) = TCPI_1_BAC(j);


      display('Still Searching for the Greater Than
Threshold Value')
 end


 %################ Calculate TSPI
############################

  % ################## New for SPI Analysis
##################

  SPI_time(j) = ES(j)./AT(j); % Schedule Performance Index
w.r.t time

  SV_time(j) = ES(j)- AT(j); % Schedule Variance w.r.t time

%   EAC_t(j) = PD(j)./ SPI_time(j); % Estimate Schedule at
Completetion w.r.t time


%@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@
@@@@@@@@@@@@

% Calculating Schedule Forescating using Simple Form(Lipke,
2009)

% IEAC_SPI_time(j) = AT(j) + ((PD(j)- ES(j))./PF(j));
%Geneal Form

IEAC_SPI_time(j) = (PD(j))./SPI_time(j); %Forecast cost
with Performance Factor = SPI(t)


VAC_schedule(j)= PD(j) - IEAC_SPI_time(j) ; %Variance at
Completion for Cost


%@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@
@@@@@

 %######### Ppercent Completion from EVM Data TABLE
```

```
 TSPI_1(j) = (PD(j) - ES(j))./(TD(j)-AT(j));%Calculate TSPI
Based on Method I Total Duration


  log_TSPI(j) = log(TSPI_1(j)); % log of TSPI

 if  TSPI_1(j) > ThreshValue ;

     %display('Found where TSPI greater than or equal
ThreshValue(1.10)')

     TSPI_ExceedCapture(j)= nonzeros(TSPI_1(j))

    SPI_time_capture(j) =  nonzeros((SPI_time(j)))


   display('Found where TCPI greater than ThreshValue(1.10)
with CPI')

  %  PercentTSPI_ExceedCapture(j) = PercentComplete(j) ;


 else
      display('Still Searching for the Value');


       TSPI_ExceedCapture(j) = 0 ;

      SPI_time_capture(j)= 0 ;


%       PercentTSPI_ExceedCapture(j) = PercentComplete(j);
%
%       TSPI_BelowCapture(j) = TSPI_1(j);
%
%       PercentTSPI_BelowCapture(j) = PercentComplete(j);


 display('Still Searching for the Value')

 end


 j = j+1;

 end
```

```
 %%
##############################################################
#############
%############################################################
################

% Export TCPI and TSPI Captured Data After Proposed
Threshold Value = 1.10

%  ###################### Write TCPI vs CPI Data to Excel
File ##########


% xls_mat_1{1,1}=('TCPI vs. CPI:');
% xls_mat_1{2,1}=(' ');
% xls_mat_1{3,1}=('TCPI_Capture');
% xls_mat_1{3,2}=('CPI_Capture');
%
% for i=1:length(TCPI_ExceedCapture(1,:))
%     xls_mat_1{3+i,1}=TCPI_ExceedCapture(i);
%     xls_mat_1{3+i,2}=CPI_capture(i);
%
% end
%
% % Compute the statistics of the export data in excel
% % Input stats values several cells below the last Excel
Entry
%
%     xls_mat_1{i+6,1}=(' ');
%     xls_mat_1{i+6,2}=(' ');
%
%
%     Mod1_TCPICapture = nonzeros(TCPI_ExceedCapture)
%
%     Mod1_CPICapture = nonzeros(CPI_capture)
%
%     xls_mat_1{i+7,1}=('Average(mean)');
%     xls_mat_1{i+7,2}=('Average(mean)')
%
%
%     xls_mat_1{i+8,1}=mean( Mod1_TCPICapture);
%     xls_mat_1{i+8,2}=mean( Mod1_CPICapture);
%
%     xls_mat_1{i+9,1}=('Standard deviation');
%     xls_mat_1{i+9,2}=('standard deviation')
%
%
```

```
%     xls_mat_1{i+10,1}=std(Mod1_TCPICapture);
%     xls_mat_1{i+10,2}=std(Mod1_CPICapture);
%
%
%     [p1,h1,stats] = signtest(
Mod1_CPICapture,Mod1_TCPICapture)
%
%     xls_mat_1{i+11,1}=(' ');
%     xls_mat_1{i+11,2}=(' ');
%
%     xls_mat_1{i+12,1}=('p-value');
%     xls_mat_1{i+12,2}=( 'h(hypothesis)')
%
%     xls_mat_1{i+13,1}=p1;
%     xls_mat_1{i+13,2}= h1;
%
%    xlswrite('TCPI_CPI_ProgramH_output.xls',xls_mat_1);
%
%
%  %% Write TSPI vs SPI_time Data to Excel File
% xls_mat_2{1,1}=('TSPI vs. SPI_time:');
% xls_mat_2{2,1}=(' ');
% xls_mat_2{3,1}=('TSPI_Capure');
% xls_mat_2{3,2}=('SPI_time_Capture');
%
% for i=1:length(TSPI_ExceedCapture(1,:))
%     xls_mat_2{3+i,1}=TSPI_ExceedCapture(i);
%     xls_mat_2{3+i,2}= SPI_time_capture(i);
% end
%


% % Compute the statistics of the export data in excel
% % Input stats values several cells below the last Excel
Entry
%
%     xls_mat_2{i+6,1}=(' ');
%     xls_mat_2{i+6,2}=(' ');
%
%     Mod2_TSPICapture = nonzeros(TSPI_ExceedCapture)
%
%     Mod2_SPICapture = nonzeros(SPI_time_capture)
%
%
%     xls_mat_2{i+7,1}=('Average(mean)');
%     xls_mat_2{i+7,2}=('Average(mean)')
%
```

```
%
%
%     xls_mat_2{i+8,1}=mean(Mod2_TSPICapture);
%     xls_mat_2{i+8,2}=mean(Mod2_SPICapture);
%
%     xls_mat_2{i+9,1}=('Standard deviation');
%     xls_mat_2{i+9,2}=('Standard deviation')
%
%
%     xls_mat_2{i+10,1}=std(Mod2_TSPICapture);
%     xls_mat_2{i+10,2}=std(Mod2_SPICapture);
%
%
%     [p2,h2,stats] = signtest(
Mod2_SPICapture,Mod2_TSPICapture);
%
%     xls_mat_2{i+11,1}=(' ');
%     xls_mat_2{i+11,2}=(' ');
%
%     xls_mat_2{i+12,1}=('p-value');
%     xls_mat_2{i+12,2}=('h(hypothesis)')
%
%     xls_mat_2{i+13,1}= p2;
%     xls_mat_2{i+13,2}= h2;
%
%
%  % % write the data to excel file
%  xlswrite('TSPI_SPI_time_ProgramH_output.xls',xls_mat_2);
%

%#########################################################
#################
%#########################################################
#################


%%
@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@
@@@@@@@@@@@@@
%@@@@@@@@@@@@@@@@@@@  NEW SECTION OF CODE
@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@
%@@@@@@@@@@@@@@@@@@ Exporting Data To Excel File
@@@@@@@@@@@@@@@@@@@@@@@@@

 % Write TCPI and TSPI Values to Excel File for Further
Data Analysis
```

```
xls_mat_1{1,1}=('Project H: Exported TCPI and TSPI Data for
Additional Analysis');


xls_mat_1{2,1}=(' ');
xls_mat_1{3,1}=('Percent Completion');
xls_mat_1{3,2}=('TCPI Values');
xls_mat_1{3,3}=('CPI Values');

%xls_mat_1{3,3}=('Log of TCPI Values');

xls_mat_1{3,4}=('TSPI Values'); ;
xls_mat_1{3,5}=('SPI(t) Values');


% xls_mat_1{3,5}=('Log of TSPI Values');


 for i=1:length(CompletionFrac_BAC(:,:))

    xls_mat_1{3+i,1}= (CompletionFrac_BAC(i)).*100;
    xls_mat_1{3+i,2}= TCPI_1_BAC(i);
    xls_mat_1{3+i,3}= CPI(i);
    xls_mat_1{3+i,4}=TSPI_1(i);
    xls_mat_1{3+i,5}= SPI_time(i);

 end



 xlswrite('ProjectH_TCPI_and_TSPI Values for
DataAnalysis',xls_mat_1);

 %@@@@@@@@@@ END OF Code for Exporting Data TCPI and TSPI
Data @@@@@@@@@@@@



%% ######################### Graph Basic EVM Data: Key Graph
###############
Index_delin =  EVM_dataH1(:,8);

timeperiod = 1:length(EVM_dataH1(:,1));



figure(1)
```

```
 plot(CompletionFrac_BAC, EV, '-.dg','MarkerSize',7)

hold on
%plot(CompletionFrac_BAC, AT, 'bd-','MarkerSize',7)
plot(CompletionFrac_BAC, PV, '-.>b','MarkerSize',7)
plot(CompletionFrac_BAC, AC, '-.sr','MarkerSize',7)
plot(CompletionFrac_BAC, BAC, '-.ok','MarkerSize',7)

% plot(CompletionFrac_BAC, EAC_table, '-
.>m','LineWidth',1,'MarkerSize',7)

plot(CompletionFrac_BAC, IEAC_CPI(:,:), '-
.>g','LineWidth',1,'MarkerSize',7)

title(sprintf('EVM Data Graph:S-Curve (Project
H)'),'Color', 'b', 'Fontsize', [12])
%xlabel('Time Periods(months)')
xlabel('Project Completion Fraction')

ylabel('Dollars($)')
legend('Earned Value(EV)', 'Planned Value(PV)', 'Actual
Cost(AC)','Budget At Completion(BAC)','Indepndent Estimate
at Completion(IEAC)_CPI','Location','Best')

%'Actual Time(AT)'

grid on



%% ######################### Forecasting Cost and VAC Graph
############################

 figure(3)

 hold on

plot(CompletionFrac_BAC, IEAC_CPI(:,:), '-
.>g','LineWidth',1,'MarkerSize',7)

plot(CompletionFrac_BAC, VAC_cost(:,:), '-
.>b','LineWidth',1,'MarkerSize',7)

plot(CompletionFrac_BAC, BAC, '-.ok','MarkerSize',7)
```

```
title(sprintf('Forecasting Project Outcomes and Variance at
Completion(cost): Project H'),'Color', 'b', 'Fontsize',
[12])
xlabel('Project Completion Fraction')
ylabel('Dollars($)')
legend('IEAC Cost','VAC Cost','Budget At Completion(BAC)'
,'ratio_CPSP','Location','Best')

grid on


%% ######################### Forecasting Schedule and
VAC(t) Graph ############################

VAC_schedule


 figure(4)

 hold on

plot(CompletionFrac_BAC, IEAC_SPI_time(:,:), '-
.>m','LineWidth',1,'MarkerSize',7)

plot(CompletionFrac_BAC, VAC_schedule(:,:), '-
.>b','LineWidth',1,'MarkerSize',7)

plot(CompletionFrac_BAC, PD, '-.ok','MarkerSize',7)



title(sprintf('Forecasting Project Outcomes and Variance at
Completion(schedule): Project H'),'Color', 'b', 'Fontsize',
[12])
xlabel('time periods')
ylabel('Indexes Values')
legend('IEAC Schedule','VAC Schedule','Planned
Duration(PD)','Location','Best')

grid on


%% ############## TCPI and CPI Indexes Graphs Comparison:
Key Graph #######
```

```
%
% figure(6)
%
% hold on
%
% plot(timeperiod, TCPI_1_BAC(:,:), '-.>b','LineWidth',1
,'MarkerSize',7)
% %plot(timeperiod, TCPI_2(:,:), '-.bo','MarkerSize',7)
% plot(timeperiod, CPI(:,:),  '-.m*','MarkerSize',7)
% plot(timeperiod, TSPI_1(:,:), '-
.>g','LineWidth',1,'MarkerSize',7)
% plot(timeperiod, SPI(:,:),     '-.dr','MarkerSize',7)
% plot(timeperiod, SPI_time(:,:),'-.sk','MarkerSize',7)
%
% figure(6)
% plot(timeperiod, ThreshValue,'-.b*','MarkerSize',7)
% plot(timeperiod, Index_delin,'-.k*','MarkerSize',7)
%
% title(sprintf('Critical EV Metric (Project H): TCPI,
TSPI, CPI, and SPI Indexes'),'Color', 'b', 'Fontsize',
[10])
% xlabel('timeperiod(months)')
% ylabel('Index Values')
% legend('TCPI_1BAC
Method','CPI','TSPI','SPI','SPI_time','Threshold
Value(1.10)','Index-delin(1.0)','Location','Best')
%
% grid on
% ylim auto
% xlim auto

%% ########### TCPI and CPI Indexes Graphs- EAC Calculation
Comparison ####

figure(7)
hold on

plot(CompletionFrac_BAC, TCPI_1_BAC(:,:), '-
.>b','MarkerSize',7)
%plot(CompletionFrac_BAC, TCPI_2_EACcalc(:,:), 'rd-.')
%plot(CompletionFrac_BAC, TCPI_3_EACtable(:,:), '-.>m')

 %plot(CompletionFrac_BAC , CostRatio,'-
.dm','MarkerSize',7)

 plot(CompletionFrac_BAC, CPI(:,:),'-.dm','MarkerSize',7)
```

```
figure(7)
plot(CompletionFrac_BAC, ThreshValue,'-.r*','MarkerSize',7)
plot(CompletionFrac_BAC, Index_delin,'-.k*','MarkerSize',7)

title(sprintf('TCPI and Cost Performance Index(CPI)w.r.t
Completion Fraction(Project H)'),'Color', 'b', 'Fontsize',
[9])
xlabel('Project Completion Faction')
ylabel('Index Values')
legend('TCPI_1BAC',' Cost Performance
Index(CPI)','Threshold Value(1.10)','Index-
delin(1.0)','Location','Best')

%,'TCPI_3EACtable'

grid on

 ylim auto
 xlim auto

%% #### TCPI and CPI Indexes Graphs w.r.t Completion
Fraction: Key Graph ###

 %######### Calcuated percent completion
%
%  figure(8)
% %
%  hold on
%
% plot(CompletionFrac_BAC , TCPI_1_BAC(:,:), '-
.>b','LineWidth',1 ,'MarkerSize',7)
% %plot(timeperiod, TCPI_2(:,:), '-.bo','MarkerSize',7)
% plot(CompletionFrac_BAC , CPI(:,:),  '-
.m*','MarkerSize',7)
% plot(CompletionFrac_BAC , TSPI_1(:,:), '-
.>g','LineWidth',1,'MarkerSize',7)
% plot(CompletionFrac_BAC , SPI(:,:),     '-
.dr','MarkerSize',7)
%
% %
% title(sprintf('Critical EV Metric (Project H): TCPI,
TSPI, CPI, and SPI Indexes'),'Color', 'b', 'Fontsize',
[10])
% xlabel('Project Completion Fraction')
% ylabel('Index Values')
```

```
% legend('TCPI_1BAC
Method','CPI','TSPI','SPI','SPI_time','Threshold
Value(1.10)','Index-delin(1.0)','Location','Best')
% ylim auto
% xlim auto

%xlim([0 100])
%
% grid on
%########################## Capture TCPI and Percent
Completion #########
%
% figure(9)
%
% hold on
%
% plot(CompletionFrac_BAC , SPI_time(:,:),'-
.sk','MarkerSize',7)
%
%  figure(8)
% plot(CompletionFrac_BAC , ThreshValue,'-
.b*','MarkerSize',7)
% plot(CompletionFrac_BAC , Index_delin,'-
.k*','MarkerSize',7)
% plot(PercentTCPI_ExceedCapture , TCPI_ExceedCapture, 'ko-
')
%
% %plot(PercentTCPI_BelowCapture , TCPI_BelowCapture, 'gd-
')
%
% plot(PercentComplete , CPI(:,:),'m-*')
%
% plot(timeperiod , CostRatio,'gd-')
%
% % figure(9)
% plot(PercentComplete , ThreshValue,'b*-')
%
% title(sprintf('TCPI Values Greater Than or Equal to
Threshold Value w.r.t Percent Completion'),'Color', 'b',
'Fontsize', [10])
% xlabel('Percent Completion(%)')
% ylabel('TCPI Values')
% legend('TCPI-
Exceed','CPI','CostRatio','ThreshValue(1.10)','Location','B
est')
%
```

```
% %legend('TCPI-Exceed','TCPI-Below','CPI',
'ThreshValue','Location','Best')
% %axis([0 100 0 2])
% ylim auto
% xlim auto



%############################ Capture TSPI and Percent
Completion #########
%  figure(10)
%
%  hold on
%
% plot(PercentTSPI_ExceedCapture , TSPI_ExceedCapture, 'ko-
')
%
% %plot(PercentTSPI_BelowCapture , TSPI_BelowCapture, 'gd-
')
%
% plot(PercentComplete , SPI_time(:,:),'r-*')
%
% figure(10)
% plot(PercentComplete , ThreshValue,'b*-')
%
%  title(sprintf('TSPI Values Greater Than or Equal to
Threshold Value w.r.t Percent Completion'),'Color', 'b',
'Fontsize', [10])
% xlabel('Percent Completion(%)')
% ylabel('TSPI Values')
% legend('TSPI-Exceed','SPI-time',
'ThreshValue(1.10)','Location','Best')
%
% %legend('TSPI-Exceed','TSPI-Below','SPI-time',
'ThreshValue','Location','Best')
% %axis([0 100 0 2])
% ylim auto
% xlim([0 100])

%############################ TCPI and Reformulated TCP
w.r.t percentEV #########
%  figure(11)
%
%  hold on
%
%  plot(timeperiod, TCPI_1_BAC(:,:), 'gd-')
%  plot(timeperiod, y(:,:), 'bo-')
```

```
% %plot(PercentTSPI_ExceedCapture , TSPI_ExceedCapture,
'ko-')
%
% %plot(PercentTSPI_BelowCapture , TSPI_BelowCapture, 'gd-
')
%
%
% plot(timeperiod, TheoryTCPI_rateIncrease(:,:), 'ko-')
%
% figure(11)
% plot(timeperiod , ThreshValue,'b*-')
%
% plot(timeperiod , CostRatio,'m*-')
%
% %plot(timeperiod , x,'y*-')
%
%
%  title(sprintf('TCPI and Reformulated TCPI w.r.t
percentEV'),'Color', 'b', 'Fontsize', [10])
% xlabel('timeperiod(months)')
% ylabel('TCPI Values')
% legend('TCPI-BAC','Y-TCPI',
'rateIncrease_TCPI','ThreshValue(1.10)','CostRatio','Locati
on','Best')
%
% %legend('TSPI-Exceed','TSPI-Below','SPI-time',
'ThreshValue','Location','Best')
% %axis([0 100 0 2])
% ylim auto
% xlim([0 100])



%EmpiricalTCPI_rateIncrease(j)


%############################ TCPI and Reformulated TCP
w.r.t percentEV #########
%  figure(11)
%
%  hold on
%
%  plot(timeperiod, TCPI_1_BAC(:,:), 'gd-')
% %plot(timeperiod, EmpiricalTCPI_rateIncrease(:,:), 'ko-')
%
% figure(11)
% plot(timeperiod , CostRatio,'m*-')
```

```
%
% plot(timeperiod , Ratio_PV2AC,'y*-')
%
% plot(timeperiod , ThreshValue,'b*-')
% plot(timeperiod, Index_delin,'k*-')
%
%
% %plot(timeperiod , x,'y*-')
% %
%  title(sprintf('TCPI and Cost Ratio w.r.t
duration'),'Color', 'b', 'Fontsize', [10])
% xlabel('timeperiod(months)')
% ylabel('TCPI Values')
% legend('TCPI-BAC','Cost
Ratio(EAC/BAC)','Ratio(PV/AC)','ThreshValue(1.10)','Locatio
n','Best')
%
% %'CostRatio'
%
% %legend('TSPI-Exceed','TSPI-Below','SPI-time',
'ThreshValue','Location','Best')
% %axis([0 100 0 2])
% ylim auto
% xlim auto

%############################ EAC, BAC, and AC W.r.t percent
Completion #########
%  figure(12)
%
%  hold on
%
% plot(PercentComplete, AC, 'ro-')
% plot(PercentComplete, BAC, 'ko-')
% plot(PercentComplete, EAC, 'b^-')
%
% figure(12)
%
%  title(sprintf('BAC AND EAC Cost w.r.t percent
Completion'),'Color', 'b', 'Fontsize', [10])
% xlabel('PercentComplete(%)')
% ylabel('Dollar Values')
% legend('AC','BAC','EAC','Location','Best')
% %'CostRatio'
% %legend('TSPI-Exceed','TSPI-Below','SPI-time',
'ThreshValue','Location','Best')
% %axis([0 100 0 2])
% ylim auto
```

```
% xlim([0 100])
%########################## EAC -time based  vs Planned
Duration #########
%  figure(13)
%
%  hold on
% %
% plot(timeperiod, TSPI_1(:,:), '-.sy')
% plot(timeperiod, SPI(:,:),     '-.sr')
% plot(timeperiod, SPI_time(:,:),'-.dg')
%
% plot(timeperiod, EAC_t(:,:),'-.om')

%% ########################## TSPI vs SPI_time #########
 figure(13)
  hold on
 plot(CompletionFrac_BAC, TSPI_1(:,:), '-
.>g','LineWidth',1,'MarkerSize',7)

 plot(CompletionFrac_BAC, SPI_time(:,:),'-
.dm','MarkerSize',7)

figure(13)
plot(CompletionFrac_BAC, ThreshValue,'-.r*','MarkerSize',7)
plot(CompletionFrac_BAC, Index_delin,'-.k*','MarkerSize',7)

title(sprintf('TSPI and Schedule Performace Index w.r.t
Completion Fraction(Project H)'),'Color', 'b', 'Fontsize',
[9])
xlabel('Project Completion Fraction')
ylabel('Index Values')
legend(' TSPI',' Schedule Performace Index(t)','Threshold
Value(1.10)','Index-delin(1.0)','Location','Best')

%,'TCPI_3EACtable'

grid on

 ylim auto
 xlim auto
```

# APPENDIX- Change Point Analysis using with R Code

```
#@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@
@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@
@@@@@@@
#  Name: Gregory C. Stamp
#  Praxis Research EVM Analysis-To Complete Performance Indexes
##@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@
@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@
@@@@@
library("strucchange")
library("ecp")
library("cpm")
library("trend")
library("xlsx")
library("bcp")
library('quantreg')
library("XLConnect")
library("ggplot2")
library("MASS")
library("AID")
library("readxl")
library("scales")
library(iplots)
```

```
library(dplyr)
library(rgr)
library(qqplotr)
library(wbs)
library(aTSA)
library(timeSeries)
library(fUnitRoots)
library(urca)
getwd()
setwd( "C:/Users/GSTAMP3/Documents")
############### Simulated Data known Change Point
################################
#Establish bayesian test data for simulation purposes
#testdata1 <- c(rnorm(50), rnorm(50, 10, 1),rnorm(100))
#############################################################################
######
#######  BEGININNING OF SIMULATED BCP FOR
RESEARCH######################
##### A random sample from a few normal distributions ############
 #testdata1 <- c(rnorm(300, 0, 1), rnorm(200, 5, 1),rnorm(200, 10, 1), rnorm(200, 15, 1))
 #hist(testdata1, col = 'gray90', freq = FALSE, main = "Histogram of Simulated Data
Values",
  #  xlab = "Data Values", ylab = "Density")
```

```
 #lines(density(testdata1), col = 4, lwd = 3)
 #qqnorm(testdata1)#TCPI Quantil-Quantile Plot
 #qqline(testdata1, col = "blue", lwd =2, na.rm=TRUE)
  ############### BCP FPR SIMULTED DATA
###################################
 #bcp.1 <- bcp(testdata1)
 #summary( bcp.1)
 #plot( bcp.1, main="Change Point Analysis: Bayesian Change Point (BCP) of Simulated
Data")
# bcp1.prob <-  bcp.1$posterior.prob
 #bcp1.prob
  ########### Wild Binary Segmentation (wbs) for multiple change-point detection
####
 #w <- wbs(testdata1)
 #w.cpt <- changepoints(w)
 #w.cpt
 #plot(x = w$x, y = NULL, type= "b", col = "blue", lwd = 2,
  #    main = "Change Point Analysis: Wild Binary Segmented (WBS) for Simulated
Data", xlab = "Location", ylab = "Data Values")
  #abline(v = c(w.cpt$cpt.th[[1]]), col = "magenta", lty = 20, lwd = 3)

############## A random sample lognormal distribution distributions
###################
```

```
 # testdata2 <- c(rlnorm(50, meanlog = 0, sdlog = 1), rlnorm(150, meanlog = 0, sdlog =
1),rlnorm(100, meanlog = 0, sdlog = 1))
 ############# create a vector of histogram breaks
##################################
  #x_brk <- seq(0,max(testdata2),length=700)

 ########## Create histogram the data
  #hst <- hist(testdata2, breaks=x_brk)
 ################  Fitting the lognormal distribution ####################
   #fit_params <- fitdistr(testdata,"lognormal")
 ####### generate values given our fit parameters ########################
     #fit <- dlnorm(x_brk, fit_params$estimate['meanlog'], fit_params$estimate['sdlog'])

############plot the fit and original distributions#############
 # plot(x_brk, fit, type="l", ylab="Density", xlab="X", ylim=c(0,max(hst$density)),
xlim=c(0,10))
 #title(main = "Density histogram with lognormal fit")
 #lines(hst$mid, hst$density, type="l", col="red")
 # legend(6,1.5,legend=c("Fit","Data"),lty=c(1,1),col=c("black","red"))
  #   bcp.2 <- bcp(fit)
  #    summary(bcp.2)
  #   plot(bcp.2, main="Lognormal Distribution Change Point Example")
  #    bcp2.prob <-  bcp.2$posterior.prob
```

```
 #    bcp2.prob
##########  END OF SIMULATED BCP FOR RESEARCH
##########################
#########################################################################
######



########### Obtain TCPI AND TSPI Data For CPA
##################################
mydata1 <-read_xlsx("NonstationaryDetect_ProjectGGG5_TCPI_Rev1.xlsx", sheet =
"Sheet1", range = NULL, col_names = TRUE)
mydata1
mydata2 <-read_xlsx("NonstationaryDetect_ProjectGGG5_TSPI_Rev1.xlsx", sheet =
"Sheet1", range = NULL, col_names = TRUE)
mydata2
################## Individual Projects Dataset Analysis
############################
x1<-mydata1[,1] # Date Range for Project
x1 <- as.matrix(x1)
y1<-mydata1[,2] # #TCPI Values
y1 <- as.matrix(y1)
y1
```

```
##@@@@@@@@@@@@@  Write Data to Excel
File@@@@@@@@@@@@@@@@@@@@
write.xlsx(y1, file = "TCPI_TSPI_Data_ProjectGGG5.xlsx",
 sheetName = "TCPI_y1", append = TRUE)
#########################################################################
######
x2<-mydata2[,1] # Date Range for Project
x2 <- as.matrix(x2)
y2<-mydata2[,2] ## TCPI Values
y2 <- as.matrix(y2)
y2

##@@@@@@@@@@@@@  Write Data to Excel
File@@@@@@@@@@@@@@@@@@@@
write.xlsx(y2, file = "TCPI_TSPI_Data_ProjectGGG5.xlsx",
          sheetName = "TSPI_y2", append = TRUE)
######################### Examine the Data Quality
################################
################################## Histogram Plots of Data
######################
##################### TCPI Data ################
```

```
hist(y1, freq = FALSE) #TCPI Histogram
lines(density(y1), col = 4)
qqnorm(y1)#TCPI Quantil-Quantile Plot
qqline(y1, col = "blue", lwd =2)
shapiro.test(y1)
ks.test(y1,pnorm,mean=mean(y1),sd=sd(y1))
acf(y1)
###################### TSPI Data #################3
hist(y2, freq = FALSE)  #TSPI Histogram
lines(density(y2), col = 4)
qqnorm(y2)#TCPI Quantil-Quantile Plot
qqline(y2, col = "blue", lwd =2)
shapiro.test(y2)
ks.test(y2,pnorm,mean=mean(y2),sd=sd(y2))
acf(y2)
##########################################################################
######
#plot(mydata$`TCPI Value`~as.Date(mydata$`Report Date`,format = ""),type="l",
xlab="Date", ylab="TCPI")
#plot(mydata$`TCPI Value`~as.Date(mydata$`Report Date`, format = "%b %y", las =
2),type="l")
#axis.Date(side = 1, dates, at = mydata$`Report Date`, format = "%b %y", las = 2)
```

```
#########################################################################
######
mydata1 = mydata1[,c("Report Date", "TCPI Value")]
mydata1 = na.omit(mydata1);
head(mydata1)
attach(mydata1)

mydata1[1]
ggplot(mydata1, aes(`Report Date`, `TCPI Value`)) +  geom_line()  + xlab("Date") +
ylab("TCPI Values")
mydata2 = mydata2[,c("Report Date", "TSPI Value")]
mydata2 = na.omit(mydata2);
head(mydata2)
attach(mydata2)
ggplot(mydata2, aes(`Report Date`, `TSPI Value`)) +  geom_line()  + xlab("Date") +
ylab("TSPI Values")
############################### Execute Bayesian Change Point Analysis
#############
#################  Determine BCP Change Point
#################################

################## TCPI BCP
################################################
```

```
TCPI_bcp <- bcp(y1, burnin=500, mcmc=5000, return.mcmc=TRUE)
summary(TCPI_bcp)
plot(TCPI_bcp,  colors = NULL, main="TCPI Project GGG5 Data: (bcp)", col = 1,
na.rm=TRUE)
##@@@@@@@  Write Data to Excel File
@@@@@@@@@@@@@@@@@@@@@@
write.xlsx(TCPI_bcp$posterior.prob, file = "TCPI_TSPI_Data_ProjectGGG5.xlsx",
        sheetName = "TCPI_postprob", append = TRUE)
write.xlsx(TCPI_bcp$posterior.mean, file = "TCPI_TSPI_Data_ProjectGGG5.xlsx",
        sheetName = "TCPI_postmean", append = TRUE)
########### Wild Binary Segmentation (wbs) for TCPI Data ####
w <- wbs(y1)
w.cpt <- changepoints(w)
w.cpt
plot(x = w$x, y = NULL, type= "b", col = "black", lwd = 3,
    main = "Change Point Analysis: Wild Binary Segmented (WBS) for Project GGG5
TCPI Data", xlab = "Location", ylab = "Data Values")

abline(v = c(w.cpt$cpt.th[[1]]), col = "magenta", lty = 3, lwd =3)
abline(h=1.00, col="blue",lwd=3, lty=2)
abline(h=1.10, col="red",lwd=3, lty=2)
```

```
##@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@ Write Data
to Excel File
@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@
write.xlsx(w.cpt$cpt.th, file = "TCPI_TSPI_Data_ProjectGGG5.xlsx",
        sheetName = "WBS_ChangePointTCPI", append = TRUE)
chpts1<-as.matrix(w.cpt$cpt.th)
######################## Set X-rane to match the number of time series data
TCPI#######
xax <- seq(0, 19, length=19)
#op <- par(mfrow=c(2,1),col.lab="black",col.main="black")
############### Posterior Mean Separate Plot with WBS Change Points imbeded
##########
#plot(xax, bcp.ri$posterior.mean, type="l", xlab="Location", ylab="Mean",
main="Posterior Means")
#lines(xax, bp.ri, col="blue")
bcp.ri<-TCPI_bcp
bp.ri <- rep(0,length(y1))
points(y1)
plot(xax, bcp.ri$posterior.prob, type="l", ylim=c(0,1), xlab="Time", ylab="Posterior
Probability", main="Posterior Probability of a Change Points: Project GGG5 TCPI", lwd
= 2)
for(i in 1:length(bp.ri)) abline(v = c(w.cpt$cpt.th[[i]]), col = "magenta", lty = 20, lwd=3)
```

```
###################################### Residuals OF TCPI Data
##################
TCPI.resid<-residuals(TCPI_bcp)
hist(TCPI.resid, freq = FALSE) #TCPI Histogram
lines(density(TCPI.resid), col = 4, lwd = 2)
#qqnorm(TCPI.resid)#TCPI Quantil-Quantile Plot
#qqline(TCPI.resid, col = "blue", lwd =2)
#shapiro.test(TCPI.resid)
#ks.test(TCPI.resid,pnorm,mean=mean(TCPI.resid),sd=sd(TCPI.resid))

#legacyplot(TCPI_bcp)
#@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@
@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@
@@@@@@
############################### TSPI BCP
##################################
TSPI_bcp <- bcp(y2, burnin=500, mcmc=5000, return.mcmc=TRUE)
summary(TSPI_bcp)
plot(TSPI_bcp,  colors = NULL, main="TSPI Project GGG5 Data: (bcp)", col = 1,
na.rm=TRUE)

##@@@@@@@@@@@@@  Write Data to Excel File
@@@@@@@@@@@@@@@@
```

```
write.xlsx(TSPI_bcp$posterior.prob, file = "TCPI_TSPI_Data_ProjectGGG5.xlsx",
        sheetName = "TSPI_postprob", append = TRUE)
write.xlsx(TSPI_bcp$posterior.mean, file = "TCPI_TSPI_Data_ProjectGGG5.xlsx",
        sheetName = "TSPI_postmean", append = TRUE)
########### Wild Binary Segmentation (wbs) for TCPI Data ####
w2 <- wbs(y2)
w.cpt2 <- changepoints(w2)
w.cpt2
plot(x = w2$x, y = NULL, type= "b", col = "black", lwd = 3,
    main = "Change Point Analysis: Wild Binary Segmented (WBS) for Project GGG5
TSPI Data", xlab = "Location", ylab = "Data Values")
abline(v = c(w.cpt2$cpt.th[[1]]), col = "magenta", lty = 3, lwd =3)
abline(h=1.00, col="blue",lwd=3, lty=2)
abline(h=1.10, col="red",lwd=3, lty=2)
##@@@@@@@@@@@@@@  Write Data to Excel File
@@@@@@@@@@@@@@@@@@
write.xlsx(w.cpt2$cpt.th, file = "TCPI_TSPI_Data_ProjectGGG5.xlsx",
        sheetName = "WBS_ChangePoint_TSPI", append = TRUE)
chpts2<-as.matrix(w.cpt2$cpt.th)



############# Set X-range to match the number of time series data for
TSPI###########
```

```
xax <- seq(0, 19, length=19)
#op <- par(mfrow=c(2,1),col.lab="black",col.main="black")
############### Posterior Mean Separate Plot
#################################
#plot(xax, bcp.ri$posterior.mean, type="l", xlab="Location", ylab="Mean",
main="Posterior Means")
#lines(xax, bp.ri, col="blue")
bsp.ri<-TSPI_bcp
bp.ri <- rep(0,length(y2))
points(y2)
plot(xax, bsp.ri$posterior.prob, type="l", ylim=c(0,1), xlab="Time", ylab="Posterior
Probability", main="Posterior Probability of a Change Points: Project GGG5 TSPI", lwd
= 2)
for(i in 1:length(bp.ri)) abline(v = c(w.cpt2$cpt.th[[i]]), col = "magenta", lty = 20)
############################ Residuals OF TSPI Data
###################33333
TSPI.resid<-residuals(TSPI_bcp)
hist(TSPI.resid, freq = FALSE) #TCPI Histogram
lines(density(TSPI.resid), col = 4, lwd = 2)
qqnorm(TSPI.resid)#TCPI Quantil-Quantile Plot
qqline(TSPI.resid, col = "blue", lwd =2)
#Check for autocorrelation of the data
acf(TSPI_bcp)
```

```
shapiro.test(TSPI.resid)

ks.test(TSPI.resid,pnorm,mean=mean(TSPI.resid),sd=sd(TSPI.resid))

legacyplot(TSPI_bcp)

##########################################################################
```