Risk Assessment Tool

(Associated with cost)

Research Draft Report from East Tennessee State University | K. Joseph Shrestha, Mohammad Moin Uddin, Pritom Paul, and Jacob Fielden | March 31, 2025

Sponsored by Tennessee Department of Transportation Long Range Planning Research Office & Federal Highway Administration

DISCLAIMER

This research was funded through the State Planning and Research (SPR) Program by the Tennessee Department of Transportation and the Federal Highway Administration under ***RES #2023-02: Risk Assessment Tool (Associated with Cost).***

This document is disseminated under the sponsorship of the Tennessee Department of Transportation and the United States Department of Transportation in the interest of information exchange. The State of Tennessee and the United States Government assume no liability of its contents or use thereof.

The contents of this report reflect the views of the author(s) who are solely responsible for the facts and accuracy of the material presented. The contents do not necessarily reflect the official views of the Tennessee Department of Transportation or the United States Department of Transportation.

Technical Report Documentation Page

1. Report No. RES2023-02	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle ***Risk Assessment Tool (Associated with Cost)***		5. Report Date March 2025
		6. Performing Organization Code
7. Author(s) K. Joseph Shrestha, Mohammad Moin Uddin, Pritom Paul, and Jacob Fielden		8. Performing Organization Report No.
9. Performing Organization Name and Address East Tennessee State University 1276 Gilbreath Dr., Box 70552 Johnson City, TN 37614-1700		10. Work Unit No. (TRAIS)
		11. Contract or Grant No.
12. Sponsoring Agency Name and Address Tennessee Department of Transportation 505 Deaderick Street, Suite 900 Nashville, TN 37243		13. Type of Report and Period Covered Draft Report July 2023 – June 2025
		14. Sponsoring Agency Code
15. Supplementary Notes Conducted in cooperation with the U.S. Department of Transportation, Federal Highway Administration.		

16. Abstract

Computing accurate cost estimates during an early phase of project development is a challenging task because of various project factors such as change in project scope, inflation, site conditions, and market conditions. This study collects and analyzes engineers' estimates throughout the project development to quantify the changes in estimates over time. The findings were used to develop a framework and tool for Preliminary Cost Estimate Adjustment (PCEA) tool. The tool enables TDOT engineers to adjust an initial engineer's estimate to make it more aligned with the likely final engineer's estimate. The use of this tool showed a promising result with the reduction of errors by 37% – 55% in validation dataset. After developing the PCEA tool, TDOT identified a need to develop another tool to produce preliminary estimates using a limited set of project characteristics that are available at the early phase of project development. This led to the development of Construction Project Cost Estimation (CPCE) tool. This data-driven tool enables TDOT engineers to calculate preliminary cost estimates quickly with acceptable error rates based on AASHTO guidelines. These two tools are expected to aid TDOT engineers in increasing the accuracy of their engineers estimates, which will enable better planning of future projects based on available funding.

17. Key Words	18. Distribution Statement
PROJECT-RISK, RISK-ASSESSMENT, RISK-MANAGEMENT, COST-ADJUSTMENT, COST-PREDICTION, COST-ESTIMATION,	No restriction. This document is available to the public from the sponsoring agency at the website http://www.tn.gov/.

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	49	

Acknowledgement

This study is funded by the Tennessee Department of Transportation (TDOT). The authors would like to thank TDOT and TDOT employees for their support for the study. Specifically, the authors would like to thank Mr. Brian Hurst for leading and coordinating the TDOT research panel. The authors would like to acknowledge valuable inputs from Mr. James D. Kelly, Mr. David A. Duncan, and Mr. Craig Parnham from TDOT for their inputs throughout the study. The authors would also like to thank Ms. Stacy Carter and Ms. Beth Jirik from TDOT for their administrative support for the study.

Executive Summary

The final engineers' estimates for construction projects are often different than the earliest engineers' estimates because of various factors, such as, changes in project scope, inflation, site conditions, and market conditions. The change in the estimates over time can have a significant impact on future project planning and execution. For example, sufficient budget may not be available when the estimates for planned project increases, and not enough projects may be in the pipeline for construction if the estimate for planned project decreases during the project development. To reduce the potential for such disruption, the DOTs strive to produce more reliable estimates at the early phase of the project development. However, current methodologies and practices at Tennessee Department of Transportation (TDOT) have resulted in over 45% cost growth from the initial to the final engineers' estimates. A part of this error may be attributed to the lack of systematic methodologies for cost estimates and heavy reliance on the subjective inputs from engineers. This error can potentially be decreased if a proper data-driven framework and tool is developed by analyzing historical cost estimates.

This study collects and analyzes historical cost estimation dataset to quantify the project cost changes over time based on various factors, such as, location of the project, work type, program type, and planned construction year. A framework and a tool are developed based on the findings to enable TDOT engineers to adjust initial cost estimates to make it more aligned with likely final engineers' estimates. This tool entitled Preliminary Cost Estimate Adjustment (PCEA) Tool can adjust the estimates in with three level of details: a) program level, b) project level, and c) item level estimates. The program level adjustment is the most basic form of adjustment that adjusts estimates based on the preliminary estimate, preliminary estimate year, planned construction year, and inflation. Project level adjustment utilizes additional project characteristics, such as, work type, program type, and location. In item level estimate adjustment, the adjustments are made for various bid items instead of the overall project costs.

As the study progressed, TDOT realized the importance of developing a systematic method to compute preliminary cost estimates during early phase of the project development. Subsequently, the research project was amended to include additional scope for developing preliminary cost estimating models. To accomplish this additional scope, another tool entitled Construction Project Cost Estimation (CPCE) tool was developed. This tool relies on project characteristics, such as, type of work, planned construction year, route type, length, Right of Way (ROW) cost, and inflation rate to compute estimates based on historical data. The tool consists of over 630 models to compute estimates for various scenarios, and estimates produced based on county, region, and state level data are aggregated to produce the estimate. This tool enables TDOT engineer to quickly and accurately compute estimates based on limited set of project characteristics.

Key Findings

The key findings of this study are:

- When all historical data provided by TDOT is analyzed, 28 bid items were responsible for more than 80% of the total cost. This highlights the potential of improving the accuracy

of project cost estimates if the quantity and unit price of these high impact items can be calculated accurately.

- TDOT's final engineers' estimate is more than 45% higher than initial engineers' estimate on average.
- There is a lack of documentation about potential reasons behind the cost growth. While the actual reason, such as change in scope, change in market conditions, etc. are not documented, the resulting impacts can be analyzed as addition, removal, and change in the bid items. The change in bid item consists of change in quantity and unit price of the bid items.
- When TDOT estimators did not know an exact variant of an item (say a culvert), its cost was not added in the estimate. This was one of the key reasons for significant changes in cost over time. Later, TDOT corrected the course by starting to add cost of likely variant of the item.
- TDOT does not have in-house highway construction cost indexes and relies on national consumer inflation factor as an inflation factor for cost estimating.

PCEA Tool

The study evaluated and quantified the cost growth over time by various project characteristics, and the PCEA tool was developed to adjust the preliminary estimate to be more reflective of the final engineers' estimate. Three level of cost adjustments can be performed depending on the level of information available: a) program level simple, b) project level, and c) item level. For program level, a single adjustment factor is applied for all projects. For project level, five adjustment factors are applied separately, and the average of all the resulting estimates are provided as the final estimate. These five adjustment factors are based on different project characteristics: a) overall, b) work type, c) program type, d) county, and e) region. For the item level adjustment, 782 adjustment factors for quantity, unit price, and extended totals are developed. Overall, the tool was able to reduce the error by 37% to 55% in the validation data provided by TDOT.

CPCE Tool

The CPCE tool was developed to compute preliminary construction cost estimates using a limited set of project characteristics that are available at the early phase of project development. This enables TDOT engineers to compute estimates quickly with significantly limited effort compared to the current estimation practice. The tool consists of 290 county level models, 205 region level models, and 135 state level models. First, multiple county-, region-, and state-level models are evaluated as relevant. The values from these models are aggregated using Relative Reliability Indexes (RRIs) for county, region, and state levels. The estimates in the three geographic levels are then aggregated to produce the final value using weighted average method. The CPCE tool was able to produce estimates for several work types that are within the error range as prescribed by AASHTO guidelines for planning phase.

Key Recommendations

The key recommendations of the study are listed below.

- TDOT should develop and maintain systematic database of project characteristics that have significant impact on the project cost.
- TDOT should develop specialized models for various project types, such as bridge replacement using the project type specific characteristics.
- TDOT should develop a structured cost estimates history database that allows tracking of changes in the estimates over time.
- TDOT should document factors responsible for changes in engineers' estimates over time. These factors may be recorded as overall factor and item specific factors anytime the estimates are updated. Such factors collected over time will enable TDOT to have a better understanding of actual reasons behind the changes in the estimates.
- The inflation of highway construction industry is reflected by a highway construction cost index. The consumer price index may not properly represent the inflation in the highway construction industry. A national highway construction cost index may not represent the market condition for the state. As such, TDOT specific highway construction index should be developed to adjust the estimates more accurately for the inflation.
- The tools and models developed in this study should be continuously updated so that the results produced from the tool is reflective of recent project data and market conditions.

The tools should not be considered as a substitute for engineering judgement. If newer projects are significantly different than historical projects, the model developed from historical projects will be less reliable for newer projects. If newer projects are similar to historical projects, the tool will likely produce more reliable results.

Table of Contents

List of Tables

List of Figures

Glossary of Key Terms and Acronyms

AASHTO	American Association of State Highway and Transportation Officials
CoV	Coefficient of Variance
CPCE	Construction Project Cost Estimation
DOT	Department of Transportation
FHWA	Federal Highway Administration
HCCI	Highway Construction Cost Index
MAPE	Mean Absolute Percentage Error
MPE	Mean Percentage Error
NCHRP	National Cooperative Highway Research Program
NHCCI	National Highway Construction Cost Index
PCEA	Preliminary Cost Estimate Adjustment
RI	Reliability Index
RRI	Relative Reliability Index
TDOT	Tennessee Department of Transportation

Chapter 1 Introduction

Construction project estimates often increase throughout the project development phases because of various internal and external factors, such as, scope changes, inflation, local competition, site conditions, and project complexities (AASHTO, 2013; Schexnayder et al., 2009; Washington State Department of Transportation, 2014). The Tennessee Department of Transportation (TDOT) has also experienced a significant increase in cost estimates from an initial engineers' estimate to the final engineers' estimate for the same project. While such changes in the estimates may be necessary to account for the changing project definitions, it inherently becomes an issue for long term planning and budgeting, which requires accurate estimates for various construction projects in the pipeline.

Several state Departments of Transportation (DOTs) have created various risk register templates to identify and evaluate project risks during planning and construction to partially address the issue of project cost changes over time (Ashuri et al., 2015; California Department of Transportation, 2012; New York State Department of Transportation, 2009; Texas Department of Transportation, 2015; Washington State Department of Transportation, 2014). While such risk register templates can be a powerful tool, they require significant inputs from the engineers to identify and evaluate various risks and their potential impacts on the project cost. However, with the shrinking size of public agencies and consequent need to accomplish more tasks by the same or smaller group of engineers, there is a need to develop an automation tool for assessing project risks associated with cost. This situation necessitates an alternative method to quickly and reliably assess the project risks associated with cost.

State DOTs, including TDOT, have historical cost estimate dataset that tracks the changes in engineers' estimate over time. Such dataset can be analyzed to identify and quantify historical trend in the cost changes over the various phases of project development using Artificial Intelligence (AI). The insights obtained from the historical trend can then be applied to the new project estimates to account for project risks associated with cost. Further, AI can also be used to produce preliminary estimates at early phase of project development with minimal project information based on the knowledge generated from the historical cost estimate dataset.

1.1 Research Objectives

The overall goal of this study is to develop a framework and tool to improve current cost estimation practices for planning and budgeting purposes at TDOT. The specific objectives of the study are:

- Review state DOT practices on quantifying and applying project risks associated with cost
- Develop a framework and tool to automatically adjust initial engineers' estimates to make it more reflective of the likely final engineers' estimates
- Develop a framework and tool to automatically compute preliminary cost estimates using limited set of project characteristics available at early phase of project development

1.2 Significance of the Research

In the strategic plan for the fiscal years 2022 - 2026, the U.S. DOT has highlighted the need to "develop and manage data systems and tools to provide objective, reliable, timely, and accessible data to support decision-making, transparency, and accountability" (US Department of Transportation, 2022). This study takes a step towards making TDOT's cost estimating practices more data-driven, more objective, and more reliable. It collects and analyzes more than a decade of existing estimating datasets from TDOT and applies the knowledge extracted from the datasets to increase the accuracy of engineers' estimate by accounting for various project risks associated with cost. Instead of developing theoretical framework only, it develops two tools that can be quickly implemented and integrated into existing TDOT estimating practices. Overall, these tools are expected to aid TDOT engineers' in improving the planning, budgeting, and execution of infrastructure projects.

1.3 Organization of the Report

This remaining part of the report is organized into six chapters. *Chapter 2 Literature Review* summarizes existing practices of addressing project costs risks at various DOTs. *Chapter 3 Methodology* details various tasks completed to achieve the objectives of the study. *Chapter 4 Framework for Preliminary Cost Estimate Adjustment (PCEA)* describes the theoretical framework developed to adjust preliminary engineers' estimates to make it more reflective of the likely final engineers' estimate. *Chapter 5 Framework for Construction Project Cost Estimation (CPCE)* develops another framework to quickly and reliably estimate the preliminary costs of new projects based on historical estimate datasets. *Chapter 6 Results and Discussion* presents the summary of data analysis as well as an overview of the two tools and their validation. The report ends with the *Chapter 7 Conclusion and Recommendation* that highlights the major findings and recommendations for continuous improvements of the cost estimating practices at TDOT.

Chapter 2 Literature Review

The literature review is divided into four sections: a) overview of risk-based cost estimating, b) project risk factors used by state DOTs, c) risk-based cost estimating practices at various state DOTs, and d) parametric cost estimating for highway projects.

2.1 Overview of Risk-Based Cost Estimating

The Project Management Body of Knowledge (PMBOK®) defines risk as "an uncertain event or condition that, if it occurs, has a positive or negative effect on one or more project objectives" (Project Management Institute, 2017). Past studies have increasingly recognized the value of probabilistic and risk-based cost estimating to account for the uncertainties associated with the construction projects (Molenaar, 2005). These studies have adopted variety of approaches and tools, such as risk register, Monte Carlo simulations, and sensitivity analysis (Molenaar, 2005; Sadeh et al., 2021). Researchers have identified major factors associated with the cost overruns, such as scope creep and inflation (Anderson et al., 2006; Shane et al., 2009).

2.2 Project Risk Factors Used by State DOTs

Many state DOTs have developed their own risk management tools and templates. **Table 2-1** summarizes the various project risk factors considered by various state DOTs. The table shows that environmental, ROW, and geotechnical / drainage are the top four risk factors considered by most of the DOTs. Next three top risk factors considered by DOTs are traffic, design, and third-party agreements.

TABLE 2-1

PROJECT RISK FACTORS USED BY STATE DOTS

DOT	Environmental	ROW	Geotechnical / Drainage	Utilities	Traffic	Design	3rd Party Agreements	Constructability	Public Perception	Payment/ Funding	Scope	Specialty	Project Management	Contractor Availability
Missouri	X	X	X	X	X	X	X		X	X		X		
California	X	X	X	X	X									
Florida	X	X	X	X	X		X		X					
Michigan	X	X	X	X	X	X		X		X		X		
Nevada	X	X	X			X		X	X	X	X	X		
Tennessee	X	X	X	X	X	X	X						X	X
Georgia	X	X		X		X	X		X					

DOT	Environmental	ROW	Geotechnical / Drainage	Utilities	Traffic	Design	3rd Party Agreements	Constructability	Public Perception	Payment/ Funding	Scope	Specialty	Project Management	Contractor Availability
Alabama	X	X	X	X			X	X		X	X			X
Mississippi	X	X	X	X	X				X	X	X			
Virginia	X	X	X	X		X	X	X	X		X		X	
North Carolina	X	X	X	X	X			X						
Kentucky	X	X	X	X	X	X	X				X	X		
Arkansas	X		X		X	X		X		X		X	X	X
South Carolina	X	X	X	X	X					X	X		X	
Federal	X	X	X			X	X	X	X	X		X	X	X

2.3 *Risk-Based Cost Estimation Practices at Various State DOTs*

This section provides an overview of the diverse risk-based cost estimation practices adopted by several state DOTS: Minnesota DOT (MnDOT), Virginia DOT (VDOT), Washington State DOT (WSDOT), Georgia DOT (GDOT), Nevada DOT (NDOT), and New York State DOT (NYSDOT). It focuses on their use of qualitative and quantitative tools, stakeholder engagement, and lifecycle-specific strategies.

Minnesota Department of Transportation (MnDOT)

MnDOT employs a structured risk-based cost estimation approach that integrates qualitative and quantitative methodologies, tailored to the complexity and development stage of each project. For simpler projects, MnDOT relies on predetermined percentage contingencies, while more complex initiatives leverage advanced tools such as risk registers, Monte Carlo simulations, and probability-impact matrices. The agency places significant emphasis on risk registers to document potential risk events and their impacts, facilitating effective tracking and management. For high-value or high-complexity projects, Monte Carlo simulations are utilized to quantify cost uncertainty, ensuring a robust estimation process. MnDOT's commitment to continuous risk reassessment, proactive management, and structured stakeholder communication enhances transparency and consistency in its cost estimation practices.

Virginia Department of Transportation (VDOT)

VDOT adopts a phased, structured approach to risk-based cost estimation, integrating qualitative and quantitative analyses based on project development phases and complexity. Following the AASHTO Risk-Based Estimating process, VDOT calculates contingencies through initial qualitative assessments or, for complex projects, detailed quantitative analyses like Monte Carlo simulations. The agency emphasizes clear documentation of assumptions, risks, and estimates via a well-defined Cost Estimate Package (CEP), supported by tools such as the Cost Estimate Workbook, Project Cost Estimating System (PCES), AASHTOWare Project Estimation, and state-developed databases. This comprehensive approach ensures consistent estimate management throughout the project lifecycle, fostering effective communication and transparency with stakeholders.

Washington State Department of Transportation (WSDOT)

WSDOT implements a phased risk-based cost estimation strategy that evolves with project stages, incorporating qualitative and quantitative methods tailored to complexity. Early planning and scoping phases utilize parametric and analogous techniques, transitioning to detailed cost-based estimating during design and contract stages, with methods like Monte Carlo simulations applied as needed. WSDOT prioritizes thorough documentation of assumptions, milestone reviews, and robust stakeholder communication, securing management endorsement to ensure estimate reliability. The agency's spreadsheet-based tool and manual provide advanced statistical measures, though their complexity poses a challenge for widespread adoption.

Georgia Department of Transportation (GDOT)

GDOT integrates risk mitigation strategies throughout the project lifecycle to perform risk-based cost estimation, emphasizing early risk identification through collaboration with stakeholders like environmental specialists and utility companies. This proactive approach involves defining assumptions, conducting regular constructability reviews, initiating early geotechnical investigations, and updating cost estimates dynamically. GDOT employs tools such as Constructability Reviews and Value Engineering (VE) studies to reduce uncertainties, with ongoing assessment and adjustment processes ensuring adaptability to changing project conditions, ultimately enhancing the accuracy of cost forecasts.

Nevada Department of Transportation (NDOT)

NDOT employs a scalable risk-based cost estimation approach, combining qualitative and quantitative analyses based on project size and complexity. For projects exceeding $100 million, NDOT uses Monte Carlo simulations to develop probabilistic cost and schedule forecasts, supported by detailed risk registers that quantify threats and opportunities with assigned probabilities and impacts. Smaller projects with a budget of $25M to $100M may use qualitative risk analysis with percentage-based risk allowances (e.g., 3%-15%, p. 92), adjusted by project phase and risk impact. This method incorporates uncertainties into project-specific contingencies, moving away from traditional lump-sum values, and emphasizes ongoing risk monitoring and stakeholder communication, promoting informed decision-making and proactive risk management.

New York State Department of Transportation (NYSDOT)

NYSDOT implements a structured risk management process for cost estimation, encompassing risk identification, qualitative and quantitative analysis, and continuous monitoring. Initial qualitative assessments classify risks using a standardized matrix (low, moderate, high) based on probability and impact. At the same time, complex projects employ Monte Carlo simulations, tornado diagrams, and decision tree analyses to establish realistic contingencies. This data-driven approach numerically estimates cost impacts, with NYSDOT's emphasis on contingency planning and stakeholder engagement enhancing its ability to anticipate and mitigate risks, ensuring resilient project budgets.

2.4 Parametric Cost Estimating for Highway Projects

At early phase of project development, parametric cost estimation models can be used to estimate project costs using a limited set of project characteristics. Many studies have developed parametric cost estimating models for estimating costs of highway construction, bridge construction, as well as preconstruction services cost (Duverlie & Castelain, 1999; Gardner et al., 2017; Gransberg et al., 2016; Kim & Hong, 2012; Liu et al., 2011). These studies have identified various project characteristics that affect project costs, including length, highway classification, ROW cost, width, bridge superstructure type, ROW constraints, project complexity, and site specific challenges (Piratla et al., 2024; Saito et al., 1991; Trost & Oberlender, 2003). Multiple Linear Regression (MLR) is often selected as a preferred modeling technique for such studies because of its simplicity, transparency, and ability to provide a reasonable performance compared to more complicated models (Piratla et al., 2024). Once MLR models are developed using specialized tools, the results can be implemented as a spreadsheet tool, which reduces the potential friction for integrating and implementing the models into existing estimating practices. These models have achieved varying level of accuracies. For example, Piratla et al. (Piratla et al., 2024) achieved accuracies ranging from 61% to 84% for various project types.

Chapter 3 Methodology

The overall methodology of the research consisted of four major tasks: a) literature review, b) collection and analysis of historical estimation datasets, c) development of theoretical frameworks for Preliminary Cost Estimate Adjustment (PCEA) and Construction Project Cost Estimation (CPCE), and d) development of PCEA and CPCE tool.

3.1 Literature review

The research team reviewed existing literature on current DOT practices for assessing project risks associated with cost and parametric cost estimating models to learn from existing body of knowledge and shape the direction of this study.

3.2 Collection and Analysis of Historical Estimation Datasets

Historical engineers' estimate dataset from 2013 to 2024 was collected as spreadsheet and PDF documents. These documents contained unstructured to semi-structured datasets. Because of the nature of these documents, extensive manual work was performed to prepare clean and structured dataset that can be used for further analysis. The projects with incomplete data attributes were removed from the dataset. Data from bundled projects were removed for systematic analysis. Data from partial years were removed, and data from 2024 was separated for validation. The clean and structure dataset was then analyzed to get insights about changes in the engineers' estimate over time, relationship between various project characteristics and increase in engineers' estimates, and relationship between the engineers' estimates and various project characteristics.

3.3 Development of Theoretical Frameworks for PCEA and CPCE

The knowledge acquired from the data analysis was used to develop a theoretical framework for PCEA and PCEA. The core of the PCEA framework is based on the insights about the changes in engineers' estimate over time based on various project characteristics, such as location, work type, and program type. The core of CPCE framework is based on modeling the impact of various project characteristics – such as, project length, location, route type, type of work, and Right of Way (ROW) cost – on the final engineers' estimate. The study introduces a new concept of Relative Reliability Index (RRI) and Reliability Index (RI) as indicators of the reliabilities of various models and factors computed for the frameworks.

3.4 Development of PCEA and CPCE Tools

The theoretical framework for PCEA and CPCE was implemented as two spreadsheet-based tools. The tool was validated with set of data that was not used for developing the models.

Chapter 4 Framework for Preliminary Cost Estimate Adjustment (PCEA)

This chapter presents a theoretical framework for Preliminary Cost Estimate Adjustment (PCEA) using historical data. The framework enables estimators to adjust initial engineers' estimates to be more reflective of the likely final engineers' estimates using various adjustment factors based on project characteristics and bid items. The adjustments can be applied at three different levels of granularity: 1) Program Level, 2) Project Level, and 3) Item Level.

4.1 Program Level Adjustment

An overview of the program level adjustment is illustrated in Figure 4-1. This simplified approach does not consider any project characteristics and hence utilizes a broader state level cost variability trend based on the entire project level dataset.

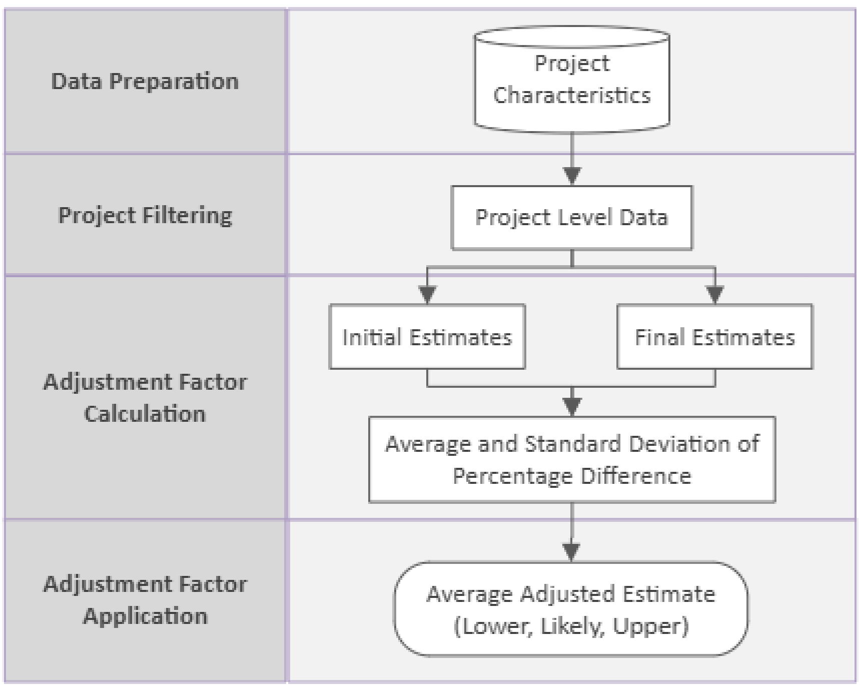

Figure 4-1 Framework of PCEA – Program Level Adjustment

The initial and final engineers' estimates are generally prepared in different year. As such, part of the cost growth is a result of inflation. This inflation related part needs to be removed before any further analysis. For that, the initial engineers' estimates are converted to the same year as the final engineers' estimates using an inflation rate between the initial and final estimate years.

The annual inflation rate between the initial and final engineers' estimate years is calculated using Highway Construction Cost Indexes (HCCIs) for the two years (Equation (1)).

(1)

$$\text{Inflation rate (i)} = \left(\left(\sqrt[N]{\frac{HCCI_f}{HCCI_i}}\right) - 1\right) \times 100\% \quad (1)$$

Where,

$HCCI_f$ = HCCI for Final Estimate Year
$HCCI_i$ = HCCI for Initial Estimate Year
N = Number of Years Between Initial and Final Estimates

The inflation adjusted initial engineers' estimate is calculated using Equation (2).

$$E_{ia} = E_{iu} * (1+i)^{Y_f - Y_p} \quad (2)$$

Where,

E_{iai} = Inflation Adjusted Initial Estimate
E_{iui} = Inflation Unadjusted Initial Estimate
i = Inflation Rate
Y_f = Final Estimate Year
Y_p = Initial Estimate Year

Now, the final engineers' estimate, and inflation adjusted initial engineers' estimate are representative of the dollar value for the same year. The estimate change percentage from the initial to final engineers' estimate for each project is calculated using Equation (3).

$$\textit{Estimate Change Percentage} = \frac{E_f - E_{iai}}{E_{ia}} \times 100\% \quad (3)$$

Where, E_f is final engineers' estimate.

Next, various statistics, including average and standard deviations, are calculated based on the cost change percentages for all projects. Finally, these statistics are used to adjust the initial engineers' estimate to be more representative of the final engineers' estimate. Further, instead of providing a single point estimate, it takes a probabilistic approach and providers a lower, likely, and upper estimates at various confidence levels.

As this simplified approach does not consider any project characteristics, it is more suitable for program level estimate adjustment for long-term planning when a group of projects are being considered all at once or if project characteristics are not available yet. Once project characteristics are available, project-level adjustment should be applied.

4.2 Project Level Adjustment

An overview of the project level adjustment is illustrated in Figure 4-2. The core of this method is same as the program level adjustment. However, in addition to generating statistics from all projects at once, projects are also segregated into various groups based on project characteristics, and statistics are generated for each group as well. For example, projects from specific county and specific region have separate sets of statistics, so does projects of specific work type and program type. To adjust the estimate of a new project, multiple set of adjustments are applied based on its project characteristics. Finally, average values of the lower, likely, and upper estimates are calculated as the final output of this method. Since this method considers various project characteristics, this will provide more reliable adjustments for preliminary cost estimates.

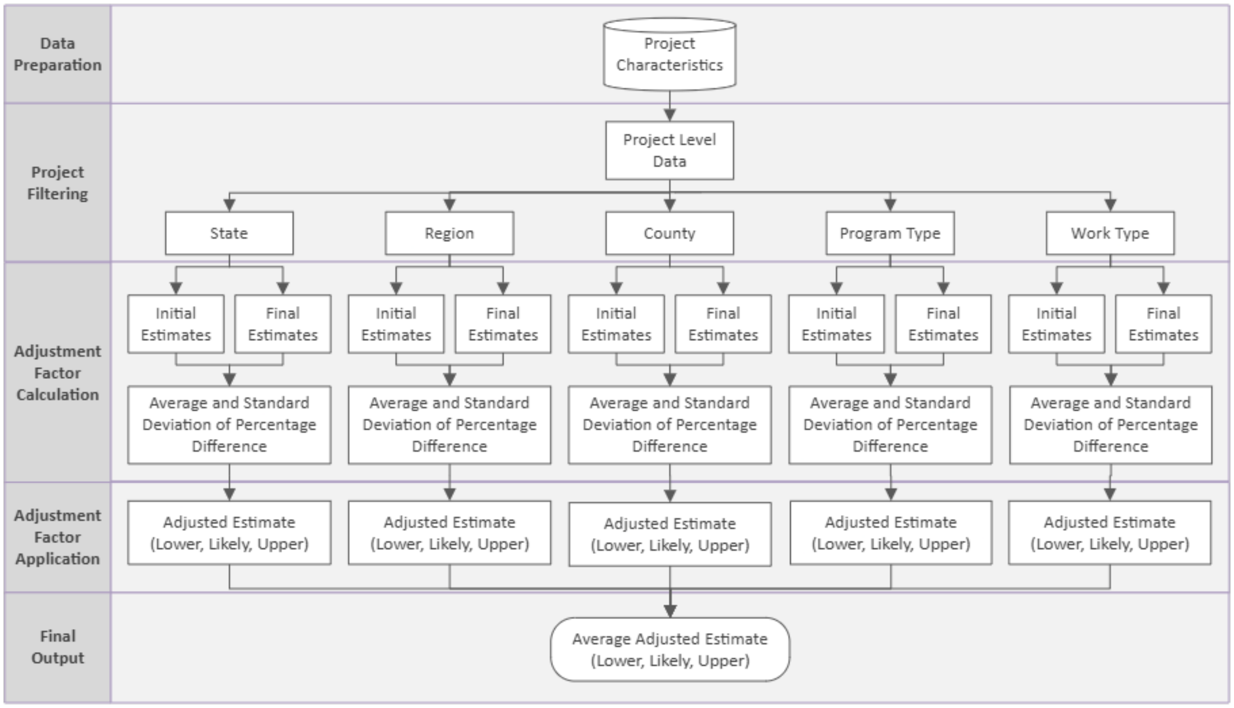

Figure 4-2: Framework of PCEA – Project Level Adjustment

4.3 Item Level Adjustment

An overview of the item level adjustment is illustrated in Figure 4-3. This method can be used only if all bid item information (quantity and unit price) is already available for the new project. In this method, change percentages are calculated for all bid items from historical data. Further, instead of relying on a single cost change metrics, it computes changes in quantity, unit price, and extended amount for each bid item.

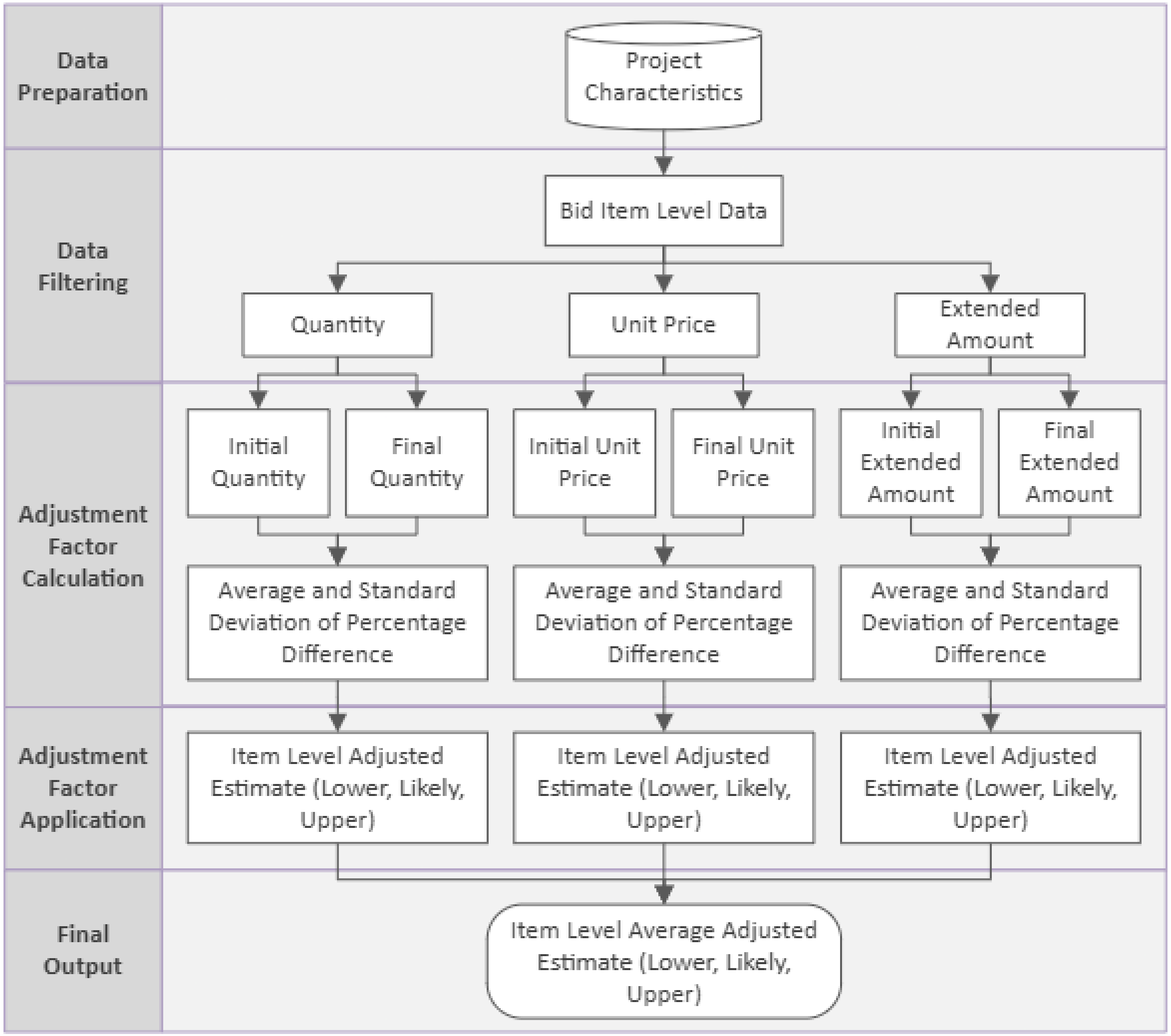

Figure 4-3: Framework of PCEA – Item Level Adjustment

Relative Reliability Index (RRI)

Several statistical measures of these change percentages are calculated for each bid item. These statistics include average, standard deviation, and Coefficient of Variance (CoV). A new metrics entitled RRI is developed in this study to indicate the reliability of the adjustments. Mathematically, the basic form of RRI can be defined as an inverse of CoV as shown in Equation (4).

$$RRI = 1/CoV \tag{3}$$

While the CoV quantifies variability in the data, the RRI quantifies the consistency in the data. Higher the RRI value, better is the likelihood of the accuracy of the predictions. As there are three separate CoVs – one each for quantity, unit price, and extended amount, the RRI for such scenario can be defined as an inverse of average CoVs, as shown in Equation (5).

$$RRI = 3/(\mathrm{CoV}_{\mathrm{quantity}} + \mathrm{CoV}_{\mathrm{unit\ price}} + \mathrm{CoV}_{\mathrm{extended\ amount}}) \quad (4)$$

For a new project, two set of estimates for each bid item are computed. First, the unit price and quantity adjustments are applied together to compute one set of lower, likely, and upper estimates. Second, the extended amount adjustment is applied to compute second set of lower, likely, and upper estimates. Average of the two set of lower, likely, and upper estimates are calculated as the final output for each bid item. Once the same process is applied to all bid items in the project, the final estimate for the project is calculated as the sum of the lower, likely, and upper values for all bid items.

Chapter 5 Framework for Construction Project Cost Estimation (CPCE)

This chapter presents a theoretical framework for Construction Project Cost Estimation (CPCE) using historical project information. The framework consists of five major components: 1) Data Preparation, 2) Project Filtering, 3) Model Development, 4) Multilevel Aggregation, 5) Inflation Adjustment, and 6) Final Output (Figure 1). The framework introduces a new concept of Relative Reliability Index (RRI) to aggregate multiple sets of adjusted estimates to a single set of adjusted estimates.

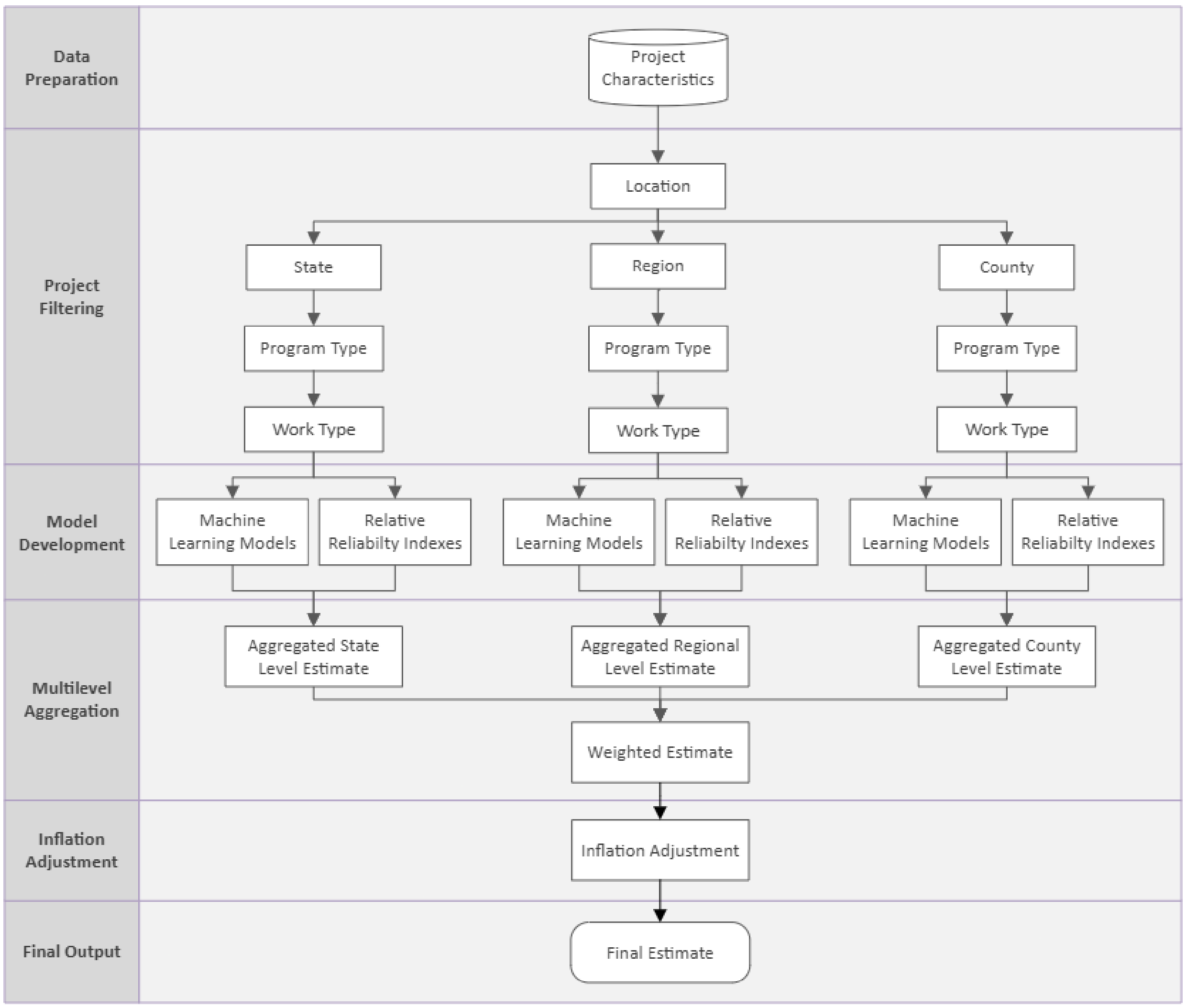

Figure 5-1: Framework of the Construction Project Cost Estimation (CPCE)

5.1 Data Preparation

This framework utilizes historical data and powerful machine learning technique to compute estimates for new construction projects. The required data attributes for the modeling includes program type, work type, project location, project length, route type, Right of Way (ROW) and final engineer's estimate. As the original dataset includes projects from multiple years, the engineers' estimates for various projects are prepared for different years. To ensure apples to apples comparison and analysis, these values are normalized to the same model year before developing machine learning models using Equation (6).

$$E_m = E_e * (1+i)^{Y_m - Y_e} \tag{5}$$

Where,

E_m	=	Estimate for model year
E_e	=	Estimate for engineers' estimate year
i	=	Inflation rate
Y_m	=	Model year
Y_e	=	Engineers' estimate year

5.2 Project Filtering

Past studies tend to focus on developing a single universal model suitable for all projects in various geographical regions and for various project types. This *one size fits* all solution can produce less reliable results as they are not able to address the important of various project characteristics in deeper levels. This study takes a different approach where multiple models are developed based on various project characteristics, such as program type, work type, and project location. These two approaches can be compared to a decision making by a single expert vs decision making based on the consensus of multiple experts in the room.

To develop multiple models, the dataset is filtered in multiple groups. For the first set of data, entire state level data is separated by program type and work type combination. For the second set of data, entire state data is first segregated by regions, then by the combination of program type and work type. For the third set of data, entire state data is first segregated by county, then by the combination of program type and work type.

5.3 Model Development

For each segregated dataset, multiple models are developed using various combination of project characteristics. Each of these sub-models will have its own strength and weaknesses in various scenarios. The development of separate models for each program type and work type ensures that future estimates are based on the historical estimate data for same program type and work type.

Reliability Index (RI)

For each model, Coefficient of Determination (R^2) and Reliability Indexes (RI) are calculated. The R^2 value is a traditional metrics that represents the portion of variance in the dependent variable (cost estimate in this case) explained by the independent variables (project characteristics in this case). A higher R^2 indicates better model performance.

The RI is a new metrics developed for this study to address the limitation of traditional R^2. The traditional R^2 value can be misleading when a very few data points are available to develop the model. Models developed from smaller number of data points are generally less reliable than the models developed from larger number of data points. Typically, 30 data points are generally considered large enough if sufficient number of independent variables are considered in the model. The RI can be defined mathematically by Equation (3).

$$Reliability\ Index\ (RI) = R^2 * Min\left(\frac{Data\ Point\ Count}{30}, 1\right) \tag{6}$$

5.4 Multilevel Aggregation

The outputs from various models are aggregated using weighted average methods in two phases: a) Aggregation within same geographic coverage and b) Aggregation of outputs for varying geographic coverages. Equation (8) shows the method used for the weighted average calculation for both phases. While same formula is used for the two phases, the weightages used in the two phases are different.

$$\text{Weighted Average Estimate} = \frac{\sum(\text{Estimate} \times \text{Weight})}{\sum\text{Weight}} \tag{7}$$

In the first phase, RIs corresponding to the models are used as the weights. The RRIs ensures that more reliable models get higher weightages. In the second phase, the three aggregated outputs from the state, region, and county levels are further aggregated using custom weights. The state level models have more data and can generally be expected to provide more reliable estimate for the state level market, but it ignores the regional and county level variabilities. The regional and county level models have less data and hence can be less reliable, but it also provides regional and county level market conditions. Aggregation of these state, region, and county level models provides a balanced result which is synonymous to the consensus of multiple experts in the room.

5.5 Inflation Adjustment

As the models are prepared using adjusted engineers' estimates for the model year, the estimates produced from the models are also representative of the market conditions for the model year. However, the actual construction cost for the actual construction year is more relevant and important value for planning and budgeting purposes. Thus, this initial estimates from the models need to be adjusted further for the planned construction year. This adjustment can be achieved using Equation (9).

$$E_c = E_m * (1+i)^{Y_p - Y_m} \tag{8}$$

where,

E_c = Estimated Cost of the Project for the Planned Construction Year
E_m = Estimated Cost of the Project for the Model Year
i = Inflation Rate
Y_p = Planned Construction Year
Y_m = Model Year

5.6 Final Output

The output after the inflation adjustment represents the final estimate produced by the framework. This value can be used for planning and budgeting purposes.

Chapter 6 Results and Discussion

The framework for Preliminary Cost Estimate Adjustment (PCEA) and Construction Project Cost Estimation (CPCE) were implemented as spreadsheet-based tools to enable TDOT to implement and integrate the findings of this study in existing cost estimating practices. This chapter presents a) an overview of the dataset used for developing the tools, b) an overview and validation of the PCEA tool, c) an overview and validation of CPCE tool, and d) Limitations and weaknesses of the study.

6.1 Overview of Dataset

For this study, a project dataset consisting of 263 bridge projects, 273 legislative projects, 284 safety projects, and 46 Economic Development projects from 2013 to 2024 were collected. The original dataset consisted of PDF files and Spreadsheets for each project at various phases of project development. Scattered data from these files were manually compiled into a dataset that can be used for further analysis. The data covering complete years were used for modeling, while the data covering partial year 2024 was used for validation purpose. A summary of the engineers' estimates for the dataset is provided in **Table 6-1**. The final dataset included various data attributes, such as, county, program type, type of work, project length, route type, ROW costs, and estimates.

TABLE 6-1

DESCRIPTIVE STATISTICS OF THE PROJECT COST

Program Type	Mean	Standard Deviation	Minimum	Maximum
Bridge	$6,876,136	$10,165,836	$265,228	$64,688,297
Safety	$1,031,756	$1,431,808	$15,053.3	$11,755,268
Legislative	$37,175,065	$32,859,456	$330,068	$180,550,835
Economic Development	$2,993,561	$2,494,421	$612,085	$6,059,895

Table 6-2 lists 28 the major bid items that covers over 80% of costs. This highlights the potential for increasing estimate accuracy by focusing on a limited number of significant cost items.

TABLE 6-2

TOP BID ITEMS BY EXTENDED AMOUNT COVERING 80% OF TOTAL COST

Item Number	Description	Unit	Item Extended Total Sum
303-01	MINERAL AGGREGATE, TYPE A BASE, GRADING D	TON	$60,926,204.12
203-01	ROAD & DRAINAGE EXCAVATION (UNCLASSIFIED)	C.Y.	$33,261,195.38
717-01	MOBILIZATION	LS	$27,364,825.00
203-03	BORROW EXCAVATION (UNCLASSIFIED)	C.Y.	$15,290,054.23

Item Number	Description	Unit	Item Extended Total Sum
307-02.01	ASPHALT CONCRETE MIX (PG70-22) (BPMB-HM) GRADING A	TON	$12,689,285.92
307-03.01	ASPHALT CONCRETE MIX (PG76-22) (BPMB-HM) GRADING A	TON	$10,596,599.45
307-01.08	ASPHALT CONCRETE MIX (PG64-22) (BPMB-HM) GRADING B-M2	TON	$9,802,460.21
411-02.10	ACS MIX(PG70-22) GRADING D	TON	$9,582,981.88
307-02.08	ASPHALT CONCRETE MIX (PG70-22) (BPMB-HM) GRADING B-M2	TON	$9,427,589.27
604-07.01	RETAINING WALL (DESCRIPTION)	S.F.	$9,287,111.90
201-01	CLEARING AND GRUBBING	LS	$9,210,500.00
712-02.02	INTERCONNECTED PORTABLE BARRIER RAIL	L.F.	$8,687,345.26
203-02.01	BORROW EXCAVATION (GRADED SOLID ROCK)	TON	$7,472,917.15
307-01.01	ASPHALT CONCRETE MIX (PG64-22) (BPMB-HM) GRADING A	TON	$7,141,313.28
501-01.03	PORTLAND CEMENT CONCRETE PAVEMENT (PLAIN) 10"	S.Y.	$6,643,974.00
411-01.10	ACS MIX(PG64-22) GRADING D	TON	$6,261,179.55
712-01	TRAFFIC CONTROL	LS	$6,150,125.00
607-03.02	18" CONCRETE PIPE CULVERT (CLASS III)	L.F.	$5,832,617.00
105-01	CONSTRUCTION STAKES, LINES AND GRADES	LS	$5,725,150.00
307-03.08	ASPHALT CONCRETE MIX (PG76-22) (BPMB-HM) GRADING B-M2	TON	$5,306,762.80
402-01	BITUMINOUS MATERIAL FOR PRIME COAT (PC)	TON	$5,246,522.63
702-03	CONCRETE COMBINED CURB & GUTTER	C.Y.	$4,752,401.25
604-07.02	RETAINING WALL (DESCRIPTION)	S.F.	$4,054,081.90
411-01.07	ACS MIX (PG64-22) GRADING E SHOULDER	TON	$3,487,870.69
701-01.01	CONCRETE SIDEWALK (4 ")	S.F.	$3,304,142.95
803-01	SODDING (NEW SOD)	S.Y.	$3,211,081.23
709-05.06	MACHINED RIP-RAP (CLASS A-1)	TON	$2,479,940.10
403-01	BITUMINOUS MATERIAL FOR TACK COAT (TC)	TON	$2,308,714.30

The increase or decrease in project costs will be reflected as an addition, removal, substitution, change in quantity, or change in unit price of bid items. The results of the top bid items that were added or removed from the initial engineers' estimate to the final engineers' estimate are identified and presented below. It was not possible to automatically identify the item substitution because of the nature of the dataset. The changes in quantity and unit price of bid items are not presented below, but the results are utilized in the PCEA tool development.

Table 6-3 lists top 20 frequently added bid items with highest impact on cost.

TABLE 6-3

FREQUENTLY ADDED BID ITEMS WITH HIGHEST IMPACT ON COST

Item Number	Description	Unit	Item Extended Total Sum	Item Added Frequency
303-01	MINERAL AGGREGATE, TYPE A BASE, GRADING D	TON	$49,630,927.45	115
203-01	ROAD & DRAINAGE EXCAVATION (UNCLASSIFIED)	C.Y.	$49,021,990.73	106
920-20.01	(DESCRIPTION)	S.F.	$37,519,260.00	11
920-20.05	(DESCRIPTION)	LS	$26,596,600.00	10
717-01	MOBILIZATION	LS	$22,299,096.00	123
307-02.01	ASPHALT CONCRETE MIX (PG70-22) (BPMB-HM) GRADING A	TON	$15,080,934.62	46
203-03	BORROW EXCAVATION (UNCLASSIFIED)	C.Y.	$13,241,460.75	43
307-01.01	ASPHALT CONCRETE MIX (PG64-22) (BPMB-HM) GRADING A	TON	$11,047,026.16	57
307-01.08	ASPHALT CONCRETE MIX (PG64-22) (BPMB-HM) GRADING B-M2	TON	$10,750,191.52	81
307-03.01	ASPHALT CONCRETE MIX (PG76-22) (BPMB-HM) GRADING A	TON	$10,676,353.50	9
411-01.10	ACS MIX(PG64-22) GRADING D	TON	$10,520,256.88	80
604-07.01	RETAINING WALL (DESCRIPTION)	S.F.	$10,188,458.00	21
307-02.08	ASPHALT CONCRETE MIX (PG70-22) (BPMB-HM) GRADING B-M2	TON	$9,735,085.02	47
201-01	CLEARING AND GRUBBING	LS	$9,476,918.94	99
411-02.10	ACS MIX(PG70-22) GRADING D	TON	$7,982,843.26	46
920-20.02	(DESCRIPTION)	S.Y.	$7,819,370.00	3
607-03.02	18" CONCRETE PIPE CULVERT (CLASS III)	L.F.	$7,193,882.04	56
105-01	CONSTRUCTION STAKES, LINES AND GRADES	LS	$6,534,838.20	104
712-02.02	INTERCONNECTED PORTABLE BARRIER RAIL	L.F.	$6,487,025.50	65
604-07.02	RETAINING WALL (DESCRIPTION)	S.F.	$6,193,790.00	10

Table 6-4 lists top 20 frequently removed bid items with highest impact on cost.

TABLE 6-4

FREQUENTLY REMOVED BID ITEMS WITH HIGHEST IMPACT ON COST

Item Number	Description	Unit	Item Extended Sum	Item Removed Frequency
303-01	MINERAL AGGREGATE, TYPE A BASE, GRADING D	TON	$116,856,552.37	184

Item Number	Description	Unit	Item Extended Sum	Item Removed Frequency
203-01	ROAD & DRAINAGE EXCAVATION (UNCLASSIFIED)	C.Y.	$96,770,016.03	158
717-01	MOBILIZATION	LS	$63,308,025.00	202
203-02.01	BORROW EXCAVATION (GRADED SOLID ROCK)	TON	$43,991,996.37	59
501-01.01	PORTLAND CEMENT CONCRETE PAVEMENT (PLAIN) 8"	S.Y.	$40,019,750.82	5
307-02.01	ASPHALT CONCRETE MIX (PG70-22) (BPMB-HM) GRADING A	TON	$35,485,104.64	93
307-03.01	ASPHALT CONCRETE MIX (PG76-22) (BPMB-HM) GRADING A	TON	$28,637,881.70	15
411-02.10	ACS MIX(PG70-22) GRADING D	TON	$25,819,834.74	96
203-03	BORROW EXCAVATION (UNCLASSIFIED)	C.Y.	$25,789,700.98	63
307-01.08	ASPHALT CONCRETE MIX (PG64-22) (BPMB-HM) GRADING B-M2	TON	$25,777,984.89	135
307-02.08	ASPHALT CONCRETE MIX (PG70-22) (BPMB-HM) GRADING B-M2	TON	$25,380,297.47	92
201-01	CLEARING AND GRUBBING	LS	$23,687,500.00	160
307-01.01	ASPHALT CONCRETE MIX (PG64-22) (BPMB-HM) GRADING A	TON	$20,511,496.45	86
712-02.02	INTERCONNECTED PORTABLE BARRIER RAIL	L.F.	$19,161,582.26	117
604-07.01	RETAINING WALL (DESCRIPTION)	S.F.	$18,096,256.90	28
501-01.03	PORTLAND CEMENT CONCRETE PAVEMENT (PLAIN) 10"	S.Y.	$18,028,672.00	26
411-01.10	ACS MIX(PG64-22) GRADING D	TON	$15,410,809.22	120
313-03	TREATED PERMEABLE BASE	S.Y.	$15,370,147.14	32
712-01	TRAFFIC CONTROL	LS	$14,413,155.00	202
307-01.21	ASP. CONC. MIX(PG70-22) (BPMB-HM) GR. A-S	TON	$13,409,675.13	39

During this analysis, one of the key reasons behind significant cost growth in TDOT estimates from initial to the final engineers' estimate was identified while discussing with TDOT. When details about any item (say a culvert) was not confirmed, TDOT did not add the cost of such items in the estimate. The result was a missing cost for potentially significant bid item. Later TDOT corrected the course by adding the cost of most likely variation of the item in the estimates so that its cost is included in the estimate. This helped TDOT reduce the magnitude of changes in engineers' estimates.

6.2 PCEA Tool

The main screen of the PCEA tool – that includes all the required inputs for adjusting cost estimates based on historical trend – is show in Figure 6-1.

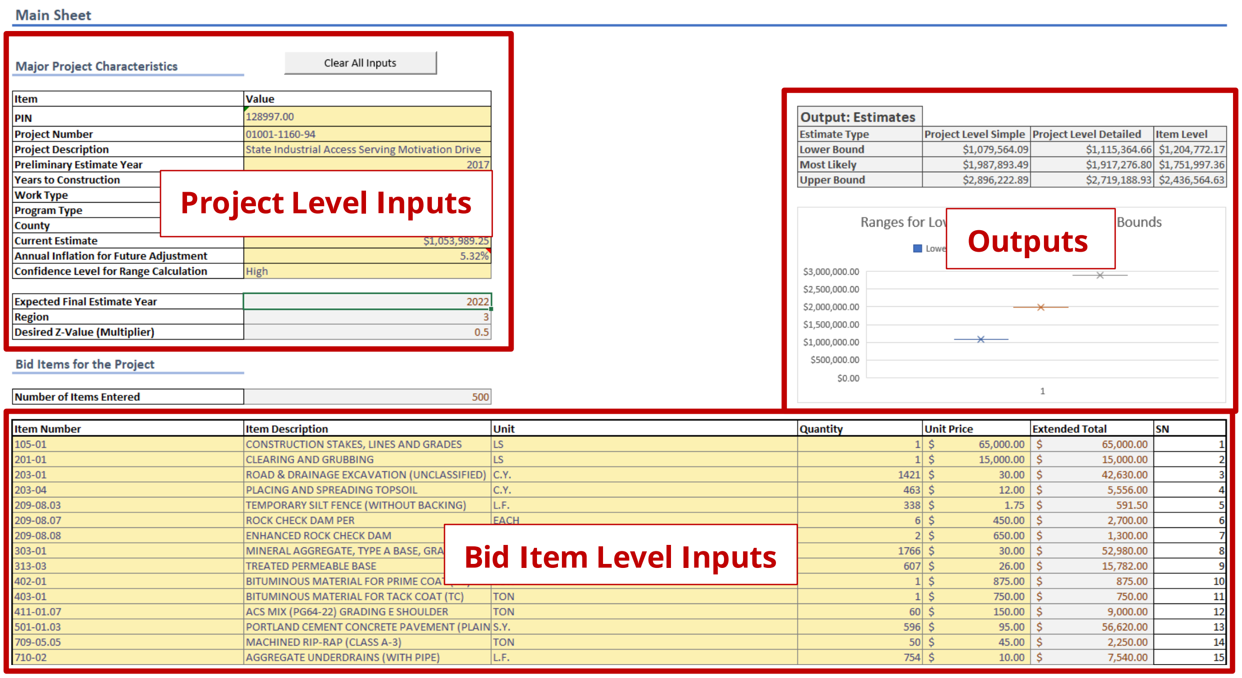

Figure 6-1: Overview of PCEA Tool

Main Sections

The PCEA tool consists of three main sections: a) project level inputs, b) bid item level inputs, and c) outputs.

Project Level Inputs

The main required project level inputs include preliminary estimate year, years to construction, work type, program type, county, and current estimate. The preliminary estimate year is the year when the current estimate was prepared for the market condition of the same year (i.e., not for planned construction year). A default annual inflation rate is provided based on historical data. This inflation year should be representative of the average inflation from the year of estimate to the year of planned construction. For example, if the original estimate was prepared for year 2023, and the planned construction year is 2030, then the inflation should be representative of the average inflation from year 2023 to year 2030. If inflation rates for later years are not available, average inflation from estimate year (2023) to current year (2025) can be used unless predictive inflation rates are available. Further, the inflation rate should be specific to highway construction industry in the state of Tennessee. If such inflation rate is not available, national inflation rate for highway construction industry can be used as less accurate substitute. Currently, National Highway Construction Cost Index (NHCCI) is available for this purpose. However, TDOT should strongly consider developing TDOT specific Highway Construction Cost Index to compute inflation rate specific to Tennessee highway construction industry.

Bid Item Level Inputs

For the bid item level inputs, all bid items along with the current quantity and unit price estimates are required.

Outputs

Based on the project level and bid item level inputs, three sets of outputs are generated: program level adjusted estimates, project level adjusted estimates, and bid item level adjusted estimates. Each set of outputs are probabilistic outputs that includes lower, likely, and upper adjusted estimates. The program level adjusted estimates are based on the overall trend of changes in estimates during the project development of all projects across the state. The project level adjusted estimates are based on the trend of changes in estimates from other projects with similar project characteristics, such as work type and location. The bid item level adjusted estimates are based on the trend of changes in quantity, unit price, and extended amount for each bid item.

Validation of PCEA Tool

A dataset consisting of 44 projects – that were not included in the model development – was used for validation. This dataset includes the initial engineers' estimate as well as the final engineers' estimate. The final engineers' estimate is considered as the desired correct estimate. Then the adjusted engineers' estimate was calculated using the tool based on the initial engineers' estimate. Now, three sets of estimates are available: a) initial engineers' estimates, b) final engineers' estimate, and c) adjusted engineers' estimate (tool's output). Subsequently two sets of errors are computed: i) error in initial engineers' estimate (compared to the final engineers' estimate) and ii) error in adjusted engineers' estimate (compared to the final engineers' estimate). If the tool improved the accuracy of the estimate, the second set of errors should be lower than the first set of errors. This reduction in the error is the improvement in accuracy by utilizing the tool. These errors and improvements are calculated in two forms: Mean Percentage Error (MPE) and Mean Absolute Percentage Error (MAPE).

The MPE calculation considers increase and decrease in estimates differently, and the increase and decrease (or errors in positive and negative directions) cancels each other out. In MAPE calculation, absolute values of the differences are considered, and hence both increase and decrease are accounted similarly. The MAPE is always equal or larger than MPE. For long term planning and budgeting where a group of projects are considered, MPE is more relevant and cost increase and cost decrease cancels each other out for total budgeting. For example, if the actual cost of one project increases by $500,000 and the cost of another project decreases by $500,000, then the total budget required for the two projects remain same.

The validation was conducted for program level and project level data. The validation result is presented in **Table 6-5**. The positive number indicates that the accuracy of the estimates increased by using the tool, while negative numbers indicate the opposite. Overall, the accuracies of all validation dataset were all improved by over 55.08% MPE and 37.89% MAPE with program level adjustment. The project level adjustment also increased the accuracy by similar percentages (54.26% MPE and 37.26% MAPE). The improvement is highest for the *Legislative – Reconstruction* and *Safety – Intersection Improvement* work types. The tool also decreased the accuracy of the estimates for in four instances. It should be noted that these improvements in the validation dataset does not necessarily mean that similar level of improvements will be seen in for other new projects. Further, these models need to be updated over time. Otherwise, changes in the estimating practices will not be reflected, and hence the accuracy of the models will decrease.

TABLE 6-5

ESTIMATE ACCURACY IMPROVEMENTS USING PCEA TOOL

Program - Work	No of Project	Program Level Accuracy Improvement (%)		Project Level Accuracy Improvement (%)	
		MPE	MAPE	MPE	MAPE
Bridge - Bridge Replacement	17	11.64	8.88	16.33	6.46
Safety - Intersection Improvements	1	159.98	159.98	**193.80**	**193.80**
Safety - Intersection Improvements and Signals	1	7.18	7.18	-6.42	-6.42
Safety - Signalization	1	18.81	18.81	16.39	16.39
Safety - Turn Lanes with Signal	2	4.10	10.71	7.05	9.88
Safety - Safety	3	14.97	11.91	21.60	8.66
Safety - Miscellaneous Safety Improvements	4	1.81	-0.29	10.91	4.12
Legislative - Widen	6	131.17	121.41	148.38	113.62
Legislative - Reconstruction	2	**174.75**	**162.06**	190.33	173.98
Economic Development - Reconstruction	4	59.30	39.47	58.68	39.89
Economic Development - Turn Lanes	3	-5.65	20.30	-2.91	23.04
Overall		55.08	37.89	54.26	37.26

* Bold values indicate the highest accuracy improvements for the column while underlined values indicate the highest accuracy deterioration for the column.

6.3 CPCE Tool

The main screen of the CPCE tool – that includes all the required inputs estimating construction project costs based on historical trend – is show in Figure 6-2. Major inputs required for the tool are planned construction year, type of work, county, route type, length, and ROW cost. The tool instantly generates the output based on the project characteristics. The tool uses county, region, and state level weights of 50%, 30%, and 20% respectively. The custom weights ensure that the reliability from larger state and regional level datasets as well as specific geographic and localized market conditions from county level dataset are represented while calculating the final estimate.

Main Sheet

Major Project Characteristics

Description	Value	
PIN	128997	
Project Number	01001-1160-94	
Project Description	Road Widening Project in Hamblen	
Planned Construction Year	2024	
Project Type of Work	Legislative - Widen	
County	Hamblen	
Route Type	State	
Length	3.8	miles
ROW Cost	$37,800,000.00	
Use Custom Inflation Rate	No	
Custom Inflation Value	3.50%	

Default Inflation Value	5.60%
Region	R1

Estimated Value for Construction Year	**$63,005,746.10**

Figure 6-2: Overview of CPCE Tool

Validation of CPCE Tool

The CPCE tools is validated in two phases: a) Statistical Validation and b) Compliance with AASHTO Practical Guide. Project information of 59 projects was used for the validation.

Statistical Validation

Mean Percentage Error (MPE) and Mean Absolute Percentage Error (MAPE) were calculated for the available validation dataset. The MPE indicates the average directional error (positive for overestimation and negative for underestimation), while MAPE measures the average magnitude of error. **Table 6-6** summarizes the results of the validation. The absolute error of estimating for the *Legislative – Reconstruction* project type was found to be lowest with -7% and 9% MPE and MAPE, respectively. This indicates the reliability of the legislative model.

TABLE 6-6

STATISTICAL VALIDATION OF THE CPCE PERFORMANCE

Project Type	MPE	MAPE
Bridge - Bridge Replacement	9%	62%
Legislative – Reconstruction	**-7%**	**9%**
Legislative – Widen	35%	51%

Project Type	MPE	MAPE
Safety - Bicycles and Pedestrians Facility	86%	86%
Safety - Intersection Improvements	-90%	90%
Safety - Intersection Improvements and Signals	-58%	29%
Safety - Miscellaneous Safety Improvements	149%	176%
Safety – Safety	80%	103%
Safety – Signalization	-93%	93%
Safety - Turn Lanes with Signal	-60%	60%
Economic Development - Reconstruction	145%	146%
Economic Development - Turn Lanes	-37%	37%
Overall	106.04%	143.91%

* Bold values indicate the lowest absolute value for any row while underlined values indicate the highest absolute value for any row.

The framework, models, and tools were initially developed for legislative projects. Consideration for other project types were added later upon TDOT request. As such, the project characteristics collected from TDOT that are used for model development were more relevant to legislative projects. Consequently, the models for legislative projects appear to be more reliable than others. These project characteristics were less impactful to describe the estimates for other project types. For example, the *Safety – Miscellaneous Safety Improvements* had the highest errors with MPE value of 149% and MAPE value of 176%. It indicates the need to utilize additional project parameters for developing models for other project types. For example, more specifics of miscellaneous safety improvements would help improve corresponding model.

Compliance with AASHTO Practical Guide.

The expected cost estimation accuracies for different phases of project development – as prescribed by AASHTO – is used as the basis for the second phase of validation (AASHTO, 2013). The range of values for estimates stated in the AASHTO guideline is presented in **Table 6-7**. As expected, the range of estimates narrows down from planning phase to final design phase, which indicates that the expected accuracy of the estimates should increase as the project development continues.

TABLE 6-7

EXPECTED COST ESTIMATION ACCURACIES BY PROJECT DEVELOPMENT PHASES - AASHTO PRACTICAL GUIDE FOR COST ESTIMATION (SOURCE: AASHTO (2013))

Phase	Project Maturity	Purpose of the Estimate	Estimation Methodology	Error Range
Planning	0% - 2%	Conceptual Estimating – Estimate Potential Funds Needed	Parametric (Stochastic or Judgement)	-50% to 200%

Phase	Project Maturity	Purpose of the Estimate	Estimation Methodology	Error Range
	1% - 15%	Conceptual Estimating – Prioritize Needs for Long-Range Plan	Parametric or Historical Bid-based (Primarily Stochastic)	-40% to 100%
Scoping	10% - 30%	Design Estimating – Establish a Baseline Cost for Project and Program Projects	Historical Bid-Based or Cost-based (Mixed but Primarily Stochastic)	-30% to 50%
Design	30% - 90%	Design Estimating – Manage Project Budgets against Baseline	Historical Bid-Based or Cost-Based (Primarily Deterministic)	-10% to 25%
Final Design	90% - 100%	PS&E Estimating – Compare with Bid and Obligate Funds for Construction	Cost-based or Historical-Bid Based Using Cost Estimate System (Deterministic)	-5% to 10%

Table 6-8 presents the overall compliance of the costs estimated by the models. The CPCE tools is a parametric-model-based tool developed for computing estimates for early phase of project development. As such, the compliance of the estimates is higher for planning phases than the later phases. Four project types had 100% compliance rate for 0% – 2% planning phase while three had 100% compliance for 1% – 15% planning.

TABLE 6-8

COMPLIANCE OF THE COST ESTIMATES IN REFERENCE TO THE AASHTO GUIDELINE

Project Type	Total Project	Percentage of Compliance (%)				
		Planning (0% - 2%)	Planning (1% - 15%)	Scoping (10% - 30%)	Design (30% - 90%)	Final Design (90% - 100%)
		-50% to 200%	-40% to 100%	-30% to 50%	-10% to 25%	-5% to 10%
Bridge - Bridge Replacement	19	73.68	52.63	42.11	21.05	10.53
Legislative – Reconstruction	2	**100.00**	**100.00**	100.00	50.00	50.00
Legislative – Widen	7	**100.00**	71.43	57.14	28.57	14.29
Safety - Bicycles and Pedestrians Facility	1	**100.00**	**100.00**	0.00	0.00	0.00
Safety - Intersection Improvements	1	0.00	0.00	0.00	0.00	0.00
Safety - Intersection Improvements and Signals	2	**100.00**	**100.00**	50.00	0.00	0.00
Safety - Miscellaneous Safety Improvements	15	46.67	46.67	26.67	20.00	6.67
Safety – Safety	5	80.00	60.00	20.00	20.00	20.00
Safety – Signalization	1	0.00	0.00	0.00	0.00	0.00
Safety - Turn Lanes with Signal	1	0.00	0.00	0.00	0.00	0.00
Economic Development – Reconstruction	4	75.00	25.00	25.00	25.00	25.00

Project Type	Total Project	Percentage of Compliance (%)				
		Planning (0% - 2%)	Planning (1% - 15%)	Scoping (10% - 30%)	Design (30% - 90%)	Final Design (90% - 100%)
		-50% to 200%	-40% to 100%	-30% to 50%	-10% to 25%	-5% to 10%
Economic Development - Turn Lanes	1	**100.00**	**100.00**	0.00	0.00	0.00

Legislative – Reconstruction, Safety – Bicycles and Pedestrians Facilities, Safety – Intersection Improvements and Signals, and *Economic Development – Turn Lanes* projects have the highest level of compliance (100%) for 0% – 2% planning phase. *Safety – Intersection Improvements, Safety – Signalization,* and *Safety – Turn Lanes with Signals* projects had 0% compliance for planning phase (0% - 2%) indicating the high level of variability in the estimates based on available parameters. The model with higher compliance rates may be acceptable, but ones with lower compliance rates are less reliable, and hence may need further refinements. Further, the results show that the tool can be used for various project types for conceptual and scoping phases, but it is less suitable for design and final design phases.

6.4 Limitations and Weaknesses

Machine learning models are not meant to substitute engineering judgement. The models and tools developed in this study are based on historical datasets. These models tend to behave well when project characteristics of new project is similar to past projects, but they may be unreliable if the new project is unique. For example, if none of the past projects included retention walls, but a new project has a retention wall, the model cannot capture and respond to the inclusion of the retention wall while predicting its cost. Each construction project is unique, and it is not possible to consider every detail of a project for developing mathematical models. Some of the additional relevant factors that affect the construction costs include local competition, local market conditions, terrain, environmental restrictions, and traffic conditions. These project characteristics were not easily accessible for modeling. Further, to develop more reliable models, project-type-specific characteristics may be required. For example, the foundation type would significantly affect a bridge replacement project cost. Such specialized models were not within the scope of the study and were not developed.

Chapter 7 Conclusion and Recommendation

This chapter presents the major conclusions and recommendations of the study.

7.1 Conclusions

This study collected and analyzed project estimate data from year 2013 to 2024 to help improve the cost estimating practices at TDOT. The analysis showed that 28 bid items were responsible for more than 80% of the total cost. This highlights the potential of improving the accuracy of project cost estimates if the quantity and unit price of these high impact items can be calculated accurately.

The study developed frameworks and tools for Preliminary Cost Estimate Adjustment (PCEA) and Construction Project Cost Estimation (CPCE) tool to improve the current cost estimation practices. The PCEA framework and tool enables adjusting initial engineers' estimates to make it more reflective of the likely final engineers' estimate by utilizing the historical trend in cost changes over time during the project development. The CPCE framework and tool enables computing preliminary estimates using various project characteristics. These two tools are expected to aid TDOT engineers in increasing the accuracy of their engineers' estimates, which will enable better planning of future projects based on available funding.

PCEA Tool

The average cost growth of construction projects from the initial engineers' estimate to the final engineers' estimate is over 45%. The PCEA tool is developed to address such discrepancy between the initial and final engineers' estimate by analyzing historical data and identifying cost growth associated with various project characteristics and bid items. Three level of cost adjustments can be performed depending on the level of information available: a) program level simple, b) project level, and c) item level. Overall, the tool was able to reduce the overall error in validation dataset by 37% to 55%.

CPCE Tool

The CPCE tool utilizes project characteristics, such as, type of work, planned construction year, route type, length, Right of Way (ROW) cost, and inflation rate to compute estimates based on historical data. The tool consists of 290 county level models, 205 region level models, and 135 state level models. The estimates from the models are aggregated in two phases using Relative Reliability Indexes and custom weights. Overall, the tool was able to compute estimates that were acceptable for planning phase of project development based on AASHTO guidelines.

7.2 Recommendations

Below are the recommendations to continuously improve TDOT's cost estimating practices.

Systematic Project Characteristics Database

More in-depth study should be conducted to identify various project characteristics that are relevant in determining the cost of various types of projects. Separate databases should be developed to collect these project characteristics for various types of projects. For example, one database may be developed to collected Bridge Replacement project characteristics and another one for Intersection Improvement projects. These systematic databases would enable automating future estimating model development.

Specialized Models for Project Types

Current models utilize a limited set of common project characteristics for all project types. More in-depth study should be conducted to develop specialized models using project characteristics that are relevant to a specific project type. Such specialized models should improve the accuracy of the estimates.

Bid Item Level Cost Estimation Models

The CPCE tool utilizes parametric models to compute estimates based on project characteristics when bid item information is not available. This tool is suitable for early phase of project development. Once bid items and their quantities are available, many state DOTs typically uses statewide average unit prices of the bid items. However, such statewide average unit prices ignore various project characteristics, such as, project location, project type, and project complexities. As such, state DOTs do not obtain the full benefit of bid item level cost estimates. TDOT should conduct a study to develop a comprehensive bid item unit price estimating tool for more accurate estimating for later stages of project development.

Systematic Cost Estimates History Database

TDOT currently stores cost estimates in spreadsheets and PDFs. Extracting relevant data from such unstructured and semi-structured documents is completely manual, time-consuming, very cumbersome, and error prone. TDOT should consider developing a more structured relational database such as MS SQL, Oracle, PostgreSQL, MySQL, or MariaDB (not just spreadsheet) and associated digital tool to store estimates produced in various phases of project development. Further, every time estimates are updated, the rationale for the changes should be noted. These notes may be related to various project risks, changes in scope, changes in market conditions, etc. These changes should be recorded in the project level as well as bid item level as relevant. For bid item level, the database should also be able to indicate the item substitution, item removal, and item addition.

Cost Estimating Performance Dashboard

An interactive dashboard showing the accuracy of estimates at various phases of project development can be developed. It can also highlight the potential improvements in the accuracy of estimates as a result of ongoing initiatives, research, and changes in the estimating practices. This would help upper management to get a high-level overview of the impacts of these activities.

Continuous Monitoring, Evaluation, and Improvement

The CPCE tool should be used alongside existing estimating practices for some time, and its performance should be continuously monitored and evaluated. The strength and weakness of the tools in computing estimates in various scenarios should be identified and documented. As newer datasets are available, the models should be continuously updated to be more reflective of the newer market conditions and construction techniques.

TDOT-Specific Highway Construction Cost Indexes

To compute a reliable estimate for future construction projects based on historical data, appropriate inflation rate must be used. The highway construction industry generally experiences different inflation rate than the cost of living. Further, the inflation rate for the construction industry varies by the location. Many state DOTs including Montana DOT calculates state specific Highway Construction Cost Indexes (HCCIs) that represents the trend of inflation and purchasing power of the agency (Shrestha et al., 2017). However, TDOT currently lacks such HCCIs and utilizes third party cost of living index, which is not suitable for construction cost estimating. The research team strongly recommends TDOT to develop a comprehensive HCCIs for estimating purposes.

Impact of Competition on Bids

Local competition can significantly impact construction costs. As such, a thorough investigation of the impact of local competition (i.e., multiple bids) on the bids and hence construction cost should be investigated. The results of such study can be integrated to the cost estimation models to improve the accuracy of the estimates further.

References

AASHTO. (2013). *Practical Guide to Cost Estimating* (First).

Anderson, S., Molenaar, K., & Schexnayder, C. (2006). *Final Report for NCHRP Report 574: Guidance for Cost Estimation and Management for Highway Projects During Planning, Programming, and Preconstruction* (Nos. 978-0-309-43213–9; Issues 978-0-309-43213–9). https://doi.org/10.17226/22045

Ashuri, B., Kingsley, G., Rogers, J., Gahrooei, M. R., Ilbeigi, M., Sung, E. J.-Y., Toroghi, S. (Sean) H., & Georgia Institute of Technology. (2015). *Streamlining project delivery through risk analysis.* (No. FHWA-GA-15-1305). https://rosap.ntl.bts.gov/view/dot/29370

California Department of Transportation. (2012). *Project Risk Management Handbook: A Scalable Approach*. http://www.dot.ca.gov/hq/projmgmt/documents/prmhb/PRM_Handbook.pdf

Duverlie, P., & Castelain, J. M. (1999). Cost Estimation During Design Step: Parametric Method versus Case Based Reasoning Method. *The International Journal of Advanced Manufacturing Technology*, *15*(12), Article 12. https://doi.org/10.1007/s001700050147

Gardner, B. J., Gransberg, D. D., & Rueda, J. A. (2017). Stochastic Conceptual Cost Estimating of Highway Projects to Communicate Uncertainty Using Bootstrap Sampling. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering*, *3*(3), Article 3. https://doi.org/10.1061/AJRUA6.0000895

Gransberg, D. D., Jeong, H. D., Craigie, E. K., Rueda-Benavides, J. A., & Shrestha, K. J. (2016). *Estimating Highway Preconstruction Services Costs—Volume 1: Guidebook*. Transportation Research Board. https://doi.org/10.17226/23590

Kim, B., & Hong, T. (2012). Revised Case-Based Reasoning Model Development Based on Multiple Regression Analysis for Railroad Bridge Construction. *Journal of Construction Engineering and Management*, *138*(1), Article 1. https://doi.org/10.1061/(ASCE)CO.1943-7862.0000393

Liu, M., Hummer, J. E., Rasdorf, W. J., Hollar, D. A., Parikh, S. C., Lee, J., & Gopinath, S. (2011). Preliminary Engineering Cost Trends for Highway Projects. *Report No. FHWA/NC/2010-10*. https://trid.trb.org/view/1141603

Molenaar, K. R. (2005). Programmatic Cost Risk Analysis for Highway Megaprojects. *Journal of Construction Engineering and Management*, *131*(3), Article 3. https://doi.org/10.1061/(asce)0733-9364(2005)131:3(343)

New York State Department of Transportation. (2009). *Risk Management for Project Development*. https://www.dot.ny.gov/divisions/engineering/design/dqab/dqab-repository/Risk%20Management%20for%20Project%20Development%20Guide_Final%20Draft_010514.pdf

Piratla, K. R., Le, T., Jamal, M. S., & Do, Q. (2024). A Preliminary Cost Estimating Model for Transportation Projects. *Report No. FHWA-SC-24-03*. https://rip.trb.org/View/1864558

Sadeh, H., Mirarchi, C., & Pavan, A. (2021). Integrated Approach to Construction Risk Management: Cost Implications. *Journal of Construction Engineering and Management*, *147*(10), Article 10. https://doi.org/10.1061/(asce)co.1943-7862.0002140

Saito, M., Sinha, K. C., & Anderson, V. L. (1991). Statistical models for the estimation of bridge replacement costs. *Transportation Research Part A: General*, *25*(6), Article 6. https://doi.org/10.1016/0191-2607(91)90012-F

Schexnayder, C. J., Molenaar, K. R., Anderson, S. D., Transportation Research Board, National Cooperative Highway Research Program, & Transportation Research Board. (2009). *Procedures Guide for Right-of-Way Cost Estimation and Cost Management* (p. 14289). National Academies Press. https://doi.org/10.17226/14289

Shane, J. S., Molenaar, K. R., Anderson, S., & Schexnayder, C. (2009). Construction Project Cost Escalation Factors. *Journal of Management in Engineering*, *25*(4), 221–229. https://doi.org/10.1061/(ASCE)0742-597X(2009)25:4(221)

Shrestha, K. J., Jeong, H. D., & Gransberg, D. D. (2017). Multidimensional Highway Construction Cost Indexes Using Dynamic Item Basket. *Journal of Construction Engineering and Management*, *143*(8), 04017036. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001307

Texas Department of Transportation. (2015). *Risk-Based Construction Cost Estimating*. 286.

Trost, S. M., & Oberlender, G. D. (2003). Predicting Accuracy of Early Cost Estimates Using Factor Analysis and Multivariate Regression. *Journal of Construction Engineering and Management*, *129*(2), Article 2. https://doi.org/10.1061/(ASCE)0733-9364(2003)129:2(198)

US Department of Transportation. (2022, March 30). *Strategic Plan: FY 2022-2026*. https://www.transportation.gov/sites/dot.gov/files/2022-04/US_DOT_FY2022-26_Strategic_Plan.pdf

Washington State Department of Transportation. (2014). *Project Risk Management Guide*. http://www.wsdot.wa.gov/publications/fulltext/cevp/ProjectRiskManagement.pdf

Appendices

Additional technical background on Preliminary Cost Estimate Adjustment and Construction Project Cost Estimation Models are provided below.

Preliminary Cost Estimate Adjustment

Additional theoretical context about the preliminary cost estimate adjustment is provided in this section.

Confidence Interval

To calculate a lower, likely, and upper estimates at various confidence levels, Equation (10) is used.

$$Confidence\ Interval = \bar{X} \pm Z * \sigma \qquad (10)$$

$\bar{X}$ = Mean of the cost estimates

Z = Z-score corresponding to the desired confidence level

σ = Standard deviation of the cost estimates

Statistical Measures

Various statistical measures calculated for the CPCE tool include mean, range, standard deviation, and coefficient of variation (CoV). These metrics are evaluated across the full dataset and segmented by factors like region, work type, and program category, revealing distinct cost behaviors tied to project traits.

1. Mean: The average percentage difference, reflecting overall cost shift tendencies.
2. Range (Min/Max): The extremes of cost deviations, identifying significant over- or underestimations.
3. Standard Deviation: The dispersion of differences around the mean indicates estimated consistency.
4. Coefficient of Variation (CoV): The ratio of standard deviation and mean – it enables variability comparisons across datasets.

Construction Project Cost Estimation Models

Theoretical background for various construction project cost estimation models is provided here for reference.

Overview of Multiple Linear Regression (MLR)

The MLR was selected for the model development. The MLR is able to utilize multiple independent variables (project characteristics in this case) to predict a dependent variable (cost estimate in this case). Further, MLR can be implemented in spreadsheet-based tools for predicting cost estimates for future projects. The general form of a MLR can be expressed by Equation (11).

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x + \epsilon \quad (11)$$

Where,

y	=	dependent variable
$x_1, x_2, \ldots\ldots x_n$	=	independent variable
β_0	=	y-intercept of the regression line
$\beta_1, \beta_1, \ldots \beta_n$	=	slope of the regression line
ε	=	error term

Through domain knowledge, exploratory data analysis (EDA), and correlation analysis, several key independent variables were identified, leading to the development of the models shown in Equation (12) - (15).

$$y = \beta_0 + \beta_1 \times \text{Project Length} \quad (12)$$

$$y = \beta_0 + \beta_1 \times \text{Project Length} + \beta_2 \times \log(\text{ROW Cost}) \quad (13)$$

$$y = \beta_0 + \beta_1 \times \text{Project Length} + \beta_2 \times \log(\text{ROW Cost}) + \beta_3 \times \text{Route Type} \quad (14)$$

$$y = e^{\beta_0 + \beta_1 \times \log(\text{Length}) + \beta_2 \times \log(\text{ROW Cost})}$$

$$y = e^{\beta_0 + \beta_1 \times \log(\text{Length}) + \beta_2 \times \log(\text{ROW Cost}) + \beta_3 \times \text{Route Type}} \quad (15)$$

Model Fitting

The model parameters ($\beta_0, \beta_1, \ldots\ldots\ldots \beta_n$) were estimated using the Ordinary Least Squares (OLS) method. The OLS minimizes the sum of squared residuals (RSS), which can be expressed by Equation (16).

$$\text{RSS} = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \quad (16)$$

Where,

y = actual observed value

$\hat{y}_i$ = predicted value from the model

A stepwise regression technique was employed to iteratively add or remove predictors based on their statistical significance and impact on model accuracy. The inclusion or exclusion of predictors was determined using p-values and adjusted R^2 values.

Development of the MLR Models

1. *Predictor Evaluation*: Statistical significance tests were conducted to determine the impact of independent variables on cost estimation.
2. *Validation of Model Assumptions*: Several regression assumptions were tested to ensure the robustness of the model.

Assumptions and Model Robustness

To ensure the validity of the MLR models, several key assumptions were tested:

Linearity: Scatter plots and residual analysis confirmed a linear relationship between the predictors and the dependent variable.

Independence of Observations: The Durbin-Watson test was conducted to detect autocorrelation in the residuals, ensuring that observations were independent.

Homoscedasticity: Residual plots were examined to verify that the variance of residuals remained constant across different levels of predictors.

Multicollinearity: The Variance Inflation Factor (VIF) was calculated to assess multicollinearity among predictors. Variables with VIF values exceeding five were removed or transformed to reduce collinearity.

Model Evaluation and Goodness-of-Fit Measures

The performance of the developed model was evaluated using several goodness-of-fit metrics:

ANOVA Test: The overall significance of the model was assessed using ANOVA, which determines whether the independent variables collectively explain a significant portion of the variation in the dependent variable.

Residual Analysis: Residual plots (e.g., residual vs. fitted plots) were used to visually assess whether the model's residuals followed the assumptions of normality and randomness.