A Protocol to Assess the Impact of Crude Oil and Fuel Price Fluctuations on Future Asphalt Prices in the State of Alabama: A Stochastic Risk Assessment Approach

by

Kaylyn M. Cardinal

A thesis submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Master of Science in Civil Engineering

Auburn, Alabama
August 4, 2018

Keywords: Asphalt, Cost Estimation, Crude Oil,
Cumulative Sum Statistical Analysis, Fuel, and
Risk Analysis

Approved by

Dr. Jorge Rueda-Benavides, Chair, Assistant Professor of Civil Engineering
Dr. David Timm, Brasfield and Gorrie Professor of Civil Engineering
Dr. Wesley Donald, Research Fellow IV of Civil Engineering

ABSTRACT

Since the introduction of asphalt in the year 1870, it has become one of the main components of horizontal construction. According to the National Asphalt Paving Association (NAPA), asphalt covers approximately 94 percent of the 2.7 million miles of paved roads in the United States (U.S.), and it contributes approximately $80 billion dollars to the horizontal construction industry annually. With bitumen, a byproduct of crude oil, as a critical ingredient of asphalt, the projected price of asphalt has been based on an assumed relationship with crude oil and fuel.

Multiple publications, researchers, estimators, engineers, and other proponents of the horizontal construction industry have mentioned this correlation, yet there is a lack of research defining these items as dependent on each other. Using crude oil and fuel price indexes, this thesis researches the long held hypothesis of their relationship with asphalt price indexes. Based on the monthly-posted asphalt, crude oil, and fuel price indexes by the Alabama Department of Transportation (ALDOT) and U.S Energy Information Administration (EIA), a stochastic risk assessment tool using a modified cumulative sum (CUSUM) statistical analysis method was created to determine their relationship.

The risk assessment tool enables the prediction of the time line of asphalt price changes along with the magnitude of the changes. Risks associated with those changes based on relationships with crude oil and fuel were answered. In terms of using crude oil price indexes to predict asphalt price index changes, the results indicated that the most likely time gap between them was 3 months and the percent change was 58 percent. This means that for a 1 percent change

in the crude oil price index, 58 percent of that change would most likely be reflected in asphalt's price index 3 months later.

In terms of using fuel to predict asphalt's price changes, the results indicated that the most likely time gap between them was 2 months and the percent change was 46 percent. This means that for any 1 percent change in the fuel's price index, 46 percent of that change would most likely be reflected in asphalt's price index two months later. In comparison with previous studies on this topic, revealed that the methods used to predict changes in asphalt price indexes in this thesis were significantly more effective ($p = 0.05$) than at least one of the previous studies.

Likewise, a cross-validation process conducted on the deterministic results revealed a 2.2 and 3.9 mean average percentage error (MAPE) in the estimation of future asphalt prices using observed crude oil and fuel price fluctuations. Although the time gap and percent change ratios between crude oil and fuel price indexes showed to be significantly more accurate at estimating future asphalt price indexes ($p = 0.05$), both MAPE values could be considered to reflect high accuracy levels when compared against the expected cost estimating errors determined by the America Association of State Highway and Transportation Officials (AASHTO). According to AASHTO, expected accuracy in final design estimates is expected to be between -5 percent and +10 percent. This study meets these requirements with MAPE values under +4 percent, in a commodity (asphalt) that sometimes represents over 80 percent of the total project construction cost.

Recognizing the unavoidable variability in the time gaps and percent change ratios mentioned above, this study conducted a stochastic analysis to create a risk assessment tool. This tool would allow estimators to make better-informed decisions by providing them with the probability of occurrence of different case scenarios in terms of potential time gaps and percent change ratios. For example, although the most likely results were 3 months and 58 percent between crude oil and asphalt price indexes, the stochastic analysis revealed that there was only a 14 percent probability of having a percent change ratio between 40 and 60 percent with a 3 month time gap.

Through updated processes and refined methods of cost estimating, the findings within this thesis are expected to improve cost estimating effectiveness in asphalt paving projects. In any project-oriented organization, such as state transportation agencies (STAs), an improvement in cost estimating effectiveness is reflected in a better allocation of available resources, which in the case of STAs, are continuously shrinking to address all transportation infrastructure needs. Although the results and risk assessment tools presented in this study are applied to assess the future asphalt price indexes in the state of Alabama, the methodology presented throughout this thesis could be replicated for other STAs with local price indexes. In conclusion, the study presented in this thesis has the potential to improve STAs' resource allocation practices through a better understanding of the factors triggering price index fluctuations in the most critical commodity used by STAs, asphalt.

ACKNOWLEDGEMENTS

The author would like to recognize the faculty that advised her through the master's degree program. The author's advisor professor, Dr. Jorge Rueda-Benavides, for providing the opportunity to pursue a master's degree and his mentorship. The author's committee, Dr. David Timm and Dr. Wesley Donald, for their time in participating and reviewing the thesis.

Additionally, the author would like to thank the Graduate Research Assistants within Dr. Jorge Rueda-Benavides' research team: Mr. Karthik Chowadary Pakalapati, Mr. Mohamed Khalafala, and Ms. Keren Xu. They have helped the author through long nights of studying, meeting deadlines, and have provided inspiration with innovative ideas.

Most notably the author would like to express gratitude to her parents, Mrs. Elvia Cardinal and Mr. Paul Cardinal. Through their examples of hard work, ambition, passion, and love none of this would have been possible. They provided the author with the means to achieve a dream they could not live for themselves.

Finally, the author's best friend Ms. Ligia Delgado, she encouraged the author on a daily basis to persevere through the master's degree program and to set an example for people who aspire to go through the same process.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AASHTO	American Association of State Highway and Transportation Officials
ALDOT	Alabama Department of Transportation
APE	Absolute Percentage Error
Brent	North Sea Brent
CUSUM	Cumulative Sum
EIA	U.S Energy Information Administration
HMA	Hot Mix Asphalt
LSD	Localized Standard Deviation
MAD	Mean Absolute Deviation
MAPE	Mean Absolute Percentage Error
NAPA	National Asphalt Pavement Association
OPEC	Organization of the Petroleum Exporting Countries
PMI	Project Management Institute
SD	Standard Deviation
STAs	State Agencies
U.S.	United States
WTI	West Texas Intermediate

CHAPTER ONE: INTRODUCTION

1.1 Motivation for Research

Asphalt is a major component of horizontal construction in the state of Alabama and across the nation in general (1). Since it came to the United States (U.S.) in 1870, it currently "covers more than 94 percent of the 2.7 million miles of paved roads in the U.S." (2). It has grown in popularity due to its "cost efficiency, reduction in noise pollution, comfort, and environmental sustainability" (3). The National Asphalt Pavement Association (NAPA) estimates public transportation agencies purchase more than 4 million tons of asphalt for every $1 billion spent on transportation infrastructure projects (2). To put these numbers in perspective, in 2014, federal, state, and local agencies invested more than $165 billion dollars on transportation construction and maintenance projects (2). Based on the national average unit cost of asphalt being approximately $75.00 per ton, the estimated total cost for asphalt paving production for one year was approximately $49.5 billion dollars (4).

The relevance of this thesis is better illustrated in the following quote from Xu (2018) in a previous study involving asphalt prices in the state of Alabama (5):

> *"The relevance of this paving material (asphalt) is not different in the state of Alabama, where this material is used in about 98 percent of the paved roads, […] and consum[es] over 40 percent of the Alabama Department of Transportation (ALDOT) annual construction budget. [Thus,] becoming an item that requires special attention during resource allocation and cost estimating procedures. Therefore, any improvements in ALDOT's HMA (hot mixed asphalt) cost*

estimating procedures is expected to be reflected in better budget control and a more effective use of ALDOT's limited available resources" (5) .

The relationship between asphalt and crude oil lays in the fact that asphalt is a byproduct of crude oil. "Almost all of the asphalt used today for paving comes from petroleum crude oil. Liquid asphalt is the heaviest part of crude oil—left over after all the volatile, light fractions are distilled off for products such as gasoline" (6). Crude oil is considered a "parent commodity" for asphalt and fuel, making these two "sibling commodities" (6). Although researchers and industry stakeholders consistently refer to a strong correlation between asphalt, crude oil, and fuel, there is a lack of formal research aimed to support those references. This thesis aims to analyze, quantify, and model the relationship between price fluctuations among these three commodities using ALDOT's price indexes for asphalt and fuel as well as national crude oil price indexes. As a result, state agencies (STAs) like ALDOT, will have the ability to develop robust cost estimating systems for asphalt paving projects; facilitating better resource allocation for taxpayers' money.

1.2 Research Objectives

The research methodology presented throughout this thesis was strategically designed to achieve the following three primary objectives:

1. Estimate the mean time gap between significant price index changes in crude oil and their reflection on asphalt price indexes, as well as the same time gap between asphalt and fuel price indexes in the state of Alabama.
2. Estimate the average percent change ratio between asphalt and crude oil price indexes, as well as between asphalt and fuel price indexes in the state of Alabama.
3. Conduct a stochastic analysis to model the variability in the time gaps and percent change ratios associated with the first two objectives. Then, synthesize the results of

this analysis into two risk assessment tools: one to assess asphalt price indexes based on fluctuations on crude oil price indexes and one to assess asphalt price indexes based on fuel price index fluctuations.

The following outline summarizes the work performed to meet the main research objectives and will be presented in detail in Chapter 3:

1. Collect over 12 years of ALDOT's monthly asphalt and fuel price indexes, as well as over 12 years of the monthly national crude oil price index. For each price index, there were 150 months collected, from January 2006 to June 2018.
2. Develop and apply a statistical analysis method to identify significant trend changes along the first 138 months for each of the three price indexes. An innovative method was developed to identify trend changes by modifying the cumulative sum (CUSUM) statistical analysis methodology presented by Boddy and Smith (2009) (7). The first 138 months were used to quantify the relationships between asphalt, crude oil, and fuel price indexes, while the last 12 years were used to validate the quantitative results.
3. Pair months with significant trend changes in the asphalt price index with their corresponding significant trend changes in the crude oil and fuel price indexes.
4. Estimate the average time gap among all pairs of significant trend changes. It should be noted that two mean values were estimated at this step: 1) a mean time gap between significant trend changes in the crude oil price index and their corresponding changes in the asphalt price index and 2) a mean time gap between significant trend changes in the fuel price index and their corresponding changes in the asphalt price index.
5. Estimate the mean percent change ratio among all pairs of significant trend changes. Two mean values have also been estimated at this step: 1) a mean percent change ratio

between significant trend changes in the crude oil price index and their corresponding changes in the asphalt price index and 2) a mean percent change ratio between significant trend changes in the fuel price index and their corresponding changes in the asphalt price index.

6. Use a cross-validation approach to compare future asphalt price indexes estimated versus observed asphalt price index fluctuations from crude oil price indexes against asphalt price indexes estimated based on fuel price index fluctuations.
7. Use a cross-validation approach to compare the results from Steps 4 and 5 against results obtained by two previous similar studies: one by Sullivan in 2014 and another by Attanasi in 2008 (8; 9). For the cross-validation process, the last 12 months of data from each price index from July 2017 to June 2018 were used.
8. Develop a risk analysis tool to account for the unavoidable variability in the time gaps and mean percent change ratios estimated at steps 4 and 5.

The objectives accomplished within this thesis were meant as a preliminary investigation for potential future objectives. Future objectives based on this research are to review this method against all STAs' price index data to eventually develop a user-friendly program that would provide a single percent change ratio and time gap for asphalt to use for contract negotiations. This user-friendly program would include an input of the price indexes used by the STA to obtain a single output value for an expected percent change and time for that change in asphalt's price index.

1.3 Organization of Thesis

To provide a comprehensive description of the research performed, this thesis was divided into five chapters. The five chapters are organized as follows:

1. Introduction
2. Literature Review
3. Methodology
4. Data Analysis
5. Conclusion and Recommendations

Chapter One: Introduction, provides a brief description of the motivation behind the thesis. This was meant to show the importance of asphalt in the horizontal construction industry, and describe the connection between asphalt, crude oil, and fuel. A brief overview of the objectives that would be accomplished through this thesis' research are also included.

Chapter Two: Literature Review, explains the importance of effective cost estimation of asphalt, and provides further details on the connections between asphalt, crude oil, and fuel. The descriptions of these connections includes important economic, manufacturing, and political factors contributing to the high volatility of asphalt prices. Other significant research is presented as a baseline for the assumptions and findings within this thesis.

Chapter Three: Methodology, discusses the process used to identify the critical significant changes related to asphalt, crude oil, and fuel price indexes. This is comprised of a discussion of the origins of the cumulative sum (CUSUM) statistical analysis method selected for this research and the reasoning behind using this method versus other options. It provides different real life applications of the statistical analysis method selected aside from the purposes of this thesis. Then, presents the modified version of the methodology developed to accommodate the needs of the data to identify and pair significant trend changes in the asphalt, crude oil, and fuel price indexes.

Chapter Four: Data Analysis, describes the actual application of the modified CUSUM statistical analysis method described in Chapter 3 to the data for this thesis. It provides a general discussion on the significant trend changes found in the three price indexes. A deterministic mean estimate for the time gap and percent change ratio of crude oil and fuel versus asphalt price indexes were answered within this chapter. This chapter also presents two different cross-validation processes.

The first cross-validation process was used to compare the accuracy of asphalt price indexes estimated from crude oil price index fluctuations against those estimated based on fuel price index fluctuations to predict observed asphalt price indexes. The second cross-validation process was used to compare the quantitative results from this study against findings from two previous similar studies. After verifying the validity of the obtained deterministic results, this chapter presents a stochastic analysis that allowed for the development of two different risk analysis tools. These two different tools represent the probabilistic relationship between asphalt price index fluctuations and previously observed fluctuations in crude oil and fuel price indexes. A discussion of the results is at the end of this chapter.

Chapter Five: Conclusions and Recommendations, provides a brief overview of the findings and main contributions of this thesis. It discusses the results presented in Chapter 4 and its potential implications on the horizontal construction industry. In conclusion, future potential research is discussed based on the findings within this thesis.

CHAPTER TWO: LITERATURE REVIEW

2.1 Introduction

Chapter 2 encapsulates the previous research and background for the relationship between asphalt, crude oil, and fuel. This comprehensive literature review begins with the significance of effective cost estimating measures in construction and their impacts on state agencies' (STAs') budgets. Maintaining effective cost estimating measures helps to mitigate the common consequences of ineffective cost estimating such as over budget and behind schedule projects (10).

After a review of cost estimation, this chapter will present a discussion of the relationship between asphalt, crude oil, and fuel. Stemming from their relationship will be a discussion of the impacts of political influences, production methods, and economic importance of crude oil on the construction industry (11). Although the end results within this thesis are in regards to asphalt, the greater investigation stems from crude oil. Crude oil has a significant worldwide impact since it is the source of many major components in the construction industry, not just asphalt (12).

Once there is a greater understanding of cost estimating and the relationships between all three commodities, similar existing research on this subject will be reviewed. From this review, a discussion of the similarities and differences between previous research efforts and this thesis' research will be provided. The previous research on this subject aids in initial ideas and validates various assumptions that will be studied in further detail in Chapter 3.

2.2 The Importance of Effective Cost Estimating for Horizontal Construction Projects

Cost estimating as defined by the Project Management Institute (PMI) is "the iterative process of developing an approximation of the monetary resources needed to complete project activities" (13). The items typically included in a cost estimate are as follows (13):

1. Labor
2. Materials
3. Equipment
4. Services
5. Software
6. Hardware
7. Facilities
8. Contingency Cost

According to both the PMI and the American Association of State Highway and Transportation Officials (AASHTO), there are four common cost estimating methods for transportation projects. These cost estimating methods are parametric, historical bid-based, cost-based, and risk-based estimating (13; 14). The definitions of the different estimating methods are as follows:

Parametric Estimates: **"**Use of a statistical relationship between historical data and other variables to calculate an estimate for activity parameters, such as scope, cost, budget, and duration" (13). This method is typically used during the early phases of construction when there is limited information of the scope of work on a project (15).

Historical Bid-Based Estimates: "Historical bid-based estimating uses data from recently let contracts as the basis for determining estimated unit prices for a future project" (14; 16). Historical bid-based estimates are the most common estimates used by STAs (16; 17).

Cost Based Estimating: "Cost-based estimating considers seven basic elements: time, equipment, labor, subcontractor, material, overhead, and profit. Generally, a work statement and set of drawings or specifications are used to 'take off' material quantities required for each discrete task necessary to accomplish the project bid items. From these quantities, direct labor, materials, and equipment costs are calculated based on assumed production rates. Contractor overhead and profit are then added to this direct cost. The total cost divided by the quantity gives the estimated unit price for the work item" (14; 16). This method of estimating is typically used when the scope of work for a given project is clearly defined (13).

Risk Based Estimating: "Probabilistic estimating that captures risk variability

associated with assumptions made and characterizes them through a range of estimated costs. Rather than using a single number to represent project cost, probabilistic approaches that involve simple or complex modeling based on the relationships among cost, schedule, and events related to the project are used to assess uncertainties and related risks and translate the risks into costs" (15). Deterministic cost estimates must have minimal variability in scope for them to be an effective estimate (15). Due to the dynamic environment of horizontal construction projects, it is hard to maintain minimal scope variability. Using risk-based estimates provides multiple case scenarios of an estimate while identifying items within the estimate that can have the largest impact on total project costs (15).

Although there are a plethora of estimating methods to determine the total cost for horizontal construction projects, it is difficult to ensure estimating accuracy and effectiveness. The two most common errors associated with cost estimation are projects being over budget or behind schedule. An international study (2002) found that 20 percent of construction projects worldwide were over budget (18). A report performed by the Quality Committee at AAHSTO (2007) reviewed nine of the largest STAs in the nation to determine their mean percentage of over budget and behind schedule projects. It was determined that 56 percent of their projects were over budget and 47 percent did not meet their original schedules (10). It also revealed that projects that cost over five million dollars to complete are more likely to be over budget and behind schedule (10). In conclusion, it was established that the mean percent of cost overruns for construction projects is approximately four percent (10). Although four percent appears to be a minimal amount, in comparison to billion dollar STA budgets the amount of money associated with four percent is significant.

The total amount of money spent on transportation construction and maintenance projects in 2014 was $165 billion dollars (2). To put the cost overrun of four percent into perspective, approximately $6.6 billion dollars were potentially lost due to ineffective cost estimating practices. According to NAPA, approximately $80 billion dollars were spent annually on asphalt for transportation construction and maintenance projects (2). That is approximately 50 percent of the money spent in 2014 for transportation construction and maintenance projects. Applying the same four percent error to asphalt, approximately $3.2 billion dollars were potentially lost due to inadequate cost estimating practices.

Another international study (2010) determined there are four main reasons behind insufficient cost estimating. They were listed as economic, political, psychological, and technical issues (11). From the four components mentioned, technical and political had the most detrimental impact on ineffective cost estimates (11). Both of these items maintained a relationship with each other. From the technical aspect, it revealed that forecasting escalations in costs for construction items such as asphalt is one of the common factors for over budget estimates (11). The technical aspect is tied to the political aspect, due to people being overly optimistic in their estimates to have politically attractive outcomes for the public (11). A discussion of the technical, economic, and political aspects of asphalt, crude oil, and fuel impacts on construction will be discussed in the following sections of this chapter.

2.3 How are Crude Oil and Fuel associated with Asphalt?

Bitumen is the binding material used in the production of asphalt and it is also a byproduct of refined heavy oil (19). Refined heavy oil forms approximately 38 percent of crude oil (20). From heavy oil, other products are created such as diesel fuel (20). The byproducts of crude oil are separated into two categories. These categories are light and heavy oil. They are placed into these categories based on density and viscosity of the oil when burned (21). A representation of all the items created from crude oil are shown in Figure 2.1 (22). Although it is assumed that the price and production of crude oil impacts the price of asphalt in a major way, it is important to note that the percent of crude oil byproduct to create asphalt is relatively small. Standard asphalt used in the industry typically has a range of 4.5 to 6.5 percent bitumen in each asphalt binder (23).

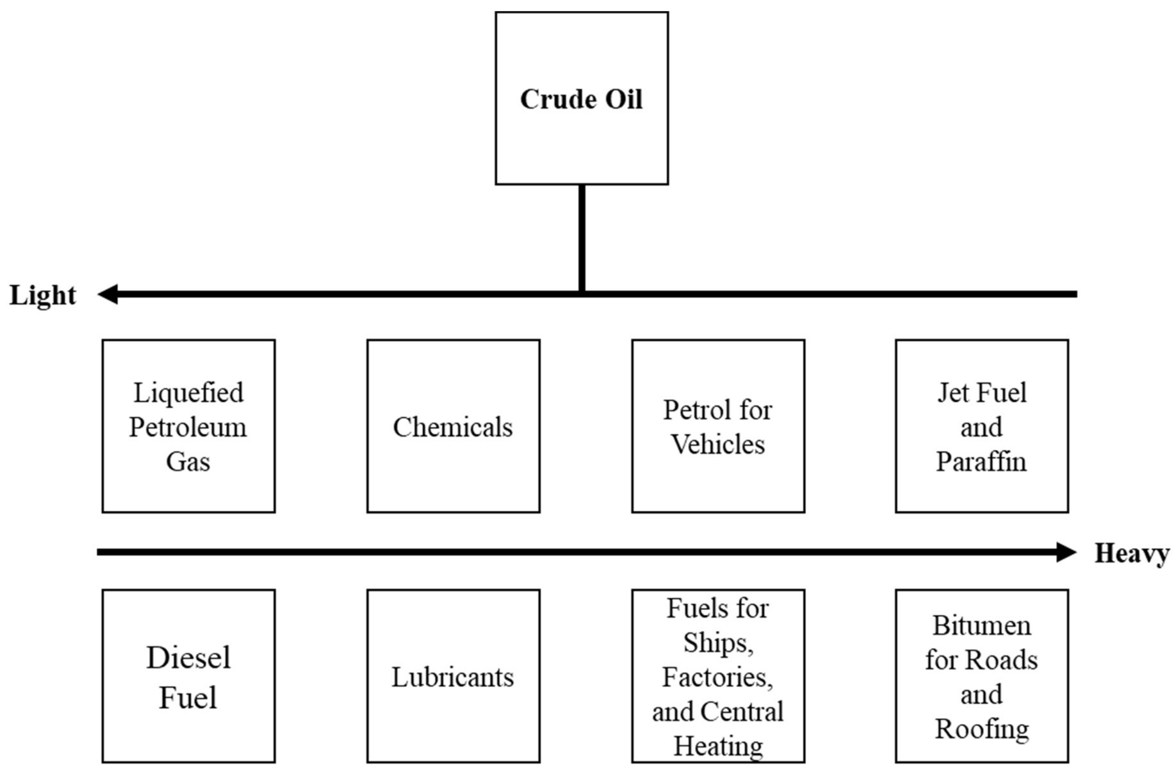

Figure 2.1 Byproducts of Crude Oil (22)

When discussing fuel for the entirety of this thesis, it is referencing diesel fuel. Fuel is often considered a relevant item in horizontal construction projects because it powers more than three-fourths of all horizontal construction equipment (24). Fuel is usually recognized as a major cost estimating challenge in the horizontal construction sector along with asphalt (24). Therefore, both asphalt and fuel costs are important to consider when creating an effective cost estimate because of their substantial impact on a successful horizontal construction project. For this reason, the majority of STAs keep track of the fluctuating prices of asphalt, crude oil, and fuel for estimating purposes.

2.4 The Economics of Crude Oil on the Asphalt Industry

In general, crude oil prices have a high impact on the construction industry (25). As shown previously in Figure 2.1, crude oil is not just a central ingredient of asphalt, but the function of machines, transportation of materials, and the creation of products used in all variations of construction projects. A study performed by Wilmot and Cheng (2003) corroborated that the leading cause of cost changes in horizontal construction is directly correlated to items that are crude oil based (26). This study was seconded by Damnjanovic and Zhou (2009) with the conclusion that "expected change[s] in oil prices impact unit bid price[s] at a level of 99.9% significance" in a horizontal construction cost estimation (12).

These types of research and studies support the theory that there is a relationship between asphalt prices and crude oil prices, but this is a blanket conclusion since the construction industry, in general, is impacted by crude oil. The strong relationship with asphalt seems obvious because it is a byproduct of crude oil, but so is diesel fuel to 18 wheelers that carry construction materials for projects. Although it provides a basis for research, to conclude that only asphalt is strongly influenced by crude oil, is misleading as other areas in the construction industry are also heavily inflicted by crude oil.

2.5 Political Influences and Volatility of the Crude Oil Industry on Asphalt Prices

The Organization of Petroleum Exporting Countries (OPEC) controls 72 percent of all the reserves in the world. Since most OPEC nations are located in what is known as unpredictable political regions, oil prices tend to be very volatile due to the authoritarian decisions made that have a detrimental impact on the oil industry worldwide (27). One of the recent decisions made in 2016, was the production cut in OPEC of medium and heavy oils (21). The cut in production of these types of oils has had a direct impact on the ability to produce asphalt since heavy oil is the type of oil needed to produce asphalt.

For the United States (U.S.), the ability to control the volatility of crude oil itself is beyond its control since they do not have the majority of oil resources. Although the U.S. cannot control the volatile cost fluctuations in crude oil, they can control the way it is produced to maximize its profitability. Unfortunately, in recent years heavy oil production has not been the most profitable item for refineries.

2.6 The Impacts of Crude Oil Production Methods on the Asphalt Industry

The variation of crude oil that has become more profitable is the lighter variation which is not the variety used to create asphalt (28). From the years 1982 to 2011, the refineries producing asphalt have been reduced from 301 to 148 (20). "New processing technologies are affecting the amount of asphalt being produced at refineries...due to the growing demand for gasoline and diesel" (28). Cokers, which are the new addition to many refineries, aid in the process of taking residual oil typically left to create asphalt "to yield more of the lighter" profitable fuels (28). Refineries are producing less asphalt based on the increased profitability and capabilities of cokers.

This asserts that not only crude oil price fluctuations control asphalt, but also the way crude oil is refined increases asphalt's price due its scarce production. These are all items to take into consideration when providing an estimate for associated asphalt costs in projects.

2.7 Asphalt and Crude Oil Relationship Studies

Background research on this topic found resources focused on various forms of production of asphalt rather than cost estimating methods. Ultimately, the majority of the sources were concerned with the more subjective picture of the relationship between asphalt and crude oil versus asphalt's competitor, concrete. The Portland Cement Association (PCA) (2014) posted a series of studies defining the relationship between asphalt and crude oil as a means to promote concrete (8).

Due to the nature of the PCA's business, these studies should be considered as potentially biased for research and academic purposes, meaning that they could be intended to promote concrete by discouraging the use of asphalt. This study intends to take a nonbiased approach to help develop more accurate cost estimating measures for STAs to be successful in their projects. This thesis does not intend to support one type of paving material over the other based on defining the relationship between asphalt and crude oil.

In one of PCA's studies, it was concluded that "roughly 60 percent of long-term asphalt price increases are accounted for by oil price changes" (8). On a national scale, it found a "six month" lag from when a change in price in crude oil occurs to when it inflicts a change in price for asphalt (8). They suggest that for every 1 percent crude oil prices increased, it caused a 70 percent increase in asphalt prices (8).

Another study conducted by Attanasi (2008), confirmed that the volatility of asphalt prices is greater than that of crude oil prices due to having no control over the access to major oil reservoirs and maximizing the production of oil refining (9). The refining industry moving towards producing more light oils than asphalt has caused a shift in the volatility of asphalt prices. Basing its data from Canadian resources, it showed that the fluctuations in crude oil prices are expected to be 130 percent greater in asphalt prices within the next month (9). Attanasi argues that the shift towards refining lighter oils over asphalt comes from asphalt's increased volatility in comparison to lighter oils when it comes to overall crude oil price fluctuations. It is less of a risk for refineries to work with lighter oils than it is producing the key component of asphalt, bitumen (9).

This study intends to review all of these factors to develop a method to estimate future asphalt prices. In predicting the volatility of asphalt from crude oil, it may help to increase the profitability of it in the refining industry. If asphalt can become less volatile by being able to understand its trends in comparison to other variations of crude oil refining, the market for it could be more accepting of producing it.

CHAPTER 3: METHODOLOGY

3.1 Introduction

The purpose of Chapter 3 is to provide a description of the modified cumulative sum (CUSUM) statistical analysis methodology used to identify the significant trend changes (hereinafter referred to as critical points) in asphalt, crude oil, and fuel price indexes. Although this thesis focuses on a trend analysis of historical price indexes for these three commodities, the proposed methodology can be used to find the significant trend changes in any type of index or time series. The critical points for asphalt, crude oil, and fuel from January 2006 to June 2017 are summarized in the final section of this chapter, (*3.5 Critical Points of Price Indexes*) for further evaluation in Chapter 4. A summary of the process being described in this chapter is shown in Figure 3.1.

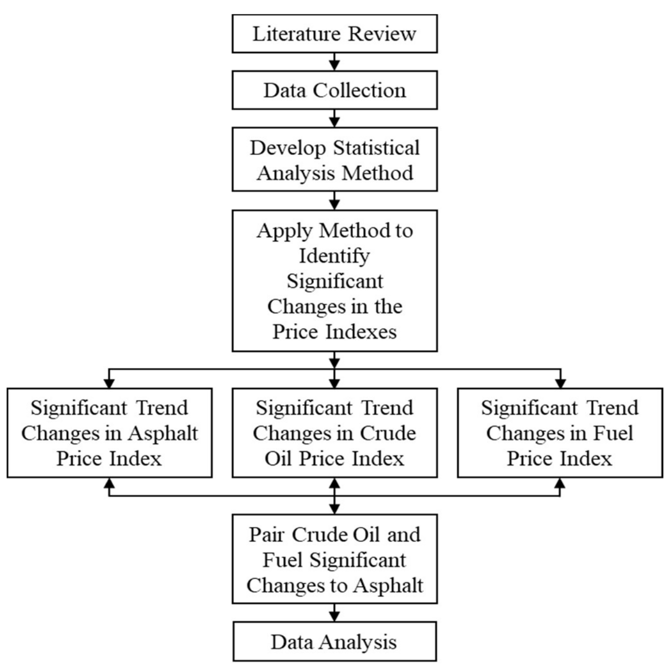

Figure 3.1 Methodology

3.2 Data Collection

This thesis used three different types of price indexes: asphalt, crude oil, and fuel. Monthly asphalt and fuel price indexes from January 2006 to June 2018 were collected from the Alabama Department of Transportation (ALDOT) website (29; 30). ALDOT price indexes were used since this was the location of where the research was being performed. The unit of measurement used in these price indexes was a dollar per gallon of asphalt or fuel. The asphalt price index is based on a standard pen grade (PG) asphalt specification (31).

There are approximately 161 different internationally traded crude oils (32). Of these 161, only a select few are used as a benchmark for predicting price changes or trends in crude oil. These are Dubai, North Sea Brent (Brent), OPEC Basket Price, and West Texas Intermediate (WTI) (33; 34). The major benchmark used for price of crude oil in the United States (U.S.) is WTI (34), which was the type of price index used in this study. Historical WTI crude oil price indexes were collected from the U.S. Energy Information Administration (EIA) website (35). Similar to ALDOT, EIA maintains a monthly record of the changes in crude oil price indexes. The price indexes collected for crude oil for the purposes of this thesis were from January 2006 to June 2018. The unit of measurement used in these price indexes was a dollar per barrel. The crude oil price indexes were converted to a dollar per gallon to match the unit of measurement for the other price indexes. A conversion factor of 42 was provided by EIA to convert dollars per barrel to dollars per gallon.

It should be noted that only the price index values for the first 138 months (January 2006 to June 2017) were used to determine the time gaps, percent change ratios, and the stochastic tables. The remaining 12 months were used for validation purposes. For calculation and visual representation purposes, all months between January 2006 and June 2017 are denoted as 1 to 138.

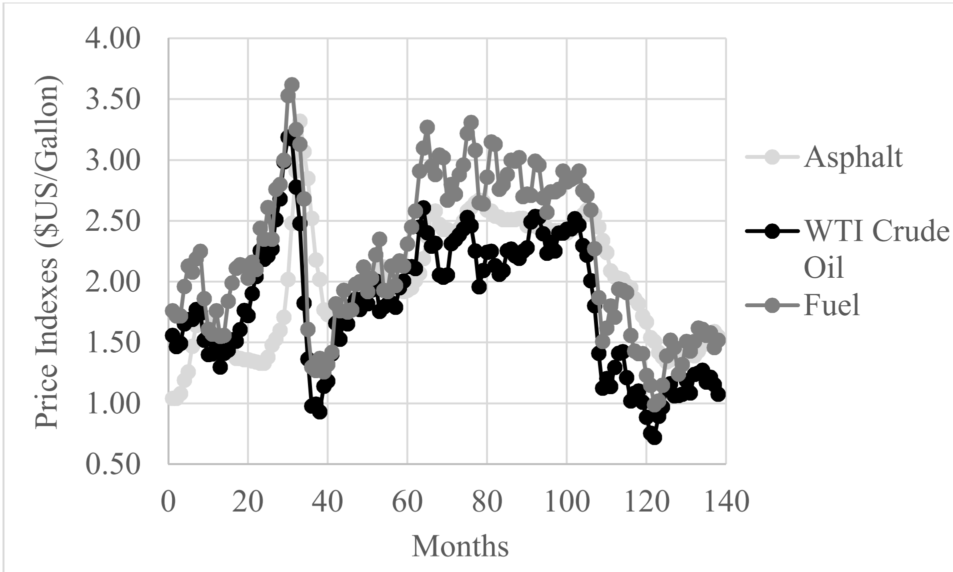

Figure 3.2 Price indexes

The hypothesis for this research stemmed from the above graphical representation. A visual inspection shows asphalt, crude oil, and fuel following similar trends. From the visual inspection it is inferred that a given price fluctuation could occur first in the crude oil price index, then in the fuel price index, and finally in the asphalt price index. Although Figure 3.2 provides some initial inferences to crude oil's and fuel's relationship with asphalt, it is not a confirmation of their connection as three independent items working together. Previous research has concluded that various statistical analysis methods could be used to connect these items (8; 9; 36) To formally determine their connection with one another, a modified version of the CUSUM statistical analysis method as described by Boddy and Smith (2009) was developed (7).

3.3 Cumulative Sum Statistical Analysis Method

The process for alignment of these three items as codependent on each other was the cumulative sum (CUSUM) statistical analysis method (7). This statistical analysis method is used when it is necessary to identify "if [a] change has taken place" and "approximately when the change took place" (7). CUSUM is a preferred method for identifying significant changes in data due to its ability to detect the slight shifts in trends that are undetectable using other methods (37). The main reason for choosing this method for the purposes of this thesis was its effectual ability to determine when a significant change has occurred.

Examples of using the CUSUM method for identifying significant changes are prevalent, but not limited to the following industry examples:

1. Evaluating surgical procedures, performance, and learning curves of surgeons (38; 39).
2. Evaluating carbon emissions and energy consumption patterns (40; 41).
3. Evaluating quality in industrial processes (7; 42)

Initial development by E.S Page in 1954, stemmed multiple variations of the CUSUM method depending on the necessary conclusions and data required for the research (43). Examples of the different CUSUM methods are as follows (44):

1. One-sided decision scheme
2. Two-sided decision scheme
3. Average run length
4. Standardized method
5. Monitoring process variability
6. Combined Shewhart-CUSUM

The method used for this thesis was a modified version of the one-sided decision scheme at a 99 percent confidence level ($p = 0.01$) based on significant test values developed by Goldsmith and Woodward (1964) (37; 45). The main objective of the CUSUM method is that it determines significant changes based on "successive deviations from a target mean for a process" (37) . CUSUM does this by smoothing the intense series of peaks seen in Figure 3.2 into important groups of significant changes in trends. It creates a series of positive to negative slopes representing the variations in trends for a process.

The CUSUM method described by Boddy and Smith (2009) is a statistical hypothesis testing technique that could be performed through 11 steps as shown in Figure 3.3 (7).

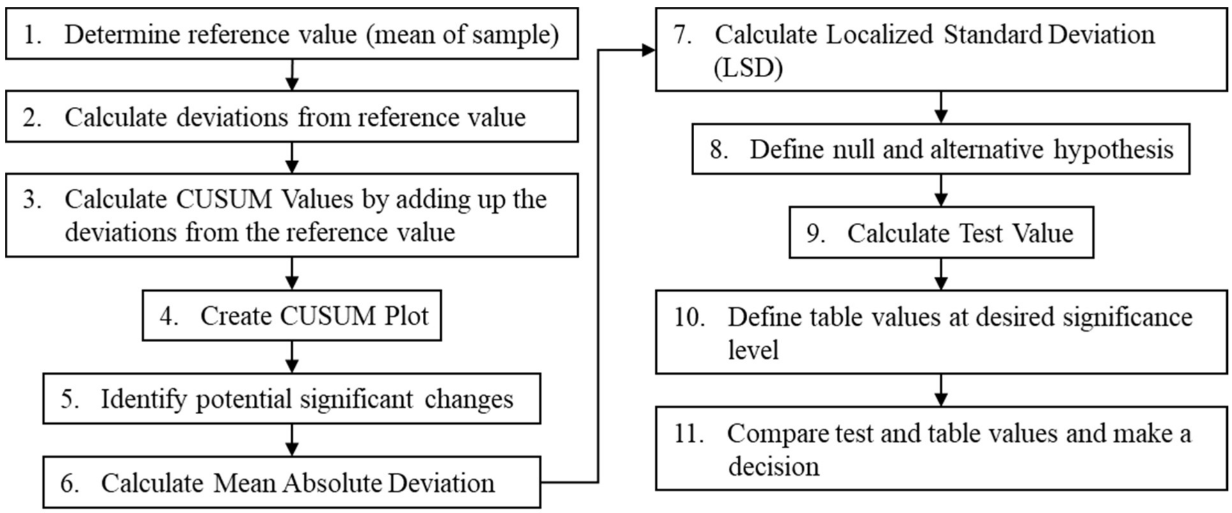

Figure 3.3 CUSUM Statistical Analysis Method

The example below is intended to better illustrate the step-by-step process shown in Figure 3.3. This example is not related to the topic of this study since the CUSUM statistical analysis method as shown in Figure 3.3 is not suitable for the purposes of this thesis (for reasons that will be explained later in this section), but it provides a clear idea about the basics of this statistical testing approach.

Example: Table 3.1 presents the daily production in linear feet (LF) of a construction crew installing water distribution piping during a period of 50 days. This time series data has been illustrated in Figure 3.4. In this example, the CUSUM method was used to determine if there had been any significant changes in the productivity (output) of this crew with respect to its mean productivity (reference value = 250 LF/day). The following was the step-by-step process followed to apply the CUSUM statistical analysis method as presented in Figure 3.3.

Table 3.1 Example – Water Distribution Piping Installation Daily Output

Day	Output (LF/day)	Day	Output (LF/day)	Day	Output (LF/day)	Day	Output (LF/day)	Day	Output (LF/day)
1	253	11	253	21	259	31	250	41	256
2	246	12	257	22	250	32	255	42	238
3	259	13	251	23	241	33	240	43	244
4	254	14	253	24	255	34	248	44	249
5	248	15	244	25	252	35	244	45	258
6	261	16	260	26	263	36	247	46	243
7	251	17	250	27	252	37	252	47	248
8	240	18	255	28	248	38	243	48	246
9	256	19	245	29	261	39	239	49	239
10	249	20	248	30	257	40	242	50	248
Mean Daily Output = 250 LF/day									

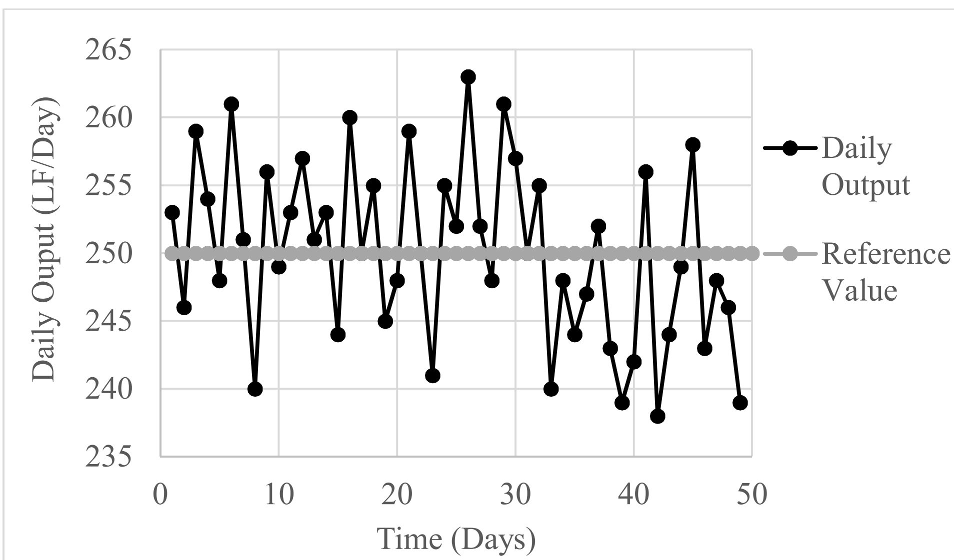

Figure 3.4 Example – Water Distribution Piping Installation Daily Output

Step 1: Determine Reference Value (Mean of the Sample)

For the purposes of this example, the sample corresponds to the piping installation daily outputs recorded for each of the 50 days. As shown in Table 3.1, the mean of all recorded daily outputs was 250 LF/day.

Step 2: Calculate Deviations from Reference Value

Deviations in this step were calculated using Equation 3.1. This equation subtracts the reference value from each sample value. The deviation of each observation is shown in Table 3.2.

Equation 3.1 $Deviation_i = X_i - Reference\,Value$

$Deviation\ _i$: Deviation for observation i

X_i: Value for observation i

Table 3.2 Example – Deviations (Dev) from Reference Value

Day	Dev	Day	Dev	Day	Dev	Day	Dev	Day	Dev
1	3	11	3	21	9	31	0	41	6
2	-4	12	7	22	0	32	5	42	-12
3	9	13	1	23	-9	33	-10	43	-6
4	4	14	3	24	5	34	-2	44	-1
5	-2	15	-6	25	2	35	-6	45	8
6	11	16	10	26	13	36	-3	46	-7
7	1	17	0	27	2	37	2	47	-2
8	-10	18	5	28	-2	38	-7	48	-4
9	6	19	-5	29	11	39	-11	49	-11
10	-1	20	-2	30	7	40	-8	50	-2

Step 3: Calculate CUSUM Values

The CUSUM values for each observation were calculated as shown in Equation 3.2. This was the cumulative sum of the deviations for all the observation previous to the current observation. The CUSUM values for all the observations are illustrated in Table 3.3.

Equation 3.2 $CUSUM_i = CUSUM_{i-1} + Deviation_i$

$CUSUM_i$: Cumulative sum value for observation i

$CUSUM_{i-1}$: Cumulative sum value for observation i-1 (previous observation)

$Deviation_i$: Deviation for observation i

Table 3.3 Example – CUSUM Values

Day	CUSUM	Day	CUSUM	Day	CUSUM	Day	CUSUM	Day	CUSUM
1	3	11	20	21	42	31	71	41	37
2	-1	12	27	22	42	32	76	42	25
3	8	13	28	23	33	33	66	43	19
4	12	14	31	24	38	34	64	44	18
5	10	15	25	25	40	35	58	45	26
6	21	16	35	26	53	36	55	46	19
7	22	17	35	27	55	37	57	47	17
8	12	18	40	28	53	38	50	48	13
9	18	19	35	29	64	39	39	49	2
10	17	20	33	30	71	40	31	50	0

Step 4: Create CUSUM Plot

Figure 3.5 corresponds to the CUSUM plot for the values shown in Table 3.3. The CUSUM plot in Figure 3.5, allows for a smoother curve than the one shown in Figure 3.4 with the original values.

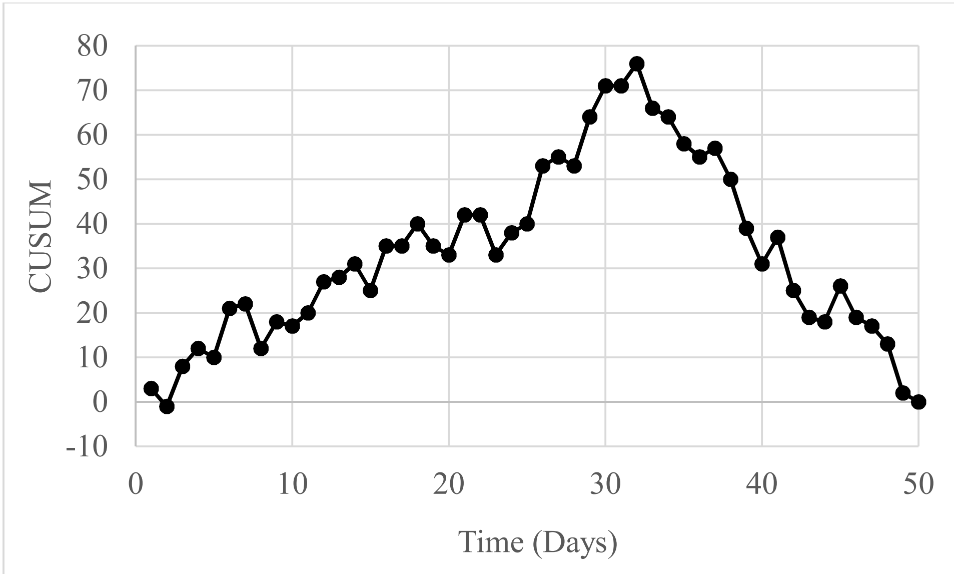

Figure 3.5 Example – CUSUM Plot

Step 5: Identify Potential Significant Changes

The smoothness of the CUSUM plot facilitated the visual identification of potential significant trend changes within the time series. For instance in Figure 3.5, it was easy to recognize a potential significant trend change at day 32 (the peak). Whether or not this change was statistically significant, will be determined in the following steps.

Step 6: Calculate Mean Absolute Difference

The mean absolute difference (MAD) was calculated as shown in Equation 3.3. This value was the average of the absolute differences between adjacent observations. The difference for each pair of adjacent observations is shown in Table 3.4. According to Equation 3.3, the MAD value for this example was equal to 8.14 LF/day. The MAD value was required to calculate the localized standard deviation (LSD) in the following step.

Equation 3.3 $MAD = \frac{\sum|X_{i-1}+X_{i-1}|}{n-1}$

MAD: Mean absolute difference

X_i: Value for observation i

X_{i-1}: Value for observation $i - 1$ (previous observation)

n: Number of observations

Table 3.4 Example – Absolute Differences (A.Diff)

Days	A. Diff	Days	A. Diff	Days	A. Diff	Days	A. Diff	Days	A. Diff
1-2	7	11-12	4	21-22	9	31-32	5	41-42	18
2-3	13	12-13	6	22-23	9	32-33	15	42-43	6
3-4	5	13-14	2	23-24	14	33-34	8	43-44	5
4-5	6	14-15	9	24-25	3	34-35	4	44-45	9
5-6	13	15-16	16	25-26	11	35-36	3	45-46	15
6-7	10	16-17	10	27-28	11	36-37	5	46-47	5
7-8	11	17-18	5	27-28	4	37-38	9	47-48	2
8-9	16	18-19	10	28-29	13	38-39	4	48-49	7
9-10	7	19-20	3	29-30	4	39-40	3	49-50	9
10-11	4	20-21	11	30-31	7	40-41	14	-	-

Step 7: Calculate Localized Standard Deviation (LSD)

Using the MAD at Step 6, the next step was to calculate the LSD as shown in Equation 3.4. This equation and value of 1.128 was a standard value that any calculated MAD was divided by (7). The LSD for this example was equal to 7.22 LF/day.

Equation 3.4 $LSD = \frac{MAD}{1.128}$

LSD: Localized Standard Deviation

MAD: Mean Absolute Difference

Step 8: Define Null and Alternative Hypotheses

For this statistical hypothesis test, the null (H_o) and alternative (H_a) hypotheses were defined as follows:

H_o = Test value is equal to or less than significant test value; there is no significant change in trend.

H_a = Test value is greater than significant test value; there is a significant change in trend.

These hypotheses referred to whether or not there was a significant trend change in productivity of the crew during the 50 days considered in the example

Step 9: Calculate Test Value

Following the guidelines provided by Boddy and Smith (2009), the test value for the CUSUM statistical analysis was calculated as shown in Equation 3.5 (7). Using the LSD calculated at Step 7 (7.22 LF/day) and the maximum absolute CUSUM value from Table 3.3 (76 LF/day), the test value for this example was equal to 10.50.

Equation 3.5 $Test\ Value = \frac{MAX|CUSUM|}{LSD}$

LSD: Localized Standard Deviation

Step 10: Define Significant Test Value at Desired Significance Level

For this example, as well as for this thesis, a 99 percent confidence level (p=0.01) was used to define the significant test value to be compared against the test value calculated at Step 9. The significant test values in this study were obtained from the Goldsmith and Woodward significant test values (1964) (7; 45). The test values were determined by the given sample size (number of observations) and the desired significance level. The table of significant test values is shown in Appendix A. According to Appendix A, the significant test value for the 50 observations used in this study at a 99 percent confidence level (p = 0.01) was equal to 10.40.

Step 11: Compare Test Values versus Significant Test Values to Make a Decision

Finally, the null and alternative hypotheses were evaluated by comparing the significant test and calculated test values at Steps 9 and 10. Given the test value was greater than the significant test value (10.50 > 10.40), the null hypothesis could be rejected at the 99 percent confidence level (p = 0.01). Therefore, it could be concluded that there was a significant change in the productivity rate of the crew within the 50 days considered in this study. Following the guidelines provided by Boddy and Smith (2009) (7), it could be assumed that the significant trend change occurred at day 32.

As previously mentioned, the reason for using an example not related to the topic of this thesis was to illustrate the CUSUM analysis process, and highlight in the next section how it is not suitable for the purposes for this study. The following two reasons led to the development of the modified version of the CUSUM analysis process presented above:

1. The visual identification of potential critical points was a subjective process that could lead to omitting significant trend changes that might not be easily identified through visual inspection. Likewise, a visual inspection prevents the automation of the CUSUM analysis process to facilitate the assessment of the historical price indexes for all three commodities across the 138 months considered in this study.
2. The reference value illustrated in Figure 3.4, which corresponds to the mean of all the 50 observations considered in the example, mediated where the significant trend changes would be determined. This means that the CUSUM approach illustrated in the example could only identify critical points following a horizontal trend. However, the three price indexes used in this study present upward and downward trends that may continue for a number of months. Using the CUSUM method as described by Boddy and Smith (2009) could pinpoint false critical points within relatively constant trends (7).

3.4 Modified Cumulative Sum Statistical Analysis Method

To combat the issues described in the previous section, a modified version of the CUSUM method by Boddy and Smith (2009) was created (7). To address the first issue stated in the previous section, an algorithm was developed that allowed for the assessment of all possible timeframes contained within the 138 considered in this study. For instance, the example in the previous section evaluates a 50-day timeframe under the assumption that a significant trend change may have had occurred at day 32. A rejection of the null hypothesis would automatically mean that a trend change had occurred at day 32.

The original method does not provide a means to evaluate every possible significant trend range within that 50 day timeframe. There was not a method to use to knowing whether a significant change occurred within a range before or after day 32. This issue could be minimal with 50 observations, but becomes more prevalent as the observations grows to a number like 138 for this thesis. A visual inspection would not be sufficient for 138 price index values since it overcomplicates the quantitative analysis in this study and prohibits the automation process to identify critical points.

A total of 9,453 different timeframes are comprised within a 138-month time series (i.e. 1-2; 1-3; 1-4; 1-5; […] 135-138; 136-137; 136-138; 137-138). Excluding all one- and two-month timeframes, for which a CUSUM would be unnecessary due to the lack of sufficient observations to define a trend within those short timeframes. A total of 9,180 timeframes were assessed for each price index. This means that the proposed algorithm performed 27,540 modified CUSUM analyses (9,180 x 3 = 27,540).

Besides evaluating all possible timeframes, the modified CUSUM method automates the process to identify the critical points within a timeframe for which the null hypothesis was rejected.

In other words, instead of relying on a visual inspection to determine that a significant trend occurred, the algorithm would automatically detect this day as the most likely critical point within the available timeframe.

To make the CUSUM method applicable to upward and downward trends, the proposed algorithm begins by transforming the trend in the timeframe under consideration into a horizontal trend. This was performed by placing the first and last observations in the timeframe on the horizontal axis with values in the vertical axis equaling to zero. The process of transforming non-horizontal trends into horizontal trends is better explained with the example shown in Figure 3.6. This figure illustrates the actual transformation process conducted for one of the 9,180 timeframes contained in the asphalt price index. This timeframe covers a 13-month period of time, from month 54 to month 66 (June 2010 to June 2011).

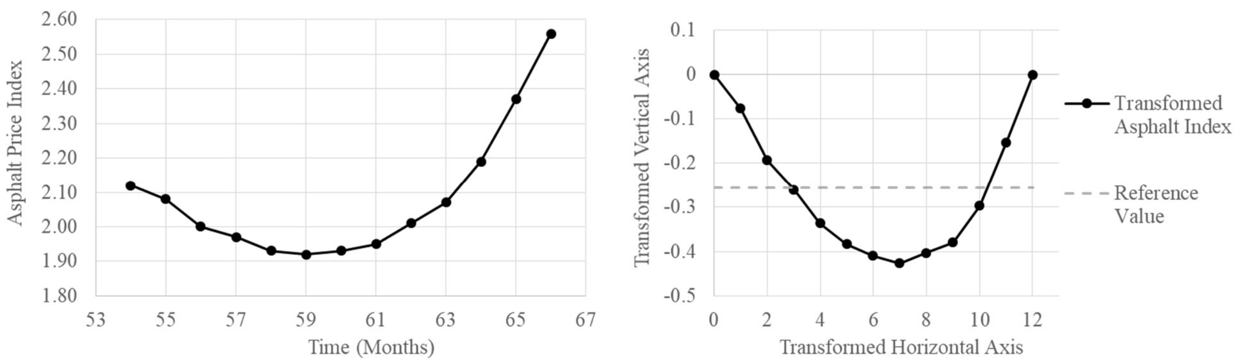

Figure 3.6 Horizontal Trend Transformation

Figure 3.6A shows the original month and asphalt price index values for the selected timeframe. Figure 3.6B shows the same observations, but in a transformed coordinate system intended to rotate the overall trend defined by the price index values into a horizontal position. The transformed coordinates for the timeframe under consideration are determined through the following three steps. It should be noted that these steps were applied to all three price indexes.

Step 1: Define Transformed Coordinates for the First Observation in the Timeframe.

The first observation in the timeframe was always at the origin of the coordinate system. Thus, the transformed coordinates for this observation were always zero in both the horizontal and vertical axes (x=0; y=0).

Step 2: Define Transformed Coordinates for the Last Observation in the Timeframe.

The last observation was always on the horizontal axis with a zero value in the vertical axis (y=0). The coordinate on the horizontal axis was equal to the straight distance calculated with Equation 3.6.

Equation 3.6 $$y_{trans;last} = \sqrt{\left(M_{last} - M_{first}\right)^2 + \left(I_{last} - I_{first}\right)^2}$$

$y_{trans;last}$: Transformed vertical coordinate for last observation in the timeframe

M_{last}: Month number for last observation in the timeframe

M_{first}: Month number for first observation in the timeframe

I_{last}: Price index value for last observation in the timeframe

I_{first}: Price index value for first observation in the timeframe

Step 3: Define Transformed Coordinates for the Remaining Observation in the Timeframe.

This step refers to the observations in-between the first and last observations. The transformed coordinates for these observations were defined using the vector dot product equation and trigonometric functions. Figure 3.6 was used to explain this process. First, the dot product equation solved for the angle between vectors $\overrightarrow{AB}$ and $\overrightarrow{AD}$ to calculate the angle theta (θ) shown in Figure 3.6 (see Equations 3.7 and 3.8). After calculating the angle theta, the trigonometric functions in Equations 3.9 and 3.10 were used to calculate the transformed coordinates for the horizontal and vertical axes. The process described at this step was repeated for all observation in-between the first and last observation for the timeframe under consideration.

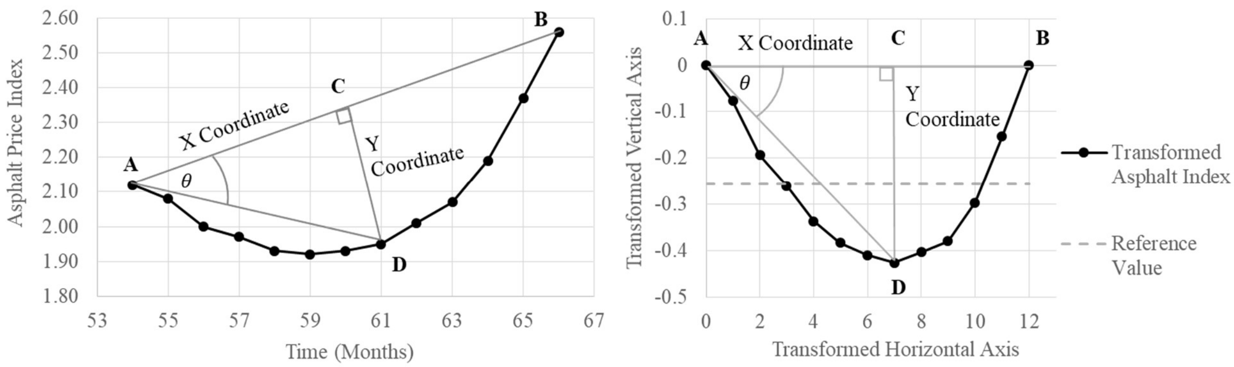

Figure 3.7 Horizontal Trend Calculation

Equation 3.7 $\theta_i = \cos^{-1}\left(\frac{\overrightarrow{AB}\cdot\overrightarrow{AD_\iota}}{\left|\overrightarrow{AB}\right|\cdot\left|\overrightarrow{AD_\iota}\right|}\right)$

Equation 3.8 $\overrightarrow{AB}\cdot\overrightarrow{AD_\iota} = \left(\left(M_{last} - M_{first}\right) \times \left(M_i - M_{first}\right)\right) + \left(\left(I_{last} - I_{first}\right) \times \left(I_i - I_{first}\right)\right)$

Equation 3.9 $x_{trans;i} = \left|\overrightarrow{AD_\iota}\right| \times \cos\theta_i$

Equation 3.10 $y_{trans;i} = \left|\overrightarrow{AD_\iota}\right| \times \sin\theta_i$

θ_i: Angle between vectors $\overrightarrow{AB}$ and $\overrightarrow{AD_\iota}$

$\overrightarrow{AD_\iota}$: Vector from first observation to observation i

$\left|\overrightarrow{AD_\iota}\right|$: Magnitude of vetor $\overrightarrow{AD_\iota}$

M_{last}: Month number for last observation in the timeframe

M_{first}: Month number for first observation in the timeframe

M_i: Month number for observation i

I_{last}: Price index value for last observation in the timeframe

I_{first}: Price index value for first observation in the timeframe

I_i: Price index value for observation i

$x_{trans;last}$: Transformed horizontal coordinate for last observation i

$x_{trans;last}$: Transformed vertical coordinate for last observation i

i: Observation under consideration

Table 3.5 shows the month and asphalt price index values for each of the observations in the timeframe shown in Figure 3.7, as well as their respective transformed coordinates.

Table 3.5 Transformed Coordinates

Original Coordinates		**Transformed Coordinates**	
Month	**Asphalt Price Index**	**Transformed x**	**Transformed y**
54	2.12	0.00	0.00
55	2.08	1.00	-0.08
56	2.00	1.99	-0.19
57	1.97	2.99	-0.26
58	1.93	3.99	-0.34
59	1.92	4.99	-0.38
60	1.93	5.99	-0.41
61	1.95	6.99	-0.43
62	2.01	7.99	-0.40
63	2.07	8.99	-0.38
64	2.19	10.00	-0.30
65	2.37	11.00	-0.15
66	2.56	12.01	0.00

Once the transformed price index for a given timeframe was created, the CUSUM process illustrated in Figure 3.3 was performed on the transformed coordinate system. As mentioned at the beginning of this section, the modified CUSUM analysis proposed in this study was designed to identify the most likely critical point within a timeframe when the null hypothesis has been rejected. Therefore, the visual identification of a potential critical point at Step 5 in the original process was not required. In the proposed approach, the critical point was identified after rejecting the null hypothesis and corresponds to the observation with the largest absolute transformed coordinate on the vertical axis.

The calculated test and significant test values were equal to 7.9 and 4.4 for the timeframe shown in Figure 3.7. Since the calculated test value was greater than the significant test value, the null hypothesis was rejected at a 99 percent confidence level (p = 0.01). Rejecting the null hypothesis meant that there was a critical change point within this timeframe. According to the proposed method, the most likely critical point would be the one that corresponds to month 61 in Table 3.5. Month 61 has the largest absolute transformed vertical coordinate of 0.43. A visual inspection would validate this conclusion since month 61 was when the slope changed from a downward trend into an upward trend as shown in Figure 3.7.

During the development of the proposed modified CUSUM methodology, there was another issue associated with the determination of the significant test values. As shown in Appendix A, the significant test values were for a select few sample sizes, which did not include majority of the sample sizes used in this study. The modified CUSUM analysis performed in this study included sample sizes that ranged from 3 to 136 observations. To determine accurate significant test values for all the sample sizes not listed in Appendix A, a polynomial regression model represented by Equation 3.11 and Figure 3.8 was used to interpolate significant test values.

Equation 3.11 $y = 1E - 05x^3 - 0.0024x^2 + 0.25x + 2.39$

x: Sample Size

y: Table Value

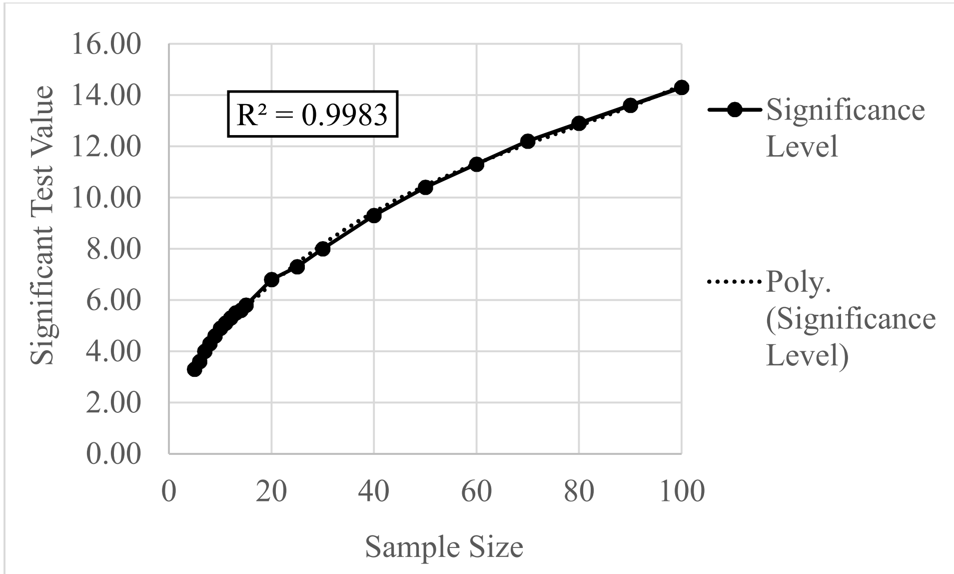

Figure 3.8 Significant Test Values

Based on the process described in this section, the modified CUSUM statistical analysis approach proposed in this thesis could be summarized into the 12 steps outlined in Figure 3.9. In comparison to the original method designed by Boddy and Smith (2009), there was one additional step added from the one shown in Figure 3.3 (7).

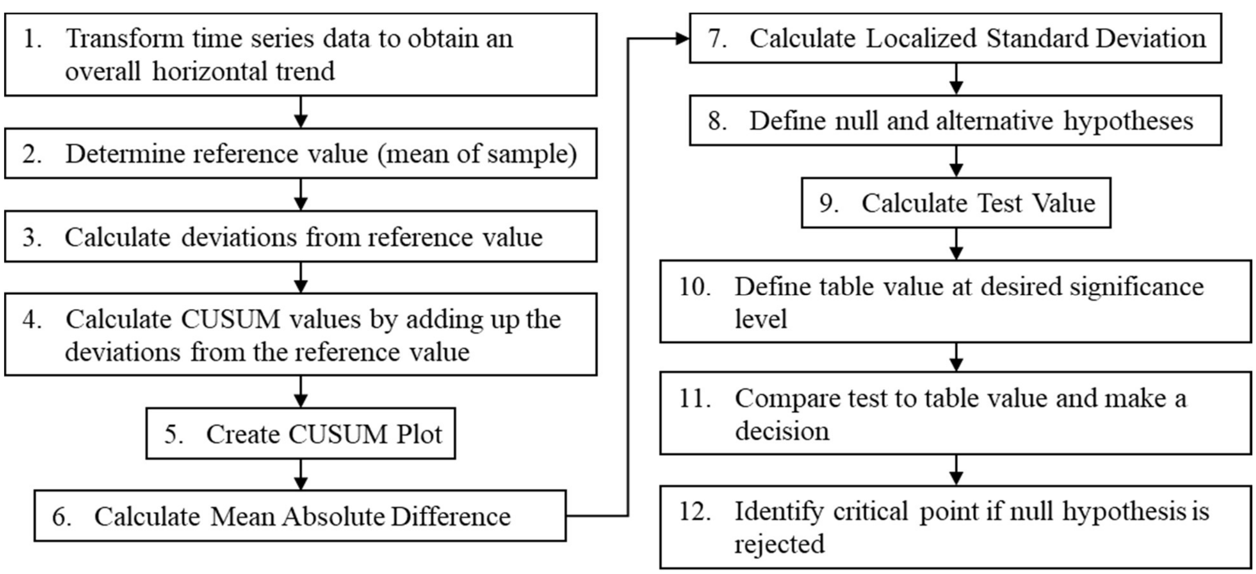

Figure 3.9 Modified CUSUM Statistical Analysis Method

3.5 Critical Points of Price Indexes

Several critical points were found for each price index after applying the proposed methodology to all the 9,180 timeframes. However, since the timeframes overlapped, it was necessary to clean the results to identify the observations that marked the end of one trend and the beginning of the next one. The flowchart in Figure 3.10 is a simplified version of the proposed algorithm used to clean the results and to identify the final set of critical points for the three price indexes.

As shown in this figure, the process starts by checking if any critical points were found in the first three-month gap, months 1 to 4. If no critical points were found, the timeframe was extended by one month (months 1 to 5) and reviewed again to verify if a critical point was determined. This process was repeated until all the possible timeframes for that trend range (month 138) were reviewed. After the trend range reached month 138, the beginning of the framework was moved one month forward (month 2), and the first three-month (months 2-4) timeframe was reviewed for critical points.

If a critical point was found, it was recorded and marked as the beginning of the next trend range. The search for significant trend changes continued from this month forward. For example, the first critical point in the asphalt price index was determined when evaluating the timeframe comprised of months 1 to 9. It revealed a critical point at month 3. Thus, the evaluation of all other timeframes beginning at months 1 and 2 would be discarded since it was already known that the next trend range began at month 3. Even if the evaluation of the timeframe between months 1 to 11 revealed a significant trend change at month 5, it would be irrelevant because a previous trend change had already been identified. The iterative process in Figure 3.10 was conducted across all three price indexes until all the critical points were identified for each of them.

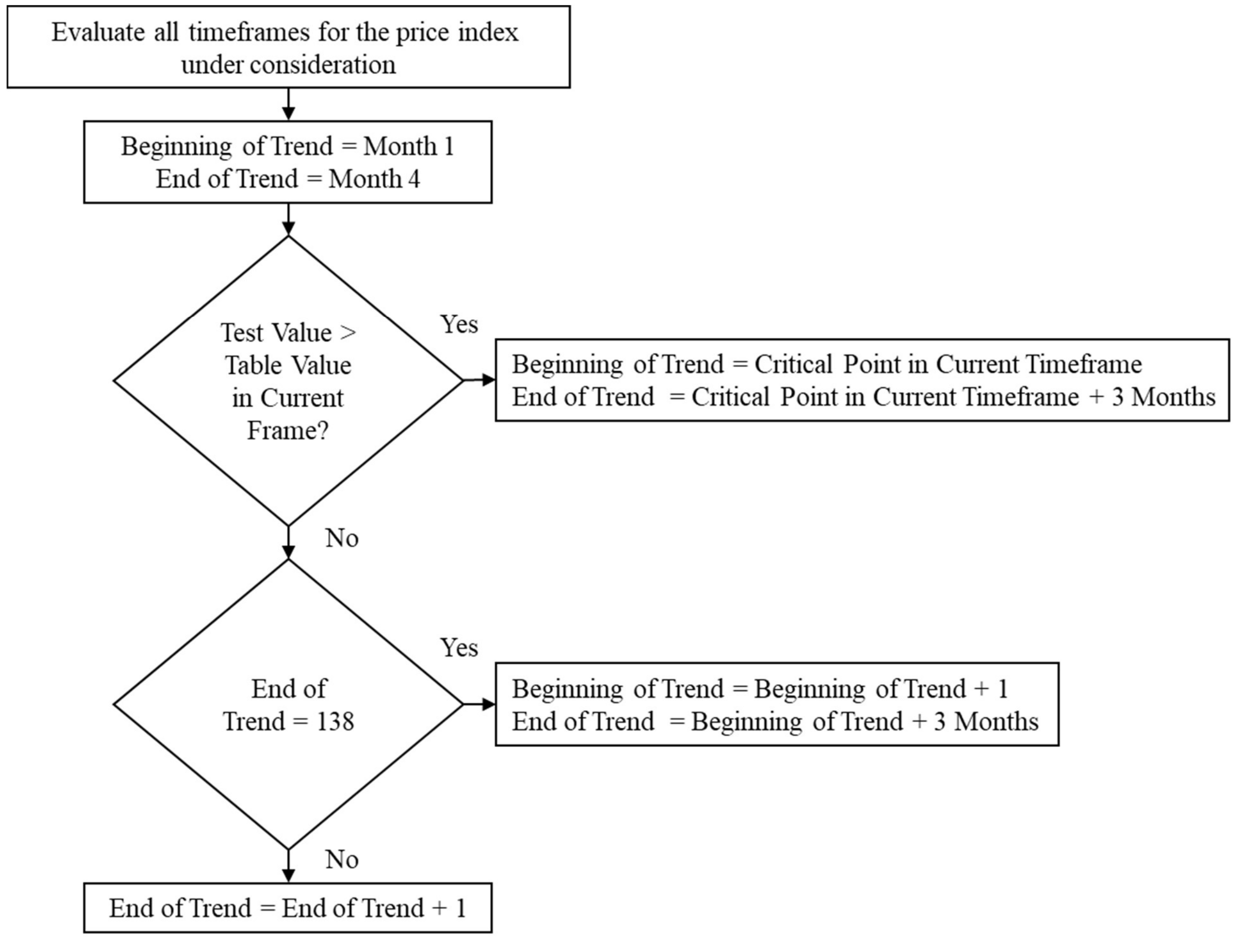

Figure 3.10 Identifying the Critical Points

Table 3.6 presents the critical points established for the asphalt, crude oil, and fuel price indexes according to the methodology described through this chapter at a 99 percent confidence level (p = 0.01). A total of 26, 23, and 20 critical points were identified for the asphalt, crude oil, and fuel price indexes, respectively. The difference in the amount of critical significant changes within the three price indexes was indicative of the level of volatility within each price index. Figure 3.11 to Figure 3.13 represent the three price indexes with their respective critical points. A visual inspection of these three figures was a preliminary validation of the effectiveness of the modified CUSUM approach from the critical points matching the major peaks and trends on each of the three time series.

Table 3.6 Critical Significant Month and Price Index Values

Asphalt		Crude Oil		Fuel	
Critical Points		Critical Points		Critical Points	
Month	Price Index	Month	Price Index	Month	Price Index
3	1.08	7	1.77	8	2.25
7	1.64	13	1.30	11	1.57
10	1.70	20	1.72	26	2.35
17	1.37	26	2.27	31	3.62
24	1.33	31	3.18	36	1.30
29	1.71	36	0.98	44	1.93
33	3.32	42	1.66	57	1.97
39	1.77	57	1.79	65	3.27
42	1.70	64	2.61	70	2.67
47	1.78	69	2.04	76	3.31
52	2.10	75	2.53	78	2.65
54	2.12	78	1.96	81	3.15
59	1.92	80	2.24	93	2.96
66	2.56	83	2.06	95	2.57
72	2.38	90	2.28	103	2.91
78	2.67	92	2.54	109	1.51
82	2.54	95	2.23	114	1.93
89	2.52	102	2.52	122	0.99
99	2.40	109	1.12	126	1.52
105	2.59	114	1.42	128	1.24
107	2.55	122	0.72	-	-
111	2.09	125	1.11	-	-
116	1.95	134	1.27	-	-
123	1.42	-	-	-	-
125	1.34	-	-	-	-
132	1.39	-	-	-	-

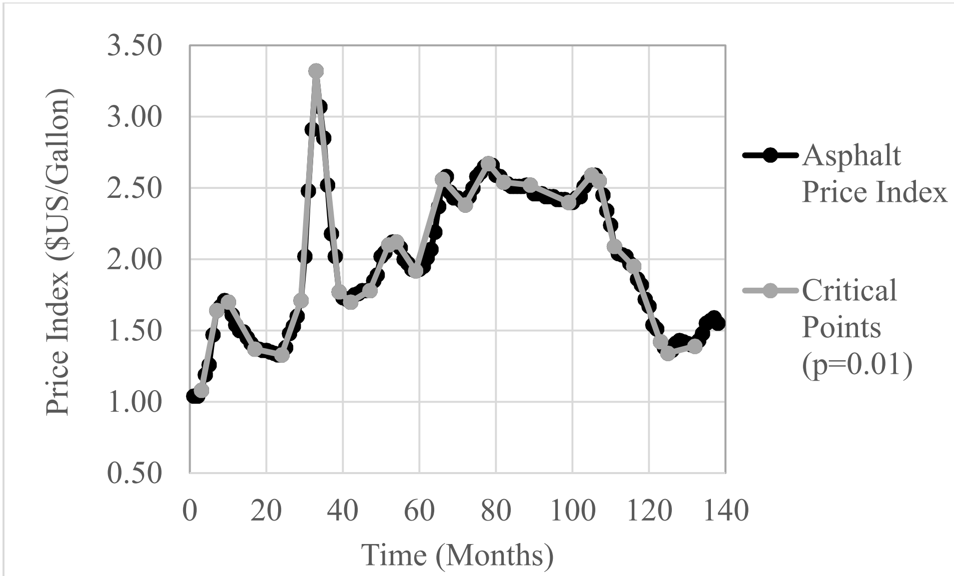

Figure 3.11 Asphalt Price Index Critical Points

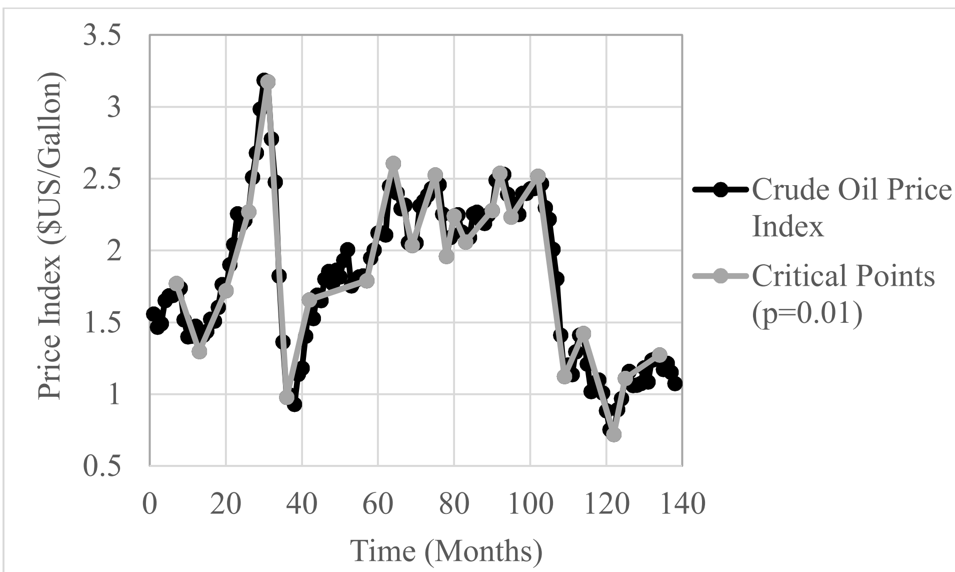

Figure 3.12 Crude Oil Price Index Critical Points

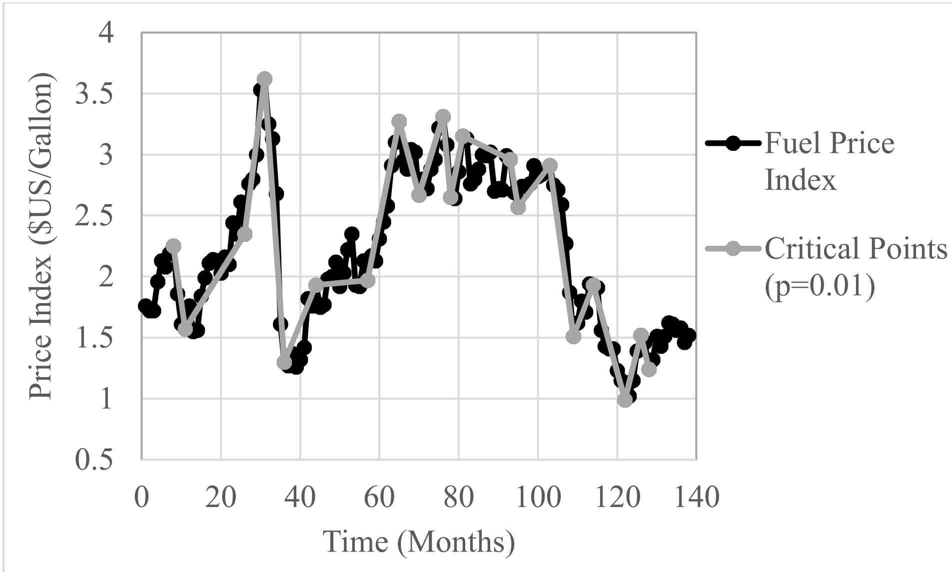

Figure 3.13 Fuel Price Index Critical Points

The next step in this study consisted of pairing the critical asphalt price indexes with their corresponding previous significant trend changes in the crude oil and fuel critical price indexes. After considering various approaches, the critical points were visually matched. They were matched based on the assumption made in the beginning that crude oil's trend changes occurred first then fuel's trend then asphalt's trend. Using this assumption, common critical peaks and trends as seen in Figures 3.11 to Figure 3.13 were matched visually.

This process reduce the amount of observations from 138 to less than 30. Through the visual inspection, only points showing a clear connection from Figure 3.11 to Figure 3.13 were kept for the data analysis process described in the following chapter. Through pairing the critical points the following questions could be answered.

1. How long does it take for crude oil and fuel price index fluctuations to impact the asphalt price indexes?
2. What proportion of the fluctuations in crude oil and fuel price indexes appear in the asphalt price indexes?

The critical asphalt to crude oil and asphalt to fuel paired points are listed in Tables 3.7 and 3.8. The alignment of the pairs are represented in Figure 3.14 and Figure 3.15. From these pairings, the 26 asphalt, 23 crude oil, and 20 fuel critical points were further reduced to 13 asphalt-to-crude oil and 11 asphalt-to-fuel paired points. These critical paired points were further evaluated in *Chapter 4: Data Analysis*.

Table 3.7 Critical Paired Points – Asphalt and Crude Oil

Asphalt - Critical Points		Crude Oil - Critical Points	
Month	**Asphalt Price Index**	**Month**	**Crude Oil Price Index**
10	1.70	7	1.77
17	1.37	13	1.30
29	1.71	26	2.27
33	3.32	31	3.18
39	1.77	36	0.98
47	1.78	42	1.66
59	1.92	57	1.79
66	2.56	64	2.61
72	2.38	69	2.04
78	2.67	75	2.53
99	2.40	95	2.23
105	2.59	102	2.52
125	1.34	122	0.72

Table 3.8 Critical Paired Points – Asphalt and Fuel

Asphalt - Critical Points		Fuel - Critical Points	
Month	**Asphalt Price Index**	**Month**	**Fuel Price Index**
10	1.70	8	2.25
29	1.71	26	2.35
33	3.32	31	3.62
39	1.77	36	1.30
59	1.92	57	1.97
66	2.56	65	3.27
72	2.38	70	2.67
78	2.67	76	3.31
99	2.40	95	2.57
105	2.59	103	2.91
125	1.34	122	0.99

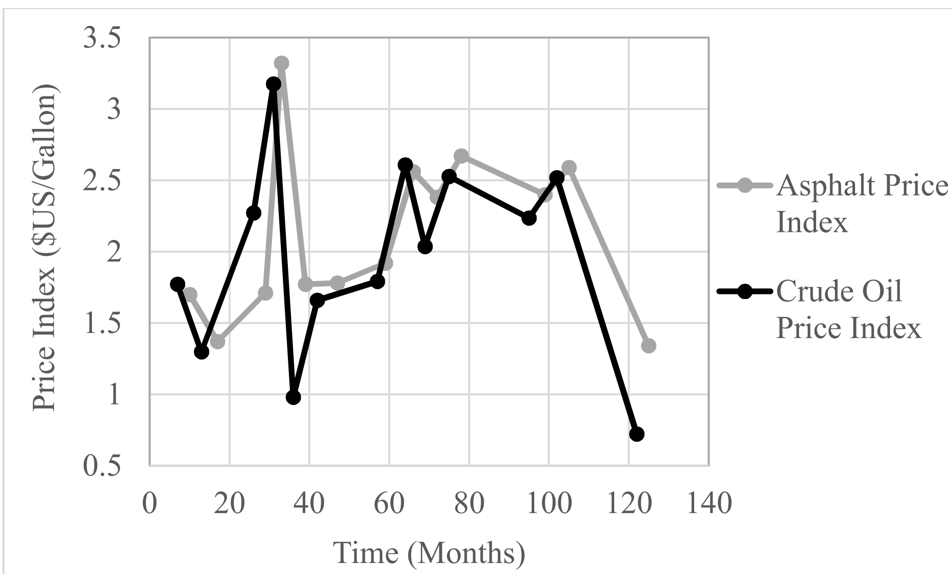

Figure 3.14 Critical Paired Points – Asphalt and Crude Oil

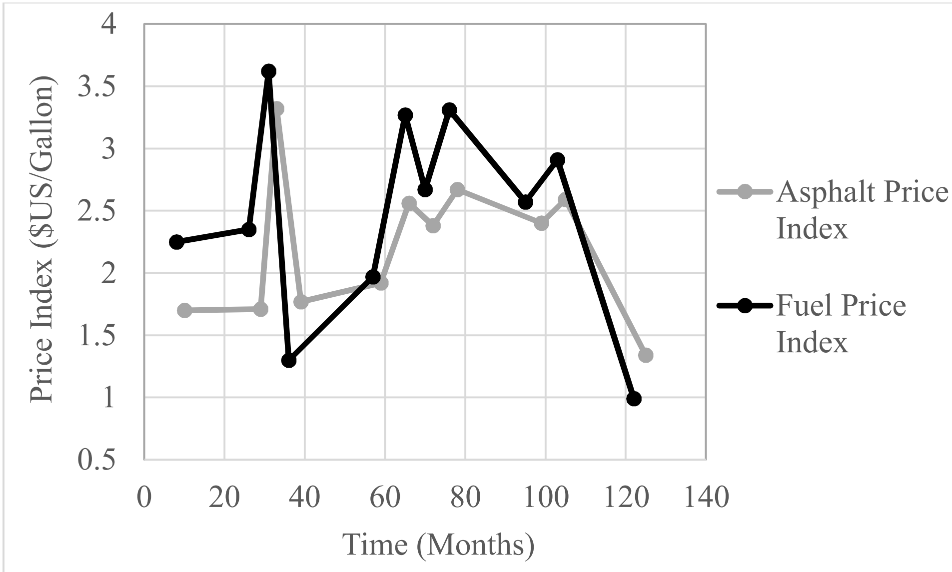

Figure 3.15 Critical Paired Points – Asphalt and Fuel

CHAPTER FOUR: DATA ANALYSIS

4.1 Introduction

The purpose of Chapter 4 is to use the findings in Chapter 3 to answer the research objectives within this thesis. From Chapter 1, the objectives that will be answered within this chapter are as follows:

1. Identify the time between when a change in crude oil's or fuel's price index will impact asphalt's price index.
2. Identify of the percentage change in the price index of crude oil or fuel, the amount of that change that impacts asphalt's price index.
3. Using the time and percent changes of the asphalt price index based on the price index of crude oil or fuel, develop risk analysis tables to evaluate all possibilities of time or percent changes in asphalt's price index in the future.

The three objectives above can be divided into a deterministic and a stochastic data analysis phase. While the deterministic data analysis phase was intended to find most likely values for crude oil-to-asphalt and fuel-to-asphalt time gaps and percent change ratios (objectives 1 and 2), the stochastic analysis was aimed to quantify and model the uncertainty associated with the deterministic values (objective 3). As seen from the paired critical points, there are variabilities in the time gaps and percent change ratios. Therefore, it would be unrealistic to assume that they can be represented with single deterministic values.

Based on the findings within the data analysis, a validation and testing of the findings were performed. These were performed through two cross-validation processes. The first cross-validation was aimed to assess the performance of the deterministic time gaps and percent change ratios in estimating asphalt price indexes, while the second cross-validation was performed to compare the results from this study against those obtained in previous similar studies.

Before the validation process was performed, the two sets of time gaps and percent change ratios (one between crude oil-to-asphalt paired points and one between the fuel-to-asphalt paired points) were reviewed for outliers. The outlier detection approach used in this study was the modified z-score method. This method is more suitable for smaller samples (46; 47), such as those corresponding to the amount of crude oil-to-asphalt and fuel-to-asphalt paired points; 13 and 11 paired points. This was to ensure that the results from this study were not impacted by atypical high or low time gaps or percent change ratios.

After presenting the final results, this chapter briefly discusses the results and their potential impact on horizontal construction. The following section addresses the first of the three objectives outlined above.

4.2 Identifying the Time Gap

This section is used to present the deterministic analysis conducted to determine the most likely time gaps between the asphalt and crude oil price indexes and the asphalt and fuel price indexes. The process to calculate the two mean time gaps were as follows. The mean crude oil-to-asphalt and fuel-to-asphalt time gaps were the arithmetic mean of the gaps between the 13 and 11 paired points listed in Tables 3.7 and 3.8. The time gap for each pair of points equal to the difference in their month values. The calculated time gaps for each pair of points and the two mean values are shown in Table 4.1.

Table 4.1 Time Gaps of Crude Oil and Fuel to Asphalt

Time Gaps (Months)	
Crude Oil to Asphalt	**Fuel to Asphalt**
3	2
4	3
3	2
2	3
3	2
5	1
2	2
2	2
3	4
3	2
4	3
3	-
3	-
Mean Time Gaps (Months)	
3	2

As shown in Table 4.1, a price fluctuation in the crude oil price index would be reflected in the asphalt price index approximately 3 months later. For fuel to asphalt, the most likely time gap was 2 months. In order to develop the risk assessment tools presented later in this chapter, it was necessary to define the probability distribution functions for these two time gaps. Those probability distributions represent all the possible time gap values with their probability of occurrence. Given that the time gaps were given by the publication frequency of price index values, which are published on a monthly basis, time gaps should be represented by discrete probability distributions.

Using @RISK, a statistical software package, the chi-square goodness of fit statistical test was used as an attempt to identify a suitable discrete probability distribution (e.g. binomial, poisson, discrete uniform, geometric) to represent the observed distribution of gaps in the two sets of paired critical points.

However, no suitable standard distributions were found to reasonably fit the observed time gaps. Therefore, empirical probability distributions given by the frequency of occurrence of the different time gaps in the paired points were used to model the variability in the values presented in Table 4.1. Figure 4.1 and Figure 4.2 illustrate these empirical distributions, showing that 3 and 2 months were the mean and most likely crude oil-to-asphalt and fuel-to-asphalt time gaps.

However, it should be noted that even though these are the most like values, these time gaps are only expected to occur about 50 percent the time. This highlights the importance of conducting a stochastic analysis. It should also be noted that time gaps between crude oil and asphalt price fluctuations tend to be larger than those between fuel and asphalt, giving ALDOT about one more month to prepare for significant asphalt price index changes.

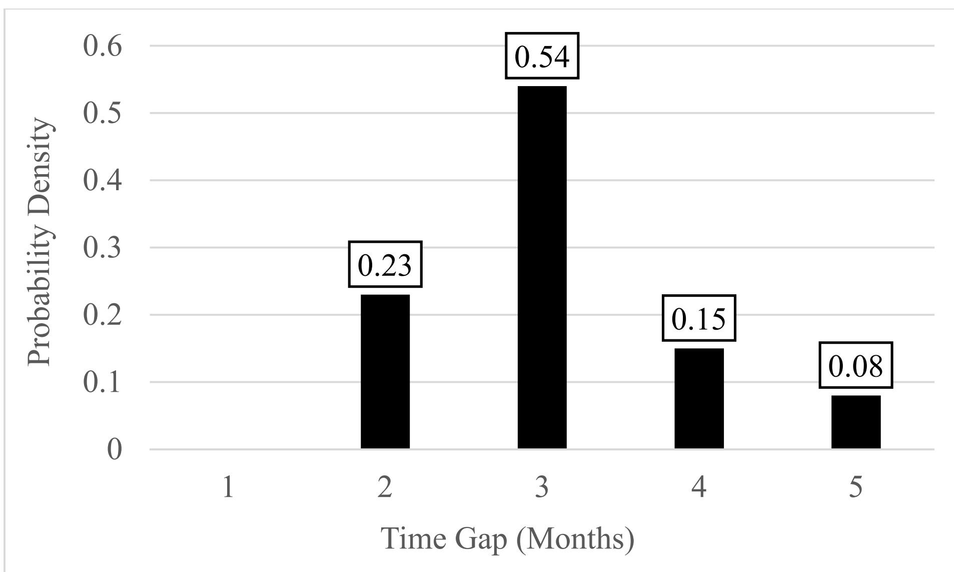

Figure 4.1 Most Likely Time Gap of Crude Oil to Asphalt

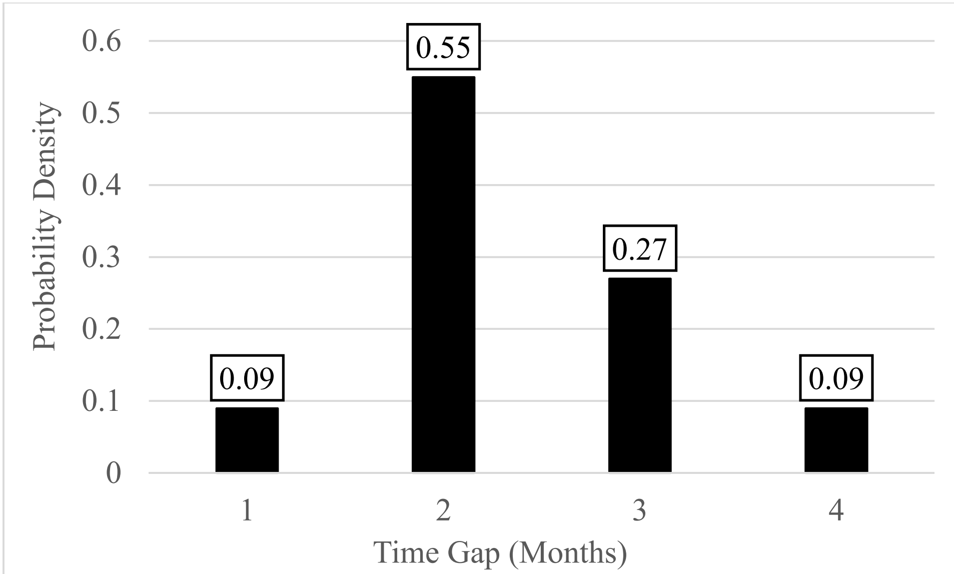

Figure 4.2 Most Likely Time Gap of Fuel to Asphalt

4.3 Identifying the Percent Change Ratios

Having addressed the first research objective, the percent change ratios between trends were defined by the critical points shown in Tables 3.7 and 3.8. The percent change ratios in this study were used to compare the slopes of the pairs of segments shown in Figures 3.14 and 3.15. Thus, the 13 crude oil-to-asphalt pairs of points define 12 pairs of segments. Similarly, there were 10 pairs of segments defined by the 11 fuel-to-asphalt critical paired points. To better understand the process followed to calculate the percent change ratios for each pair of segments, it is illustrated in this section by showing the calculations for the segments shown in Figure 4.3, which correspond to the first crude oil-to-asphalt pair of segments.

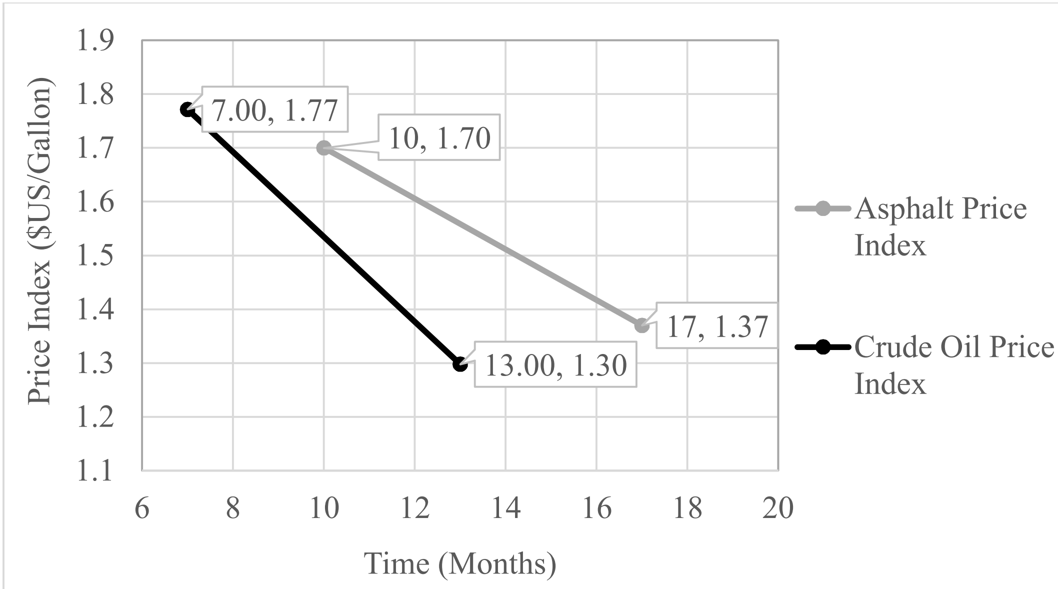

Figure 4.3 First Segment – Crude Oil to Asphalt

Figure 4.3 shows a 27 percent decrease between the first two critical points in the crude oil price index ([1.77-1.30]/1.77 = 0.27), while the first two critical points in the asphalt price index show a 19 percent decrease ([1.70-1.37]/1.70 = 0.19). If these two percent changes are compared for conclusion purposes, it would be concluded that a 0.70 percent change should be expected in the asphalt price index for each 1 percent change observed in the crude oil price index (0.19/0.27 = 0.7). However, it should be noted that the measured change in the crude oil price index occurs along a six-month time gap (month 13 – month 7 = 6 months), while the one measured in the asphalt price index occurs during a seven-month gap (month 17 – month 10 = 7 months). Therefore, these percent change ratios are not comparable. To address this issue, percent changes in this study were mean monthly percent changes calculated as shown in Equation 4.1.

Equation 4.1 $\boldsymbol{PC = \left(\frac{I_{second}}{I_{first}}\right)^{\frac{1}{M_{second} - M_{first}}} - 1}$

PC: Percent change for the segment under consideration

I_{first} : Price index value of the first observation in the segment

I_{second} : Price index value of the second observation in the segment

M_{first} : Month number of the first observation in the segment

M_{second} : Month number of the second observation in the segment

According to Equation 4.1, the percent changes for the crude oil and asphalt segments would be -5 percent and -3 percent (the negative sign represents a decrease in index values). The percent change ratio for the segments in Figure 4.3 was then calculated by dividing the asphalt change by the crude oil change, giving a percent change ratio of 0.60 (-0.03/-0.05 = 0.6). If a conclusive statement were to be made based only on these two segments, it would be said that a 60 percent change should be expected in the asphalt price index for every 1 percent change observed in the crude oil price index.

However, conclusions in this study were made after evaluating all pairs of segments shown in Figure 3.14 and Figure 3.15. Table 4.2 shows the percent change ratios for all crude oil-to-asphalt and fuel-to-asphalt pairs of segments. Based on the mean values at the end of Table 4.2, it can be concluded that a mean percent change of 58 and 46 percent could be expected in the asphalt index prices in Alabama for every 1 percent change observed in the crude oil and fuel price indexes.

Table 4.2 Percent Changes for the Price Indexes

Pairs of Segments	Crude Oil-to-Asphalt			Fuel-to-Asphalt		
	Asphalt Percent Change	Crude Oil Percent Change	Percent change Ratio	Asphalt Percent Change	Fuel Percent Change	Percent change Ratio
1	-0.03	-0.05	0.60	0.00	0.00	0.13
2	0.02	0.04	0.42	0.09	0.18	2.00
3	0.18	0.07	2.60	-0.19	-0.10	0.54
4	-0.10	-0.21	0.48	0.02	0.00	0.20
5	0.00	0.09	0.01	0.07	0.04	0.64
6	0.01	0.01	1.22	-0.04	-0.01	0.30
7	0.04	0.06	0.76	0.04	0.02	0.53
8	-0.01	-0.05	0.25	-0.01	-0.01	0.38
9	0.02	0.04	0.53	0.02	0.01	0.82
10	-0.01	-0.01	0.83	-0.06	-0.03	0.59
11	0.01	0.02	0.74	-	-	-
12	-0.06	-0.03	0.54	-	-	-
Average			**0.58**	**Average**		**0.46**

As shown in Table 4.2, the percent change ratio for the third pair of segments was not considered in the calculation of the mean crude oil-to-asphalt percent change ratio due to its outlying condition, which was determined by the modified z-score method. For the same reason, the second pair of segments was excluded from the calculation of the mean fuel-to-asphalt change ratio.

The modified z-score method was applied using Equations 4.2 and Equation 4.3. This method was used due to its suitability for small samples given that outliers are identified using the sample median ($\tilde{x}$) and the median absolute deviation (MAD), unlike other more commonly used outlier detection methods that are based on mean and standard deviation values (46; 47). It should be noted that in this study the modified z-score method was applied to small samples; the two sets of percent change ratios are listed in Table 4.2. Mean and standard deviation values are more sensitive to extreme values in small samples, increasing the risk of not detecting outliers that should be discarded (46; 47).

Equation 4.2 $MAD = median\{|x_i - x_o|\}$

x_i: Point in the data set being evaluated

x_o: Median of the original data set

Equation 4.3 $M_i = \frac{.6745(x_i - x_o)}{MAD}$

x_i: Point in the data Set being evaluated

x_o: Median of the Original Data Set

Observations with absolute modified z-scores greater than 3.5 ($|M_i|$>3.5) were discarded. After applying Equations 4.2 and Equation 4.3 to the percent change ratios in Table 4.2, modified z-score values of 7.5 and 5.2 were obtained for the third crude oil-to-asphalt pair of segments and the second fuel-to-asphalt pair of segments. Both values were greater than 3.5 and were discarded.

The outlying condition of this observation seems to be the result of an abnormal exponential increase in the asphalt price index between months 29 and 33 in comparison to its corresponding increases in the crude oil and fuel price indexes. This abnormal increase was easier to identify when the three price indexes were scaled to start with a price index value equal to one (index value at January 2006 = 1) and plotted on the same graph, as shown in Figure 4.4.

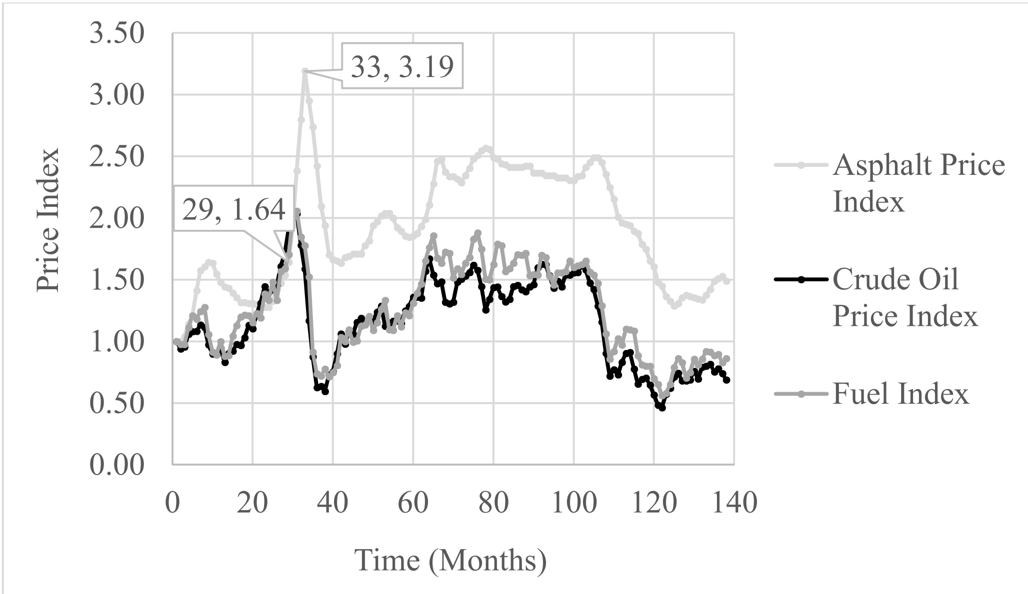

Figure 4.4 Scaled Price Indexes

The reason for this rapid increase in asphalt price indexes during this period of time is unknown and it is not required to be known for the purposes for this study. However, future research should be focused on analyzing these abnormal fluctuations and their causal factors. That could generate additional knowledge to refine the results and risk analysis tools presented in this thesis.

As done in the previous section for the time gap, the chi-square goodness of fit statistical test was used to find a suitable standard probability distribution to represent both sets of percent change ratios in Table 4.2, but in this case, it was applied to find a continuous probability distribution due to the continuous nature of the variable under consideration. The goodness of fit test revealed that Gamma distributions reasonably fit both sets of percent change ratios. Gamma distributions provided the lowest chi-square test statistics: 0.09 for crude oil-to-asphalt and 0.11 for fuel-to-asphalt. Figure 4.5 and Figure 4.6 show the Gamma distributions for the crude oil-to-asphalt and fuel-to-asphalt percent change ratios. These distributions will be used later in this chapter to develop the risk assessment tools, but first it is necessary to validate the results of the deterministic analysis.

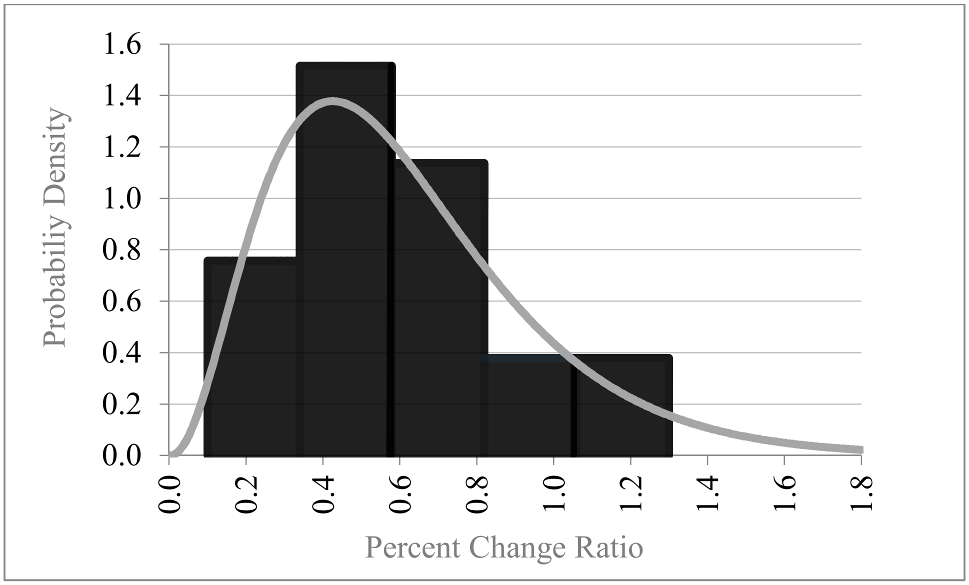

Figure 4.5 Crude Oil to Asphalt Percent Impact

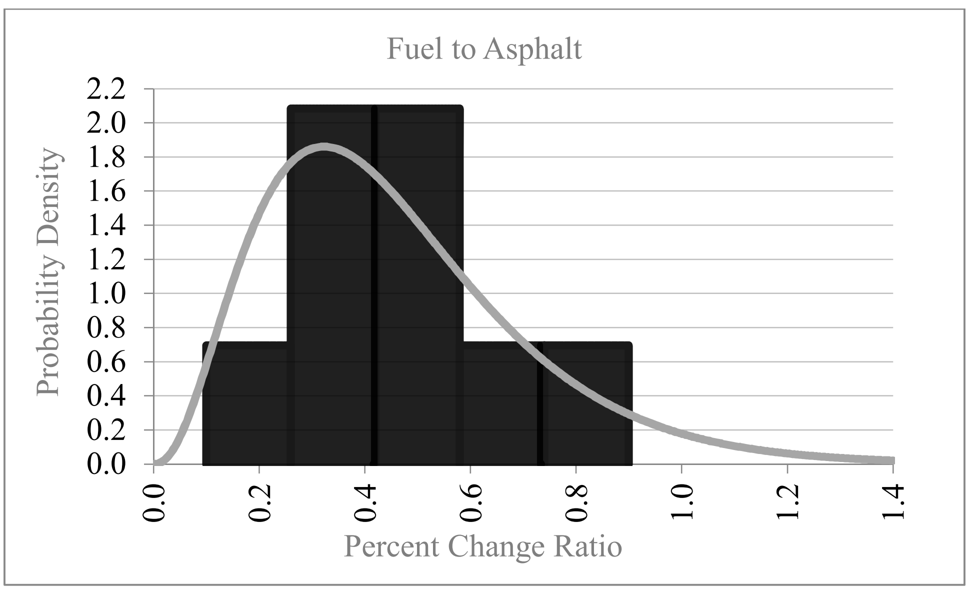

Figure 4.6 Fuel to Asphalt Percent Impact

To summarize the findings for the time gap and percent changes determined within the price indexes, the percent change ratios obtained through the paired critical points represent the level of impact of crude oil and fuel on asphalt. The time gap represents the time it would take for those percent changes to appear in the asphalt's price index. By stating that the percent change ratio was 58 for crude oil to asphalt means if there were a 1 percent change in crude oil's price index, 58 percent of that 1 percent change would appear in asphalt's price index. For example, if the crude oil's price index were to increase by 2, then asphalt's price index would then increase by 1.16 in 3 months. This same practice would apply to the percent change ratio and time gap calculated for fuel to asphalt.

4.4 Cross-Validation

The expected asphalt price indexes from June 2017 to July 2018 created by the crude oil price indexes had a mean absolute percentage error (MAPE) of 2.2 percent in comparison to the observed asphalt price indexes.

1. The expected asphalt price indexes from June 2017 to July 2018 created by the fuel price indexes had a MAPE of 3.9 percent in comparison to the observed asphalt price indexes.
2. The time gaps and percent change ratios obtained from the study by Attanasi (2008), had a significantly higher ($p = 0.05$) MAPE value of 5.5 percent for the expected asphalt price indexes created by the crude oil price indexes (9).
3. The time gaps and percent change ratios obtained from the PCA study (2014), had a higher MAPE value of 3.0 percent for the expected asphalt price indexes created by the crude oil price indexes (8).

The validation portion of the data analysis proved that the methodology and data analyses used to obtain the time gap and percent change ratios were effective methods to predict future asphalt price indexes. Comparing them to previously performed studies on the same topic, proved that the methods used in this thesis were significantly more effective and had a higher accuracy than at least one of the previous studies performed.

As mentioned before, two different cross-validation processes were performed in this thesis. The first cross-validation was aimed to assess the performance of the deterministic time gaps and percent change ratios in estimating asphalt price indexes, as well as to determine which of the two price indexes, crude oil or fuel, would better serve as a benchmark to anticipate asphalt price index fluctuations. On the other hand, the second cross-validation was intended to compare this study's results against those obtained in previous similar studies.

Cross-validation is commonly used in data-driven modeling and quantitative research to ensure the integrity and reliability of the obtained results. More specifically, this study used the holdout cross-validation approach, in which the available data was split into a training and a testing dataset (48; 49). The training data set was then used to create the desired models or estimate the intended values, while the testing dataset was used to assess the performance and/or accuracy of those models or values. The main principle of cross-validation was to not use the same observations to obtain and validate research results. False positive validation results may be obtained when this issue is ignored and the same data set is used for training and testing.

As explained earlier in this thesis, price index values for 150 months (January 2006 – June 2018) were initially collected for the asphalt, crude oil, and fuel price indexes. The first 138 months (January 2006 – June 2017) were used to accomplish the research objectives and the remaining 12 months (July 2017 – June 2018) were used to validate those accomplishments. In other words, the first 138 index values were used to assess time gaps and percent change ratios, which were presented in Sections 4.2 and 4.3, and the last 12 months were used for the cross-validation presented in this section.

First Cross-Validation:

For the first cross-validation, Equation 4.4 shows how the crude oil-to-asphalt mean time gap and percent change ratios were used to estimate asphalt price index values for each of the 12 months in the testing data set. The estimated asphalt price index for each month was calculated by applying a 58 percent (mean crude oil-to-asphalt percent change ratio) of the fluctuation observed in the crude oil price index three months before (mean crude oil-to-asphalt time gap) the asphalt price index from the previous month. In the same way, Equation 4.5 was used to estimate asphalt price index values for the testing dataset using observed fluctuations in the fuel price index.

Equation 4.4 $$API_i = API_{i-1} \times \left(1 + \left(\frac{OPI_{i-3} - OPI_{i-4}}{OPI_{i-4}} \times 0.58\right)\right)$$

Equation 4.5 $$API_i = API_{i-1} \times \left(1 + \left(\frac{OPI_{i-2} - OPI_{i-3}}{OPI_{i-3}} \times 0.46\right)\right)$$

API_i: Asphalt price index at month i

OPI_{i-3}: Crude oil price index at month $i-3$

i: Testing month under consideration

The asphalt price indexes estimated for the 12 testing months using Equations 4.4 and Equation 4.5 are presented in Table 4.3. This table shows the actual asphalt price indexes recorded for each month, the absolute percentage error (APE) for each estimated value, the mean absolute percentage errors (MAPEs) for each set of errors, and the APE standard Deviation (SD). APE and MAPE values were calculated as shown in Equations 4.6 and Equation 4.7.

Table 4.3 First Cross-Validation: Actual vs. Estimated Asphalt Price Index

Month	Actual Asphalt Price Index	Estimated Asphalt Price Index		Absolute Percentage Error	
		From Crude Oil Index	From Fuel Index	From Crude Oil Index	From Fuel Index
17-Jul	1.54	1.58	1.5	2.70%	2.90%
17-Aug	1.5	1.49	1.57	0.30%	4.60%
17-Sep	1.49	1.44	1.47	3.30%	1.20%
17-Oct	1.48	1.52	1.56	2.50%	5.40%
17-Nov	1.48	1.51	1.58	1.80%	6.90%
17-Dec	1.52	1.51	1.39	0.50%	8.70%
18-Jan	1.57	1.55	1.59	1.20%	1.00%
18-Feb	1.64	1.66	1.58	1.20%	3.50%
18-Mar	1.69	1.66	1.67	1.70%	1.50%
18-Apr	1.76	1.79	1.74	1.60%	1.30%
18-May	1.83	1.74	1.69	5.10%	7.90%
18-Jun	1.93	1.84	1.89	4.70%	2.00%
		MAPE		**2.20%**	**3.90%**
		SD		**1.50%**	**2.80%**

Equation 4.6 $APE_i = \frac{|API_{actual;i} - API_{estimated;i}|}{API_{actual;i}} \times 100\%$

Equation 4.7 $MAPE = \frac{\sum_{i=1}^{n} API_i}{n}$

APE_i: Absolute percentage error for month i

$API_{actual;i}$: Actual asphalt price index at month i

$API_{estimated;i}$: Estimated asphalt price index for month i

$MAPE$: Mean absolute percentage error

n: Number of testing months

i: Testing month under consideration

A statistical F-test to compare the SDs of the APEs obtained from the crude oil and fuel price indexes revealed a statistical significant difference between the SD values at a 95 percent confidence level (p=0.05). This means that the estimating uncertainty was significantly lower when asphalt price index values were estimated from the crude oil price index. Likewise, a paired two-sample t-test revealed that the MAPE value from the crude oil price index was significantly lower than the MAPE yielded by the fuel price index at a 95 percent confidence level (p=0.05). Meaning that the overall accuracy was significantly higher when the asphalt price indexes were estimated from the crude oil price index. Thus, these two statistical tests showed the estimation of asphalt price index values was significantly more effective, in terms of accuracy and reliability, when estimated from observed crude oil price index fluctuations.

However, even though the crude oil-to-asphalt time gap and percent change ratios proved to be more accurate at estimating future asphalt price indexes, both MAPE values could be considered to reflect high accuracy levels when compared against the expected cost estimating errors determined by AASHTO. According to AASHTO, expected accuracy in final design estimates is expected to be between -5 percent and +10 percent (14). Thus, this study is considered a positive result based on the MAPE values being under 4 percent in a commodity (asphalt) that in some projects could represent over 80 percent of the total construction cost.

Second Cross-Validation:

From the previous studies on this topic presented in Chapter 2, two different crude oil-to-asphalt time gaps and percent change ratios were compared through a second cross-validation to the ones determined in this thesis from the crude oil price index. The first time gap and percent change ratio came from a study performed by Attanasi (2008). Attanasi estimated a mean time gap of 1 month and percent change ratio of 130 percent (9). The second similar previous study in the second cross-validation process corresponds to a study conducted by Sullivan and Chiappe (2014). Sullivan and Chiappe's estimated an average time gap of 6 months and a percent change ratio of

70 percent (8). The process for the second cross-validation is the same as the one described above for the first cross validation. The only difference was in the time gaps and percent change ratios used to estimate asphalt price index values. Table 4.4 summarizes the results for the second-cross validation.

Table 4.4 Second Cross-Validation: Actual vs. Estimated Asphalt Price Index

Month	Actual Asphalt Price Index	Estimated Asphalt Price Index			Absolute Percentage Error		
		From Current Study	From Attanasi	From Sullivan and Chiappe	From Current Study	From Fuel Index	From Sullivan and Chiappe
17-Jul	1.54	1.58	1.41	1.56	2.70%	8.30%	1.40%
17-Aug	1.50	1.49	1.60	1.56	0.30%	7.00%	4.00%
17-Sep	1.49	1.44	1.56	1.42	3.30%	4.60%	4.80%
17-Oct	1.48	1.52	1.56	1.53	2.50%	5.50%	3.10%
17-Nov	1.48	1.51	1.55	1.43	1.80%	4.60%	3.50%
17-Dec	1.52	1.51	1.67	1.41	0.50%	9.80%	7.30%
18-Jan	1.57	1.55	1.56	1.55	1.20%	0.40%	1.00%
18-Feb	1.64	1.66	1.78	1.60	1.20%	8.20%	2.20%
18-Mar	1.69	1.66	1.59	1.68	1.70%	5.90%	0.40%
18-Apr	1.76	1.79	1.71	1.73	1.60%	3.00%	1.60%
18-May	1.83	1.74	1.89	1.88	5.10%	3.20%	2.80%
18-Jun	1.93	1.84	1.83	1.86	4.70%	5.20%	3.70%
			MAPE		**2.20%**	**5.50%**	**3.00%**
			SD		**1.50%**	**2.60%**	**1.90%**

As shown in Table 4.4, both of the previous studies showed higher estimating MAPE and SD values than those obtained from the crude oil price index. A paired two-sample t-test and F-test showed that the MAPE and SD from the crude oil price indexes were significantly lower than those obtained from Attanasi's estimated time gap and percent change ratio. In other words, the results from this thesis were more effective at estimating asphalt price index values in the state of Alabama than Attansi's results.

On the other hand, the paired two-sample t-test and the F-test between this thesis' results and Sullivan and Chiappe's time gap and percent change ratios failed to find a significant difference between the MAPE and SD at a 99% confidence level (p=0.01). In spite of the results from these statistical tests, the use of this study's crude oil-to-asphalt time gap and percent change ratios to estimate asphalt price indexes in Alabama would be a better fit than those provided by Sullivan and Chiappe.

This recommendation is based on two facts. First, a failure to reject the null hypothesis in these tests does not mean that the estimating performance of both approaches were the same. It means that it was not possible to find a significant difference between these approaches with the observations provided. The MAPE value for this study's crude oil-to-asphalt time gap and percent change ratio was 25 percent lower than the MAPE obtained using Sullivan and Chiappe's time gap and percent change ratio. Likewise, the SD was 19 percent lower, in favor of this thesis. Therefore, it would be reasonable to assume that it is more likely to achieve higher accuracy and reliability if using a crude oil-to-asphalt time gap and percent change ratio of 3 and 58 percent.

The second reason that it is recommended to use this study's time gap and percent change ratio over those estimated by Sullivan and Chiappe is because of the results applying these two approaches to the 138 months in the training dataset. It showed lower MAPE and SD values, in favor of this thesis, and in this case, the paired two-sample t-test and F-test showed a significant difference between the values at a 99 percent confidence level (p=0.01). MAPE and SD values obtained from the training data set are shown in Table 4.5. This study was not intended to make conclusive statements from validation results from the training dataset since it was against the cross-validation principles. However, it is seen as an indication that a time gap of 3 and percent change ratio of 58 percent could actually provide higher accuracy and reliability in asphalt price index estimating.

Table 4.5 MAPE and SD Values from Training Dataset

	From Current Study	From Sullivan and Chiappe
MAPE	3.70%	5.70%
SD	3.10%	4.80%

After validating the results of the deterministic analysis presented in Sections 4.2 and 4.3, the risk assessment tools intended to quantify the uncertainty associated with crude oil-to-asphalt and fuel-to-asphalt time gaps and percent change ratios were created. These risk assessment tools are presented in the following section.

4.5 Risk Assessment Tools

Although it was already determined that the most likely time gaps between crude oil and asphalt and fuel and asphalt price index fluctuations are 3 and 2 months, it was also found that the occurrence of those time gaps should only be expected about 50% of the time (see Figure 4.1 and Figure 4.2). This means that there was a considerable amount of uncertainty associated with these time gaps, which should be considered by STAs to make asphalt pricing decisions. Likewise, Figure 4.5 and Figure 4.6 show some variability to be considered in the crude oil-to-asphalt and fuel-to-asphalt percent change ratios.

For example, even though the most likely crude oil-to-asphalt percent change ratio was 58 percent, the Gamma distribution in Figure 4.5 indicates that there was only a 26 percent probability of having an actual ratio between 40 and 60 percent. Table 4.6 shows the probability of occurrence of different crude oil-to-asphalt and fuel-to-asphalt percent change ratios within six different ranges according to the Gamma distributions illustrated in Figures 4.5 and 4.6.

Table 4.6 Percent Change Ratios Variability

Percent Change Ratio Range	Probability of Occurrence	
	Crude Oil-to-Asphalt	Fuel-to-Asphalt
>1.0	0.12	0.03
0.8 - 1.0	0.12	0.06
0.6 - 0.8	0.20	0.15
0.4 - 0.6	0.26	0.28
0.2 - 0.4	0.24	0.35
0.0 - 0.2	0.07	0.13

The proposed risk assessment tool was intended to provide STAs with the probability of occurrence of 24 different case scenarios given by four possible time gaps in Figure 4.1 and Figure 4.2 and the six ranges in Table 4.6 (4 time gaps x 6 percent change ratio ranges = 24 possible case scenarios). According to probability theory principles, the probability of occurrence of each of the 24 case scenarios was calculated as the joint probability of two independent events as shown in Equation 4.8. Thus, the probability of having a given observed change in the crude oil price index affecting the asphalt market 3 months later (P(A=3) = 0.54) and with a percent change ratio between 40 and 60 percent (P(B=[0.4;0.6]) = 0.26) would be equal to 14 percent (0.54 x 0.26 = 0.14). It should be noted that this is the most likely crude oil-to-asphalt scenario with only a 14 percent probability of occurrence. It shows one more time the importance of this stochastic analysis and the value of the proposed risk assessment tools.

Equation 4.8 $P(A \cap B) = P(A) \times P(B)$

$P(A \cap B)$: Probability of having an observed fluctuation in crude oil prices reflected in asphalt prices “A” months later and with a percent change ratio within range “B”

$P(A)$: Probability of having an observed fluctuation in crude oil prices reflected in asphalt prices "A" months later

$P(B)$: Probability of having and observed fluctuation in crude oil prices reflected in asphalt prices with a percent change ratio within range "B"

A: Time gap under consideration

B: Percent change ratio range under consideration

The proposed risk assessment tools organize the probability of occurrence of all possible case scenarios into a matrix to facilitate the use of these tools by a STA. Table 4.7 and Table 4.8 show the crude oil-to-asphalt and fuel-to-asphalt risk assessment tools. These tables use a color scale to show the change in the probability of occurrence across the matrices. The color scale goes from dark red for the highest probability of occurrence to dark blue for the lowest one.

Table 4.7 Crude Oil-to-Asphalt Risk Assessment Tool

	Change Ratio Range	**Percent Change Range**	**Probability**				
Percent Change Ratio	**>1.0**	>	0.12	2.80%	6.60%	1.80%	1.00%
	0.8-1.0	;	0.12	2.70%	6.40%	1.80%	1.00%
	0.6-0.8	;	0.19	4.50%	10.60%	2.90%	1.60%
	0.4-0.6	;	0.26	6.00%	14.20%	3.90%	2.10%
	0.2-0.4	;	0.24	5.40%	12.70%	3.50%	1.90%
	0.0-0.2	;	0.07	1.50%	3.50%	1.00%	0.50%
			Probability	0.23	0.54	0.15	0.08
				2	**3**	**4**	**5**
				Crude Oil-to-Asphalt Time Gap (Months)			

Table 4.8 Fuel-to-Asphalt Risk Assessment Tool

	Change Ratio Range	**Percent Change Range**	**Probability**				
Percent Change Ratio	**>1.0**	>	0.03	0.30%	1.80%	0.90%	0.30%
	0.8-1.0	;	0.06	0.50%	3.30%	1.60%	0.50%
	0.6-0.8	;	0.15	1.30%	8.00%	3.90%	1.30%
	0.4-0.6	;	0.28	2.50%	15.50%	7.60%	2.50%
	0.2-0.4	;	0.35	3.20%	19.50%	9.60%	3.20%
	0.0-0.2	;	0.13	1.10%	6.90%	3.40%	1.10%
			Probability	0.09	0.55	0.27	0.09
				1	**2**	**3**	**4**
				Fuel-to-Asphalt Time Gap (Months)			

The risk assessment tools in Table 4.7 and Table 4.8 are to be used once a given fluctuation in the crude oil or fuel price index has been observed. Once the magnitude and direction of the fluctuation has been determined, an STA can proceed to fill out the "Percent Change Range" column (third column – highlighted) before using the tool to support an asphalt pricing decision. This column uses Equation 4.9 and Equation 4.10 to transform the percent change ratio ranges into ranges of actual percent changes based on the observed fluctuation in the crude oil or fuel price index.

Equation 4.9 $PCR_{low,i} = CRR_{low,i} \times OIC$

Equation 4.10 $PCR_{high,i} = CRR_{high,i} \times OIC$

$PCR_{low,i}$: Low end of percent change range for range i

$PCR_{high,i}$: High end of percent change range for range i

$CRR_{low,i}$: Low end of change ratio range for range i

$CRR_{high,i}$: High end of change ratio range for range i

OIC: Observed index change (crude oil price index or fuel price index, as applicable)

For example, between January and February 2018, there was a decrease of 2.3 percent in the crude oil price index. Equation 4.9, Equation 4.10, and Table 4.7 would be filled out as shown in Table 4.9 to determine the asphalt price index change. From the 2.3 percent change in crude oil's price index would mean that there was a 14.2 percent probability of having this fluctuation reduce the asphalt price index between 0.9 and 1.4 percent by May 2018 (three months later).

Table 4.9 Example - Crude Oil-to-Asphalt Risk Assessment Tool

	Change Ratio Range	**Percent Change Range**	**Probability**				
Percent Change Ratio	**>1.0**	**2.3>**	0.12	2.80%	6.60%	1.80%	1.00%
	0.8-1.0	**-1.8%;-2.3%**	0.12	2.70%	6.40%	1.80%	1.00%
	0.6-0.8	**-1.4%;-1.8%**	0.20	4.50%	10.60%	2.90%	1.60%
	0.4-0.6	**-0.9%;-1.4%**	0.26	6.00%	14.20%	3.90%	2.10%
	0.2-0.4	**-0.4%;-0.9%**	0.24	5.40%	12.70%	3.50%	1.90%
	0.0-0.2	**0.0%;-0.4%**	0.07	1.50%	3.50%	1.00%	0.50%
			Probability	0.23	0.54	0.15	0.08
				2	**3**	**4**	**5**
				Crude Oil-to-Asphalt Time Gap (Months)			

4.6 Discussion of Findings

From the analyses performed within this chapter, the research objectives presented within the introduction of this chapter have been accomplished. Using the methods and analyses from the asphalt, crude oil, and fuel price indexes the following items were established:

1. The time it would take for a change in crude oil's price index to impact asphalt's price index was 3 months. Likewise, the time it would take for a change in fuel's price index to impact asphalt's price index was 2 months.
2. For every 1 percent change in crude oil's price index, 58 percent of that change would appear in the asphalt price index. Likewise, for every 1 percent change in the fuel price index, 46 percent of that change would appear in the asphalt price index.
3. From the findings in research objectives 1 and 2, a risk analysis tool was created to determine the probabilities of various time gaps and percent change ratios of asphalt price index changes based on potential observed changes in either crude oil or fuel price indexes.

From the cross-validation analysis, the methods to identify the time gap, percent change ratios, and risk analysis tools were accurate and reliable in determining asphalt price index changes. The analyses within this chapter were broken into two parts: deterministic and stochastic. The deterministic values being the most likely time gaps and percent changes ratios, could provide STAs with improved engineering judgment when evaluating changes in asphalt price indexes based on crude oil or fuel price indexes without a multitude of calculations to determine these changes. The stochastic aspect being the risk analysis tools, provide STAs with the means to manage the volatility in asphalt with realistic probabilities of major changes occurring within a horizontal construction project.

CHAPTER FIVE: CONCLUSION

5.1 Introduction

Asphalt's production has become one of the main components of the horizontal construction industry. According to the National Asphalt Paving Association (NAPA), asphalt covers approximately 94 percent of the 2.7 million miles of paved roads in the United States (U.S.), and every year, the transportation construction industry sent approximately $80 billion dollars on this commodity (2). Given that one of asphalt's critical ingredient is bitumen, a byproduct of crude oil, it is usually assumed that fluctuations in crude oil and fuel (another byproduct of crude oil) prices are also reflected in the asphalt market. Even though it seems to be common knowledge, a lack of research has been done on better understanding and modeling market connections among these three commodities towards the improvement of pricing procedures associated with the most important material in the horizontal construction industry. This gap in the body of knowledge was the motivation behind the study presented in this thesis.

In this thesis, the following three questions were answered through a quantitative analysis of the relationship between asphalt price indexes maintained by the Alabama Department of Transportation (ALDOT) and two other price indexes; a national crude oil price index published by the EIA and ALDOT fuel price index ALDOT:

1. When there is a change in crude oil or fuel price indexes, how long does it take for that change to be reflected in the asphalt price index? This question was formulated under the assumption that price index fluctuations occur first in the crude oil price index, then

in the fuel price index, and finally in the asphalt price index. This assumption has been validated in this thesis.

2. What is the proportion of crude oil or fuel price index fluctuations that impact asphalt price indexes?
3. How much uncertainty is associated with the deterministic answers to the first two questions?

Monthly asphalt, crude oil, and fuel price index values, from January 2006 to June 2018, were collected and cleaned in this study. The first 138 months, of the 150 months of available data, were processed in Chapter 3 of this thesis to pair changes in the crude oil and fuel price indexes with their corresponding changes in asphalt price indexes. These paired significant changes were then used in Chapter 4 to answer the first two research questions listed above. The identification of corresponding fluctuations among these price indexes in Chapter 3 was performed using an innovative modified cumulative sum (CUSUM) statistical analysis method developed. The modified CUSUM method was used to recognize significant critical points along the three price indexes, facilitating a pairing process. The original 138 observations were reduced to 26, 23, and 20 observations in the asphalt, crude oil, and fuel price indexes, respectively. A visual inspection of this reduced sample allowed for the identification of 13 crude oil-to-asphalt and 11 fuel-to-asphalt corresponding points.

The two sets of paired critical points were analyzed in Chapter 4 to determine the crude oil-to-asphalt and fuel-to-asphalt time gaps (research question 1), as well as the percent change ratios that represent the magnitude of asphalt price changes as a function of previously observed changes in the crude oil and fuel price indexes (research question 2). The results of the deterministic data analysis were positively validated through cross-validation, by applying these results, as well as results from two previous similar studies, to estimate asphalt price index values for the last 12 months of data; the testing dataset. MAPE values were used to measures the estimating accuracy, and were calculated by comparing estimated values against the actual index values for each of the months in the testing dataset.

Chapter 4 presents the stochastic analysis conducted to answer the third research question. The analysis of the paired critical points used probability distribution functions to represent the time gaps and percent change ratios between the asphalt price index and the indexes for the other two commodities. Finally, these probability distributions were integrated to create two risk assessment tools that allow STAs to evaluate the probability of occurrence of all possible reactions of the asphalt price index to observed changes in the crude oil and fuel price indexes.

5.2 Conclusions

Based on the analysis of the paired critical points, this study found that the average and most likely crude oil-to-asphalt and fuel-to-asphalt time gaps in Alabama are 3 and 2 months. This means that the use of the crude oil price indexes would give ALDOT one more month to be prepared for upcoming asphalt price index changes than if fuel price indexes were used as a benchmark.

In a similar way, the study has found that a 58 percent change should be expected in the asphalt price index for every 1 percent change observed in the crude oil price index, while a 1 percent change in the fuel index would cause a 46 percent change in the asphalt price indexes in Alabama. These results show what seems to be a higher correlation between crude oil and asphalt price indexes than the correlation between fuel and asphalt price indexes. This could be explained by the fact that there is a "parent" relationship between crude oil and asphalt, while fuel is considered a "sister commodity" to asphalt. Although both fuel and asphalt are a byproduct of the refinement of crude oil, their market prices are affected in different ways by political and socio-economic issues, weakening the relationship between these two commodities.

Chapter 4 presents two cross-validation processes. The first cross-validation found that the use of the crude oil price index would not only give ALDOT an additional month to estimate asphalt price changes, but also would allow for the more accurate and reliable estimation of asphalt price indexes. This was proven through statistical significance testing at a 95 percent confidence level (p=0.05). However, asphalt price index estimates obtained with the fuel-to-asphalt time gap and percent change ratio found in this study should still be considered as reasonably accurate under AASHTO standards. According to AASHTO, accuracy in effective construction cost estimates at design completion could vary between -5 to +10 percent, and the absolute average estimated accuracy offered by the fuel-to-asphalt time gap and percent change error was under +4 percent.

In the second cross-validation, the results from this thesis were compared against crude oil-to-asphalt time gaps and percent change ratios from two previous studies performed on the same topic by Attanasi (2008) and Sullivan and Chiappe (2014) (8; 9). The comparison against Attanasi's results revealed that the asphalt price indexes estimated were significantly more accurate, at a 95 percent confidence level (p = 0.05), when estimated with the results from this study. When compared against Sullivan and Chiappe's findings, MAPE values and the SD of the APEs were also lower for asphalt price indexes estimated with the time gap and percent change ratio calculated in this study, but in this case, the superior performance could not be statistically demonstrated. However, the fact that this study's MAPE and SD values were 25 and 19 percent lower as well as statistically superior in estimating performance when applied, both approaches to the training dataset could be used from the crude oil-to-asphalt time gap and percent change ratio calculated in this study over those proposed by Sullivan and Chiappe.

After validating the deterministic results, a stochastic analysis was conducted to quantify and model the uncertainty associated with the time gaps and percent change ratios found in this study. As part of this analysis, @RISK, a statistical software was used to apply the chi-square goodness of fit statistical test in an attempt to identify suitable standard discrete probability distributions to represent the variability in the time gaps, as well as to find suitable continuous distributions for the percent change ratios.

None of the discrete distributions evaluated by @RISK showed a good fit to the distribution of time gaps showed by the paired critical points. Therefore, for these two variables (the crude oil-to-asphalt time gap and the fuel-to-asphalt time gap) empirical distributions given by the observed frequency of occurrences of the different time gap values across the paired critical points were used. On the other hand, the chi-square goodness of fit statistical test found that Gamma distributions were suitable to model the variability in both percent change ratios (the crude oil-to-asphalt percent change ratio and the fuel-to-asphalt percent change ratio). Both Gamma distributions, with their respective parameters, are shown in Chapter 4.

Finally, the defined probability distributions were used to create two risk assessment tools; one for crude oil-to-asphalt values and one for fuel-to-asphalt values. These tools were matrices that can provide STAs with the probability of occurrence of all possible reactions of the asphalt price index to changes in the crude oil and fuel price indexes in terms of when and by how much would the asphalt price index change. These tools are used to easily communicate the uncertainty between the asphalt price index and the price indexes for crude oil and fuel. A better understanding of this uncertainty would facilitate better decision-making on issues that might depend on asphalt price index fluctuations.

5.3 Research Limitations

The purpose of this thesis was to establish the relationships between asphalt, crude oil, and fuel for the purposes of predicting future asphalt price index values and to anticipate market changes. In order to achieve the research objectives, this study analyzed the relationship between three price indexes: two state price indexes for asphalt and fuel maintained by ALDOT and a national crude oil price index published by the EIA. Thus, the main limitation in this study was given by these three indexes.

The time gaps and percent change ratios calculated in this study should only be applied to correlate these three specific price indexes. Different asphalt, crude oil, or fuel indexes could be connected by different time gaps and ratios due to political, economic, or geographic considerations. Likewise, since the asphalt price index was tracking changes in the state of Alabama, the results of this study and the proposed risk assessment tools should only be used to estimate asphalt prices in the state of Alabama.

Another important limitation of this study was related to the visual inspection required to pair the significant critical points along the price indexes. Although this process was considerably simplified by the reduction in the number of observations from 138 to less than 30 for each price index, the visual inspection added some subjectivity to the process, meaning that different researchers could select different pairs of points, and therefore obtain different results. Further research on this topic should invest efforts on the development of objective processes to pair critical points.

5.4 Recommendations for Future Research

The objectives accomplished in this thesis were meant as a preliminary investigation for further research efforts intended to better understand the factors affecting asphalt prices in horizontal construction projects and learn how to read the behavior of other related markets to anticipate fluctuations in asphalt prices. Future research efforts derived from this study and anticipate are as follows:

1. The replication of this study for other STAs.
2. The development of an objective system to pair critical points across price indexes to eliminate the subjectivity of the visual inspection.
3. An in-depth analysis of abnormal fluctuations in asphalt price indexes. This refers to the significant changes in the asphalt price index that are related to crude oil price index changes. This research could help to refine the asphalt pricing process by improving estimating accuracy and reliability.
4. The development of an user-friendly system that considers not only observed crude oil or fuel price index fluctuations, but also other factors affecting prices in the asphalt market, to produce reliable time gaps, percent change ratios, and complete risk analysis reports.

From these findings, more effective asphalt pricing procedures for STAs are reflected in improved budget control and resource allocation procedures. Objectives STAs are constantly trying to improve in order to cope with the shrinking budgets and expanding infrastructure systems managed by these agencies. The research efforts presented in this thesis are expected to open the door to new research initiatives intended to develop and promote innovative ideas to facilitate a more effective use of taxpayers' money in the transportation construction industry.

REFERENCES

1. **Davis, J and Griffin, N.** *Celebrating 90 Years of Service to the Asphalt Industry.* Lexington : Asphalt Institute, 2009.

2. **National Asphalt Pavement Association.** *Fast Facts.* Lanham : s.n., 2018.

3. **European Asphalt Pavement Association .** *Driving Ahead with Sustainable Asphalt Roads.* Brussels : s.n., 2015.

4. **Washington Department of Transportation.** *WSDOT Highway Construction Costs: Hot Mix Asphalt.* Olympia : s.n., 2016.

5. **Xu, K.** *A Stochastic Three-Dimensional Cost Estimation System for Hot Mix Asphalt in the State of Alabama.* Auburn : Auburn University, 2018.

6. **Lindquist, Peter.** *Role of the Maritime Transportation System in Ohio's Economy.* Toledo : CPCS Transcom, 2018.

7. **Boddy, R and Smith, G.** *Statistical Analysis Methods in Practice for Scientists and Technologists.* West Sussex : John Wiley & Sons, 2009.

8. **Sullivan, E.** *PCA Outlook: 2015 - 2019, Asphalt and Oil Price Correlation.* New York City : Portland Cement Association, 2014.

9. **Attanasi, E D.** *Volatility of Bitumen Prices and Implications for the Industry.* Springer : National Resources Research, 2008.

10. **Crossett, J and Hines, L.** *Comparing State DOTs' Construction Project Cost and Schedule Performance - 28 Best Practices from 9 States.* Washington D.C : American Association of State Highway and Transportation Officials, 2007.

11. **Cantarelli, C C, et al.** *Cost Overruns in Large-Scale Transportation Infrastructure Projects: Explanations and Their Theoretical Embeddedness.* Jaffalaan : European Journal of Transport and Infrastructure Research, 2010.

12. **Damnjanovic, I and Zhou, X.** *Impact of Crude Oil Market Behavior on Unit Bid Prices: The Evidence from the Highway Construction Sector.* College Station : Construction Management and Economics, 2009.

13. **Project Management Institute.** *A Guide to the Project Management Body of Knowledge.* Newtown Square : s.n., 2017.

14. **American Association of State Highway and Transportation Officials.** *Practical Guide to Cost Estimating.* Washington D.C. : s.n., 2013.

15. **Shane, J S.** *Risk-Based Engineers Estimates.* Ames : Iowa State University, 2015.

16. **Pakalapati, K C.** *Data Usage Optimization for Cost Estimating in Asphalt Paving Projects Using a Cost Indexing System.* Auburn : Auburn University, 2018.

17. **Anderson, S D, Molenarr, K R and Schexnayder, C J.** *Guidance for Cost Estimation and Management for Highway Projects during Planning, Programming, and Preconstruction.* Washington D.C. : Transportation Research Board, 2007.

18. **Buhl, S, Flyvbjerg, B and Holm, M S.** *Underestimating Costs in Public Works Projects: Errors or Lies?* Boston : Journal of the American Planning Association, 2002.

19. **Huang, Y H.** *Pavement Analysis and Design.* Upper Saddle River : Pearson Education, 2004.

20. **Canfield, R.** *Asphalt/Binder Update.* Kansas City : Missouri Asphalt Pavement Association, 2015.

21. **U.S Energy Administration .** *Changing Quality Mix is Affecting Crude Oil Price Differentials and Refining Decisions.* Washington D.C : s.n., 2017.

22. **Tipton Corporation.** *Oil Refining and Petrochemical Basics.* Toyoda : s.n., 2017.

23. **Indian Institute of Technology Bombay.** *Conventional Asphalt Mix Design.* Bombay : Department of Civil Engineering, 2017.

24. **Diesel Technology Forum.** *Diesel Powers the U.S Economy.* Sacremento : Aspen Environmental Group, 2011.

25. **Xu, Y J.** *Determination of Various Operation Parameters for the Asphalt Mix Pavers.* s.l. : Engineering Machine Repair, 2003.

26. **Cheng, G and Wilmot, C G.** *Estimating Future Highway Construction Costss.* Baton Rouge : Journal of Construction Engineering and Management, 2003.

27. **Tsui, K K.** *More Oil, Less Democracy: Evidence from Worldwide Crude Oil Discoveries.* Oxford : The Economic Journal, 2011.

28. **Formier, P.** *What's Driving Asphalt Prices Up?* s.l. : Associated Construction Publications, 2008.

29. **Alabama Department of Transportation.** *ALDOT Asphalt Index Summary.* Montgomery : s.n., 2006 - 2018.

30. —. *ALDOT Fuel Index Summary.* Montgomery : s.n., 2006 - 2018.

31. —. *ALDOT Standard Specifications for Highway Construction.* Montgomery : s.n., 2012.

32. **Energy Intelligence.** *International Crude Oil Market Handbook.* New York City : s.n., 2010.

33. **Bacon, R and Tordo, S.** *Crude Oil Price Differentials and Differences in Oil Qualities: A Statistical Analysis.* s.l. : Energy Sector Management Assistance Program, 2005.

34. **U.S. Energy Information Administration.** *Price Differences Among Various Types of Crude OIl.* Washington D.C : s.n., 2006.

35. **U.S Energy Information Administration.** *Petroleum and Other Liquids.* Washington D.C : s.n., 2017.

36. **Ashuri, B, Illbeigi, M and Shayegh, S.** *A Data Mining Approach for Analyzing the Effects of Price Adjustment Clauses on Bid Prices Considering Material Quantity and Project Size.* Washington D.C : Journal of Construction Engineering and Management, 2015.

37. **Thompson, J.** *Development of Statistical Approaches to the Handling and Analysing of Large Occupational Data Sets.* s.l. : Health and Safety Executive, 2005.

38. **Lerch, L, et al.** *Measuring Surgeon Performance of Sentinel Lymph Node Biopsy in Breast Cancer Treatment by Cumulative Sum Analysis.* Spokane : American Journal of Surgery, 2006.

39. **Maguire, T, et al.** *Analysis of the Surgical Learning Curve Using the Cumulative Sum Method.* s.l. : Neurology and Urodynamics, 2013.

40. **Jalil, A and Mahmud, S.** *Environmental Kuznets Curve for CO2 Emissions: A Cointegration Analysis for China.* s.l. : Energy Policy, 2009.

41. **Acaravci, A and Ozturk, I.** *On the Relationship between Energy Consumption, CO2 Emissions, and Economic Growth in Europe.* s.l. : Energy, 2010.

42. **Kohsti, V.** *Cumulative Sum Control Chart.* s.l. : International Journal of Physics and Mathematical Sciences, 2011.

43. **Page, E.S.** *Continuous Inspection Schemes.* Oxford : Oxford University Press, 1954.

44. **Guo, R and Galeboe, K.** *Cumulative Sum Quality Control Charts Design and Application.* Cape Town : University of Cape Town, 2006.

45. **Goldsmith, P.L. and R.H., Woodward.** *Mathematical and Statistical Techniques for Industry: Monograph No. 3: Cumulative Sum Techniques.* s.l. : Oliver and Boyd, 1964.

46. **Rueda, J.** *Outlier.* Auburn : Auburn University, 2017.

47. **Seo, S.** *A Review and Comparison or Methods for Detecting Outliers in Univariate Data Sets.* Pittsburgh : University of Pittsburgh, 2006.

48. **Kohavi, R.** *A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection.* Stanford : Standford University, 1995.

49. **Rueda, J.** *Cross Validation.* Auburn : Auburn University, 2017.

APPENDIX A: GOLDSMITH AND WOODWARDS CUSUM TEST VALUES

	Significance Level		
Sample Size	**0.10**	**0.05**	**0.01**
5	2.40	2.70	3.30
6	2.70	3.00	3.60
7	2.90	3.20	4.00
8	3.20	3.50	4.30
9	3.40	3.70	4.60
10	3.60	3.90	4.90
11	3.80	4.10	5.10
12	3.90	4.30	5.30
13	4.00	4.50	5.50
14	4.10	4.60	5.60
15	4.20	4.80	5.80
20	5.20	5.60	6.80
25	5.60	6.00	7.30
30	6.20	6.70	8.00
40	7.20	7.80	9.30
50	8.00	8.60	10.40
60	8.80	9.50	11.30
70	9.50	10.30	12.20
80	10.10	10.80	12.90
90	10.50	11.30	13.60
100	11.00	11.80	14.30