SCHOOL OF GRADUATE STUDIES

DEPARTMENT OF CONSTRUCTION TECHNOLOGY AND MANAGEMENT

Developing construction duration prediction model for Ethiopian Road Administration Road construction projects

By: Yafet Girma

Advisor: Dr. Dagnachew Adugna (Associate professor)

August, 2025

Addis Ababa, Ethiopia

SCHOOL OF GRADUATE STUDIES

DEPARTMENT OF CONSTRUCTION TECHNOLOGY AND MANAGEMENT

Developing construction duration prediction model for Ethiopian Road Administration Road construction projects

By: Yafet Girma

Advisor: Dr. Dagnachew Adugna (Associate Professor)

A Thesis Submitted to Addis College School of Graduate Studies in Partial Fulfillment of the Requirements for the Degree of Master of Science in Construction Technology and Management

August, 2025

Addis Ababa, Ethiopia

ADDIS COLLEGE

School of Graduate Studies

APPROVAL SHEET

Department of Construction Technology and Management

Developing construction duration prediction model for Ethiopian Road Administration Road construction projects

A Thesis Submitted for the Partial Fulfillment of the Requirements for the Degree of master's in Construction Technology and Management

By: Yafet Girma

Endorsed by:

------------------------------------	-----------------------------------	---------------------
Main advisor	Signature	Date
------------------------------------	-----------------------------------	---------------------
External Examiner	Signature	Date
-------------------------------------	----------------------------------	---------------------
Internal Examiner	Signature	Date
-------------------------------------	----------------------------------	---------------------
Chairperson	Signature	Date
------------------------------------	-----------------------------------	---------------------
School Dean	Signature	Date

DECLARATION

I, the undersigned, declare that the study entitled "Developing construction duration prediction model for Ethiopian Road Administration Road construction projects" is my original work and has not been submitted for any Degree or Diploma in any University. To the best of my knowledge, all sources of materials used for the study have been duly acknowledged. I have prepared this thesis paper with the guidance and support of the research advisor.

Name: ______________________________

Signature: __________________

Date of submission: August 2025

Place: Addis Ababa, Ethiopia

STATEMENT OF CERTIFICATION

This is to certify that Yafet Girma has carried out his research paper entitled "Developing construction duration prediction model for Ethiopian Road Administration Road construction projects". This work is original in nature and is suitable for thesis submission for the award of a Master of Science in Construction Technology and Management.

______________________	____________________________	____________________
Advisor Name	Signature	Date

ACKNOWLEDGEMENT

First and foremost, I thank Almighty God for granting me the commitment and perseverance to overcome various obstacles and successfully complete this thesis. Next, I extend my sincere gratitude to my advisor, Dr. Dagnachew Adugna (Associate Professor), for his invaluable guidance and advice throughout the preparation of this paper. Finally, my appreciation goes to all ERA engineers and officers who assisted me by providing necessary project documents.

ABSTRACT

Inadequate early-stage project planning often undermines overall project performance, making accurate duration estimation crucial, particularly during the bidding phase of road construction. Currently, the Ethiopian road construction industry largely relies on subjective, experience-based estimations, highlighting a critical lack of robust, data-driven planning tools. Therefore, the objective of the study is to develop a practical-based construction duration prediction model specifically for road projects managed by the Ethiopian Road Administration (ERA). The study utilized a quantitative research approach, analyzing historical data from 30 ERA road projects completed from 2001 to 2016. Simple and multiple regression analyses were used for road duration model development based on untransformed and natural logarithm-transformed datasets. The simple regression models—Bromilow's Time–Cost and various Time-Cost and Time-Road Length relationships—demonstrated limited explanatory power (BTC model R^2 *=0.151, untransformed Compound/Growth/Exponential Time-Cost* R^2*=0.285, Ln-transformed Time-Road Length Quadratic model* R^2*=0.209). Conversely, the multiple regression analysis yielded significantly higher explanatory power. Specifically, the Ln-transformed Multiple Linear Regression (MLR) model emerged as the optimal choice, incorporating Actual Cost (AC), Contractor Category (CC), Project Scope (PS), Highway System (HS), and Site Accessibility (SA). This model explained a substantial 78.6% of the variance in project duration (*R^2*=0.786) and exhibited a low Mean Absolute Percentage Error (MAPE) of 1.43%. The model's strong explanatory power and high precision were further substantiated by rigorous validation of regression assumptions, including linearity, homoscedasticity, normality of residuals, and the absence of significant multicollinearity. The developed model is applicable for front-end predictions of construction duration, which can then be used as a policy-setting tool or as a basis by which to evaluate either the planner's duration estimate or construction performance. However, it was not intended as a replacement for estimates of project duration developed using detailed construction scheduling techniques.*

Key words: *Construction Duration Prediction, Bromilow's Time–Cost model, Time-Cost relationship, Time-Road Length relationship, Road Construction*

LIST OF TABLES

LIST OF FIGURES

Table of Contents

LIST OF ABBREVATIONS

AA	Average Accuracy Percentages
AADT	Average Annual Daily Traffic
AC	Asphalt Concrete Surfaced Road Project
AC	Actual Cost
AD	Actual Duration
ANN	Artificial Neural Network
ANOVA	Analysis of Variance
BG	Bridge
BTC	Bromilow's Time-Cost
BY	Bays
CC	Contract Cost
CD	Contract Duration
CM	Contract Method
COM	Compound equation
CON	Contractor Category
CPI	Consumer Price Index
CPM	Critical Path Method
CUB	Cubic Equation
DB	Design-Build
DBB	Design-Bid-Build
DBST	Double Surface Asphalt Treatment Road Project
DC	Domestic Contractors
DS	Delivery System
ERA	Ethiopian Roads Authority
ETB	Ethiopian Birr
EXP	Exponential Equation
FDRE	Federal Democratic Republic of Ethiopia
GDP	Gross Domestic Product
GDS	Geometric Design Standard
GRO	Growth Equation
HS	Highway System
IC	International Contractors
IHA	Imperial Highway Authority
IN	Intersection
INV	Inverse Equation
LIN	Linear Equation
LN	Lanes
LOG	Logarithmic Equation
LR	Linear Regression

MAPE	Mean Absolute Percent Error
ML	Machine Learning
MLR	Multiple Linear Regression
MLRM	Multi-Linear Regression Model
MSE	Mean Squared Error
NLRM	Nonlinear Regression Model
NN	Neural Network
PERT	Program Evaluation Review Technique
POW	Power Equation
PS	Project Scope
QUA	Quadric Equation
RC	Road Classification
RL	Road Length
RMSE	Root Mean Square Error
RSDP	Road Sector Development Program
RW	Road Width
SA	Site Accessibility
SCU	S - curve Equation
SPSS	Statistical Package for the Social Sciences
SRA	Step Wise Regression Analysis
SVM	Support Vector Machine Model
VIF	Variance Inflation Factor
WBS	Work Breakdown Structures

Chapter 1: Introduction

1.1. Background of the study

The construction industry is one of the most significant contributors to a country's economic growth. The significant influence of the construction industry has been widely observed in countries worldwide (Finkel, 2015). In developed countries, it accounts for approximately 7–10% of the Gross Domestic Product (GDP) growth, whereas in developing countries the contribution is 3–6% (Ortiz, Castells, & Sonnemann, 2009). According to the first Growth and Transformation Plan (2010/11–2014/15) performance report, the Ethiopian construction sector grew on average by 28.7%, accounting for 8.5% of the country's GDP.

Ethiopia's modern road construction began in the early 20th century with basic efforts under Emperor Menelik II, followed by the first systematic program during the Italian occupation (1936-1941), which built key routes for colonial use (Girma, 2011). After regaining independence, the Imperial Highway Authority (IHA), formed in 1951, focused on rehabilitating and expanding the network, despite financial and geographical hurdles (Girma, 2011). Road development stagnated under the Derg regime (1974-1991). A transformative shift occurred post-1991 with the Ethiopian Roads Authority (ERA) and the Road Sector Development Program (RSDP) in 1997. This program, backed by significant investment, led to unprecedented network expansion and upgrades (Tegegne & Mengistu, 2019). Currently, Ethiopia continues intensive RSDP efforts, aiming to expand all-weather roads, enhance quality, and develop expressways. However, the road construction sector in Ethiopia is critiqued for its poor performance in delivering projects, often resulting in poor quality, delays, cost overruns, safety issues, community dissatisfaction, and other performance deviations that have a significant macroeconomic impact (Ayele, 2019). Project delays and cost escalation are considered the most frequently recurring problems observed in every infrastructure project, particularly in road and railway construction (Kassa, 2020). Scholars argue that inadequate project planning during the early stage of a project is the primary cause of poor project performance.

Early understanding of construction cost and time represents a critical factor of a feasibility study in the early design phase of a project (Attal, 2010) (Williams, 2008) (Njeem, 2012). This is because major decisions on continuation or dismissal of the project are made based these factors (Njeem, 2012). All parties involved in the construction of a project; owners, contractors, and services companies are in need of reliable information about the cost and time in the early stages of the project. In practice, there are two common methods of estimating project completion time: (1) according to the client's time constraints e.g. Occupancy need, or (2) through a detailed analysis of work to be done and resources available, using estimates of the time requirements for each specific activity (Telford, 1994) quoted in S. Ng (2000). Estimates based on client's occupancy need may overlook actual project requirements and result in overly optimistic values and force a project into a desired, rather than realistic, time mold (Bromilow, 1969). The current common scheduling methods using Gantt bar charts or critical path method/program evaluation review technique (CPM/PERT) have been widely used for schedule estimation in the construction industry (Jin et al., 2016). These approaches, however, can only be used when a significant portion of the detailed design is completed (Bhokha & Ogunlana, 1999). (Helvacı, 2008) and (Williams, 2008) stated that in some cases these techniques may be performed for pre-design duration estimates with multiple assumptions, personnel experience, or comparison of similar projects, however the accuracy of these estimates mainly depends on the experience of the planning engineer and the planning process becomes an intuitive and subjective process. Thus to get over this subjective approach, numerous approaches have been suggested to systematically estimate construction duration in an early stage without detailed design information.

The current practice of early duration estimation within the Ethiopian road construction industry is largely based on individual experience, and is highly influenced by the skill, experience, and intuition of the planning engineer. Although there is a high interest and need by both researchers and the construction industry as a whole and road projects, fewer published studies were obtained in Ethiopian construction context; such as (Alemu 2021) and (Liben, and Belachew et al. 2024) focused on building projects, whereas (Assefa 2008), (Menbere 2019), and (Menbere 2022) were focused on the road construction projects. This highlights the lack of a robust early-time planning tool. Consequently, this research aimed to provide more accurate construction time estimates for road projects, intended for use in the early planning phase.

1.2. Statement of the problem

The Ethiopian road construction industry continues to face significant challenges in delivering projects on time and within budget, with project delays and cost escalations being the most frequent issues observed in the sector. Project delays and cost escalation are the most frequently recurring problems observed in every infrastructure project, particularly in road and railway construction. Kassa (2020) confirmed the widespread nature of this problem, attesting that 88% and 80% of road construction projects suffered from poor time performance and cost overruns, respectively. Tadewos & Patel (2018) found that all projects had experienced time overruns with rates ranging from 25% to 264.38% and cost overruns ranging from 4.11% to 135.06%.

The core of this issue lies in the continued reliance on a subjective, experience-based approach to project time estimation. This lack of a standardized, data-driven methodology has been a persistent issue, leading the government to previously waive the use of completion time and allow a low-evaluated cost award system for tender evaluation in 1993. While current scheduling methods like Gantt charts or CPM/PERT are widely used, their accuracy for pre-design estimates depends heavily on the planning engineer's experience, making the process intuitive and subjective.

Ahmed et al. (2023), Musarat et al. (2021), and Miressa & Shimelis (2022) further emphasize that this problem is compounded by a range of factors, including inadequate planning and scheduling, difficulties in project financing, escalation of material prices, slow land acquisition, and poor site management.

To address this critical gap, this research aimed to develop a systematic, quantitative model to support or replace the current intuitive estimation techniques. The study utilized multiple regression analysis, which proved to be a powerful tool for this purpose. This approach allowed for the incorporation of several key variables to create a more accurate and reliable construction duration prediction model. Such a model, grounded in historical project data, is essential to provide more accurate and reliable construction time forecasts. This would help stakeholders improve project performance, enhance competitiveness during bidding, and ultimately contribute to more efficient infrastructure development in Ethiopia.

1.3. Research Objective

1.3.1. General Objective

The general objective of the study is to develop a practical construction duration prediction model for road construction projects in Ethiopia.

1.3.2. Specific Objectives

The specific objectives of the study include:

- To test the predictability, and the applicability, of Bromilow's time-cost model (BTC) to predict the construction durations for Ethiopian road project.
- To analyse empirical relationships between the construction duration-road length for Ethiopian road construction projects.
- To develop a practical conceptual construction duration prediction model for Ethiopian road construction projects based on the gaps of the BTC model.

1.4. Research questions

The following research questions were drawn to address the objective of the study:

- How predictable and applicable is Bromilow's Time-Cost (BTC) model in estimating construction durations for the Ethiopian road projects?
- What is the empirical relationship between construction duration and road length in Ethiopian road construction projects?
- How can a more practical and accurate construction duration prediction model be developed for Ethiopian road projects by addressing the limitations of Bromilow's Time-Cost (BTC) model?

1.5. Scope of the study

1.5.1. Spatial Scope

The study focused mainly on the development of a conceptual construction duration prediction model for Ethiopian road construction projects. Thus, spatially the study was limited to road construction projects administered under the Ethiopian Road Administration. The data used for model development and the resulting output were also applied specifically to Ethiopian road projects.

1.5.2. Thematic Scope

The study's thematic scope was limited to the development of a construction duration prediction model for Ethiopian road construction projects. The study tried to verify the applicability of the BTC model for Ethiopian road projects and explored for best-fit duration-cost relationships and duration road length relationships. The study further extended to develop multiple regression analysis using duration determinant factors which were known early in the project-planning phase. Finally, through comparison among the various models, the study recommended the best practical duration prediction model for Ethiopian road construction projects.

1.5.3. Temporal Scope

The study used extensive reviews of past research related to duration prediction models for road projects to compile the duration prediction factors potentially known during the early planning phase. The reviews covered published journal articles, research reviews, and academic thesis/dissertations in worldwide and local contexts. As this type of research required archives of past-completed projects, the study used data from ERA road projects completed within the period ranging from 2001 to 2016.

1.6. Limitation of the study

This study was primarily constrained by the limited existing research on road construction duration models within Ethiopia. To compensate, the researcher identified key variables by reviewing global studies and drawing on personal experience. In addition, a major challenge was the lack of well-recorded project documentation, which restricted the number of case projects used for model development. Furthermore, because the model depends on past project performance, it will require regular checks and updates to ensure its continued accuracy and relevance.

1.7. Significance of the study

This research output would significantly contribute to the stakeholders (contractors, consultants, owners, and regulatory bodies) who participated in road construction projects. The research was generally intended to give the road construction industry the chance to support their decision-making processes with reasoned information. The study is significant to the construction industry as a whole and road projects in particular, through providing a systematic tool or procedure to support the prevailing subjective decision made by planners for construction duration predictions during the early project phase. The study will assist clients and consultants by providing a systematic tool for an early-time model used for feasibility checking, to incorporate a precise construction time in the bid package, and for preparation the financial and budgeting aspects ahead. Contractors also used the models to check the adequacy of the construction duration initially settled by the client. Furthermore, the study is beneficial to students, Professionals, and researchers as an input for future studies in construction time forecast and related topics.

1.8. Organization of the document

This research paper was organized into five chapters, which include:

Chapter 1: Introduction: This chapter provides the introductory part of the research, which includes the research background, research problem statement, research questions, research objectives, significance of the study, scope of the study, and limitations of the study.

Chapter 2: Literature Review: This chapter presents a conceptual and empirical review of past research works in a global and local context. The review contains contents in relation to the factors affecting road construction duration and various parametric models for road construction time prediction.

Chapter 3: Research Methodology: This chapter presents the methodological framework of the research. It includes research location, research approach and design, research population and sampling technique, the method of data collection, and the processing and analysis of data.

Chapter 4: Research Results and Discussion: This section presents the research result interpretation and discussion based on the findings from the processing and analysis of the data.

Chapter 5: Conclusion and Recommendation: Finally, this chapter presents the summary, conclusions of the study, and recommendations for further research and industry.

Chapter 2: Literature review

2.1. Theoretical literature review

2.1.1. Definition of Project and Project Management

PMI (2004) defines a project as 'a temporary endeavor undertaken to create a unique product, service, or result'. Projects shall accordingly have the following characteristics (PMI, 2004):

- **Temporary**: Temporary means that every project has a definite beginning and a definite end. The end is reached when the project's objectives have been reached, or it becomes clear that the project objectives will not or cannot be met, or the need for the project no longer exists and the project is terminated. Temporary does not necessarily mean short in duration; many projects last for several years. In every case, however, the duration of a project is finite.
- **Unique Products, Services, or Results**: A project creates unique deliverables, which are products, services, or results. Uniqueness is an important characteristic of project deliverables. The presence of repetitive elements does not change the fundamental uniqueness of the project work.
- **Progressive Elaboration**: is a characteristic of projects that accompanies the concepts of temporary and unique. Progressive elaboration means developing in steps and continuing in increments. For example, the project scope will be broadly described early in the project and made more explicit and detailed as the project team develops a better and more complete understanding of the objectives and deliverables. PMI (2004) concludes projects can be undertaken at all levels of the organization and they can involve a single person or many thousands. Their duration ranges from a few weeks to several years. Projects can involve one or many organizational units, such as joint ventures and partnerships (PMI, 2004).

Project management is the application of knowledge, skills, tools, and techniques to project activities to meet project requirements (PMI 2004)). Project management is accomplished through the application and integration of the project management processes of initiating, planning, executing, monitoring and controlling, and closing. The project manager is the person responsible for accomplishing the project objectives. Project managers also manage projects in response to uncertainty. Project risk is an uncertain event or condition that, if it occurs, has a positive or negative effect on at least one project objective. PMI (2004) managing a project includes: Identifying requirements, Establishing clear and achievable objectives, Balancing the competing demands for quality, scope, time, and cost, and adapting the specifications, plans, and approach to the different concerns and expectations of the various stakeholders.

PMI (2004) Project managers often talk of a 'triple constraint'- project scope, time, and cost-in managing competing project requirements. Project quality is affected by balancing these three factors. High-quality projects deliver the required product, service, or result within scope, on time, and within budget. The relationship among these factors is such that if any one of the three factors changes, at least one other factor is likely to be affected. Thus, projects can better be managed in ways that balance these constraints. But, overly emphasizing one of these aspects may compromise the other. Changing any of the three without adjusting one or all of the others may affect the quality of the project outputs, thus making the task of having a balance almost impossible. This result in a focus on one of the constraints, either time or cost.

2.1.2. Project Life Cycle

According to PMI (2004), Organizations performing projects usually divide each project into several project phases each marked with a completion of one or more deliverables mainly for management control. A deliverable is a tangible, verifiable work product such as a feasibility study, a detailed design, or a working prototype. Project Life Cycle serves to define the beginning and the end of a project or Project Life Cycle - is the collective name for the different project phases.

PMI 2004 adopted Morris's life cycles for construction. According, Morris identified the following Project Life Cycles for Construction Projects:

- *Feasibility* – project formulation, feasibility studies, and strategy design and approval. A go/no-go decision is made at the end of this phase.
- *Planning and Design* – base design, cost and schedule, contract terms and conditions, and detailed planning.
- *Construction* – manufacturing, delivery, civil works, installation, and testing. The facility is substantially completed at the end of this stage.

- *Turnover and Start-Up* – final testing and maintenance. The facility is in full operation at the end of this phase.

2.1.3. Construction Project Planning

Planning in general, and construction planning in particular, plays a pivotal role in the actualization of a construction project. According to Hendrickson (2000), Construction planning is a fundamental and challenging activity in the management and execution of construction projects, which involves the choice of technology, the definition of work tasks, the estimation of the required resources and durations for individual tasks, and the identification of any interactions among the different work tasks. Adequate preconstruction planning is recognized as essential to limit the potential for later construction delays and cost overruns. In developing a construction plan, it is common to adopt a primary emphasis on either cost control or schedule control (Hendrickson, 2000). Some projects are primarily divided into expense categories with associated costs. In these cases, construction planning is cost or expense oriented. Within the categories of expenditure, a distinction is made between costs incurred directly in the performance of an activity and indirectly for the accomplishment of the project. For other projects, scheduling of work activities over time is critical and is emphasized in the planning process. In this case, the planner ensures that the proper precedence among activities is maintained and that efficient scheduling of the available resources prevails. However, most projects require consideration of cost and schedule over time, so planning, monitoring and record-keeping must consider both dimensions

Construction planning is more difficult in some ways since the building process is dynamic as the site and the physical facility change over time as construction proceeds. According to Hendrickson (2000), from the standpoint of construction contractors or the construction divisions of large firms, the planning process for construction projects consists of three stages that take place between the moments in which a planner starts the plan for the construction of a facility to the moment in which the evaluation of the final output of the construction process is finished.

The ***estimate stage*** involves the development of a cost and duration estimate for the construction of a facility as part of the proposal of a contractor to an owner. It is the stage in which assumptions of resource commitment to the necessary activities to build the facility are made by a planner. A careful and thorough analysis of different conditions imposed by the construction project design and by site characteristics is taken into consideration to determine the best estimate. The success of a contractor depends upon this estimate, not only to obtain a job but also to construct the facility with the highest profit. The planner has to look for the time-cost combination that will allow the contractor to be successful in his commitment. The result of a high estimate would be to lose the job and the result of a low estimate could be to win the job, but to lose money in the construction process. When changes are done, they should improve the estimate, taking into account not only present effects but also future outcomes of succeeding activities. It is very seldom the case in which the output of the construction process exactly echoes the estimate offered to the owner.

In the ***monitoring and control stage*** of the construction process, the construction manager has to keep constant track of both activities' durations and ongoing costs. It is misleading to think that if the construction of the facility is on schedule or ahead of schedule, the cost will also be on the estimate or below the estimate, especially if several changes are made. Constant evaluation is necessary until the construction of the facility is complete.

The ***evaluation stage*** is the one in which the results of the construction process are matched against the estimate. It is in this last stage of the planning process that he or she determines if the assumptions are correct. If they were not or if new constraints emerge, he/she should introduce corresponding adjustments in future planning.

2.1.4. Concepts of Construction Duration

There are many definitions of construction duration. This review includes the most widely recognized definitions: Bhokha (1998) as cited by Odabaşi (2009) defines construction duration as: "The time frame given by the owner for the contractor to complete the project under normal work conditions, normal practice of construction, and based on the minimum costs. It starts when the contractor receives the instruction to proceed and ends after construction works on site. It also includes delays caused by unanticipated circumstances, e.g. alteration of works (changed conditions and change orders), extra works, and supply of materials, location, weather, and site work conditions. Major changes that after the scope of work significantly are not included."

Nkado (1995) and Chan and Kumaraswamy (2001) also defined construction duration as: "Construction duration is the period from the moment that the works commences at the site to the completion and handing over of the facility to the client."

Martin et.al (2006) defined building project duration as “the period between the date of client sanction to the project and the date of practical completion”. Whereas, building construction duration was defined as “the period between the date of the construction contract start on site and the date of practical completion”.

Construction durations can be expressed in different forms: Working day, Calendar day, or Completion date (Ahmed Saleh, 2004). A Working day is a time that the contractor will be working on the project, excluding weekends, state-recognized legal holidays, and adverse weather-related non-working days; Calendar day is elapsed time without regard to the contractor’s necessarily being on the job; and Completion date is a specific date in the calendar year by which the project is to be completed.

2.1.5. Conceptual Construction Duration Prediction

Construction duration is identified as one of the most important criteria along with cost and quality for measuring the overall success of construction projects. Completion of projects within the planned time is very critical for successful projects. This is partly because late completion of projects increases the cost of the work and the client loses potential revenue (Mensah, Adjei-Kumi, et al., 2016). The issue of time in construction projects is vital for both the owner and the contractor. It determines the date on which the project will be in use, the cost to be paid, and the amount and density of resources needed to execute the job in the specified time (Al-Sultan, 2004). Construction delay, which is a common problem in the construction sector, may occur due to some factors such as poor productivity, variation of orders, unforeseen weather conditions, or simply due to underestimated construction duration (Helvacı, 2008). From the contractors’ point of view, an underestimation of construction duration leads to reorganization and reallocation of resources which was not planned initially. On the other hand, overestimation of construction duration can also be as bad as an underestimation of construction duration. It may lead to spending extra money on the allocation of extra resources which is unnecessary, in fact (Helvacı, 2008). duration underestimating would raise events of penalty, disputes, etc. for the contractors (Golizadeh et al., 2015). Furthermore, when the project timetable is loose (the progress of the work is behind the schedule), construction workers are more likely to have a low level of arousal, which results in declining in labor productivity (Park et al., 2010; Kim et al., 2011). On the other hand, overestimating may increase the chances of material and equipment damage by bad weather or loss of organizational competitiveness (Le-Hoai et al., 2008). Both of the situations could have counter-productive and undesirable effects on the project performance and achievement of the project objectives.

Time prediction aims to forecast the most likely number of working days required to complete a project (Abdel Aziz, 2009). Construction duration estimations can be made either after the detailed design phase or pre-design stages (William, 2008; Skitmore and Ng, 2006). But, both are required for different purposes. The early estimate allows the construction owner or designer to check the project scope early in the project lifecycle, while the cumulative project cost is relatively low, and the ability to influence the outcomes of the project is relatively high (Hendrickson, 2000). Duration estimations in different stages of construction projects, according to the project data availability and time constraints, are very important for the planning phase of construction (Helvacı, 2008). Krokowski (1992) cited by (Williams 2008) defines the primary objective of these early, conceptual, estimates being: “to provide a basis for the capital budget medium-term plan, and to provide a basis for the definition technical package (±20%) for project feasibility studies and upper management approval.” In pre-design stages, forecasting of construction duration is very difficult with minimum design information (Helvacı, 2008; Odabasi, 2009).

The importance of ensuring the accuracy and reliability of construction time estimates at the tendering stage cannot be over-emphasized (Ahmadu et al., 2015). Such early estimates provide a basis for evaluating the success of a project and the efficiency of the project organization (Nkado, 1995). They also provide useful grounds for ascertaining logistical and cash flow implications for feasibility, budgeting, planning, monitoring, and even litigation purposes (Skitmore and Ng, 2003). Accurate time planning and estimation of construction duration are essential to prevent delays and ensure the timely completion of projects (Tıratacı and Yaman 2023).

In general, construction durations are estimated by using planning and scheduling techniques such as Gannt or bar chart, the Critical Path Method (CPM), and the Program Evaluation and Review Technique (PERT) (Helvacı, 2008). However, these techniques usually require detailed design information for the estimation of activity durations and determination of the sequencing of the activities. In some cases, pre-design duration estimates may be performed by using these techniques, however, the accuracy of these estimates mainly depends on the experience of the planning engineer.

Bromilow (1969) developed the first construction time estimating model, following a pioneering investigation of the time performance of 309 building projects in Australia. The model is often referred to as Bromilow's time-cost model which expresses time as a function of the cost of a building. However, the model had limited applicability due to differences in projects' characteristics and the uniqueness of various construction industries. This necessitated replication of the model by several researchers (Ireland, 1983; Kaka and Price, 1991; Yeong, 1994; Chan, 1999; Ng et al., 2001; Ojo, 2001; Ogunsemi and Jagboro, 2006; Long and Young, 2009; Ameyaw et al., 2012).

The need for improved preliminary construction duration estimates has been noted by numerous authors and research has taken place in a number of construction industry sectors to support the development of these estimates (Williams, 2008). However, industry research suggests that there is a noticeable lack of research on the ability to reliably and historically predict highway construction duration using early known project design details (Williams, 2008). Work has been performed in the residential and building construction arenas internationally [Bromilow 1980, Nkado 1992, Chan and Kumaraswamy 1995, Chan 1999, Skitmore and Ng 2001, Burrows et al. 2005]. However, the domestic highway construction industry has not seen the same level of attention.

2.1.6. Bromilow's Principle

A literature review disclosed that the first significant recorded detailed assessment of the construction time performance of building projects was initiated in Australia in the late 1960s (Chan, 1998). A relationship between completed construction cost and the time taken to complete a construction project was first mathematically established by Bromilow (1974) and subsequently updated by Bromilow, et al.(1980) from a survey of 370 Australian building projects, developed a model which predicts construction time known by Bromilow Time Cost Model (BTC) described in the form of the formula:

$$T = KC^B$$

Where: T is the duration of the construction period from the date of site possession to practical completion, in working days; C is the final cost of building in millions of dollars, adjusted constant labor and material prices; K is a constant describing the general level of time performance for a

$1 million project, and B is a constant describing how the time performance is affected by project size, as measured by cost.

The model indicated that one factor (the scope of the project as measured by construction costs in 1972 Australian dollars) partially predicts construction time. However, it did not incorporate the effects of other factors besides cost and hence was chosen as a first approximation for general scenarios (Chan, 1998).

Following the time-cost model introduced by Bromilow (1969), several studies have been conducted for either building or Civil Engineering projects around the globe. For instance in; Australia (Mak et al., 2000; Walker, 1994), United Kingdom (Kaka & Price, 1991), Hong Kong (Chan, 1999; Chan & Kumaraswamy, 1995), Malaysia (Chan, 2001; Endut et al., 2006), Bangladesh (Choudhury et al., 2002; Rahman et al., 2014), Texas (Choudhury, 2013), Vietnam (Le-Hoai et al., 2013; Long et al., 2009), and Korea (Le Hoai & Dai Lee, 2009). A few studies were found in Africa related to time – cost model; Nigeria (Ogunsemi & Jagboro, 2006); Ghana (Ameyaw et al., 2012; Mensah et al., 2016; Mensah et al., 2016), and (Alemu, 2021; Assefa, 2008; Liben et al., 2024; Menbere, 2019; Menbere, 2022) are specific from Ethiopia.

2.2. Empirical literature review

2.2.1. Factors Affecting Construction Duration

Construction projects are complex and dynamic, with their successful completion hinging on the effective management of various influencing factors. A key indicator of a project's success is its completion within the planned timeframe. Project delays, which directly impact construction duration, are a significant and recurring problem globally, leading to financial losses, disputes, and reputational damage (Doloi et al., 2012). This literature review synthesizes scholarly research from diverse international contexts to identify and categorize the most prevalent factors contributing to extended construction timelines.

Table 2-1 Studies on key factors influencing construction duration

Category	Researcher	Key Factors and Examples
Project Management & Planning	Abdellatif & Alshibani (2019), Pamidimukkala & Kermanshachi (2021), Gündüz et al. (2013), Zidane & Andersen (2018), Aibinu & Jagboro (2002), Doloi et al. (2012), Abdul Aziz & Mohamed Abdel-Hakam (2016), Dissanayaka & Kumaraswamy (1999)	Inadequate planning and scheduling, lack of proper project management procedures, absence of clear project scope, poor communication and coordination among all parties, inefficient site management and supervision.
Client & Owner-Related	Abdul Aziz & Mohamed Abdel-Hakam (2016), Zidane & Andersen (2018), Abdellatif & Alshibani (2019), Sami Ur Rehman et al. (2022), Assaf & Al-Hajji (2006), Pamidimukkala & Kermanshachi (2021)	Slow decision-making by the client, delays in approvals for design documents, permits, and change orders, late or delayed payments to the contractor.
Design & Technical	Abdellatif & Alshibani (2019), Doloi et al. (2012), Sepani Senaratne & Kushan Balasuriya (2017), Abdul Aziz & Mohamed Abdel-Hakam (2016)	Frequent changes in design and scope during construction, and errors or incomplete drawings.
Resource-Related	Adnan Enshassi & Rafiq Choudhry (2009), Hwang et al. (2013), Fashina et al. (2020), Aibinu & Jagboro (2002), Pamidimukkala & Kermanshachi (2021), Zidane & Andersen (2018)	Shortage of skilled labor, inadequate contractor experience, delays in material delivery and supply chain issues, and unavailability of equipment.
Financial & Economic	Adnan Enshassi & Rafiq Choudhry (2009), Gündüz et al. (2013), Sepani Senaratne & Kushan Balasuriya (2017), Aibinu & Jagboro (2002)	Financial difficulties or cash flow problems of the contractor, and economic fluctuations like inflation.
External & Site Conditions	Adnan Enshassi & Rafiq Choudhry (2009), Sepani Senaratne & Kushan Balasuriya (2017), Hwang et al. (2013)	Unforeseen site conditions and poor performance of subcontractors.

2.2.2. Previous studies on factors affecting construction duration

A review of the literature suggests that the construction duration of a project is affected by a vast number of factors and to varying extents. (Nkado, 1995) found that there is no consensus in the literature on the identification of factors that influence planned or actual construction times. However, (Chan & Kumaraswamy, 1995) illustrated the factors affecting construction duration based on the general international literature, observed common construction practices, and survey results. These factors include both qualitative and quantitative contributors. Hence, the construction duration can be regarded as a function of all these factors, that is, construction time = f (all the factors in the hierarchy of Figure (2-1) below). The factors can be reviewed into three levels namely: Primary, Secondary, and Tertiary levels. For instance, the Primary factors include construction cost, type of construction, location, stakeholder's priorities, productivity, type of contract, and post-contractual developments.

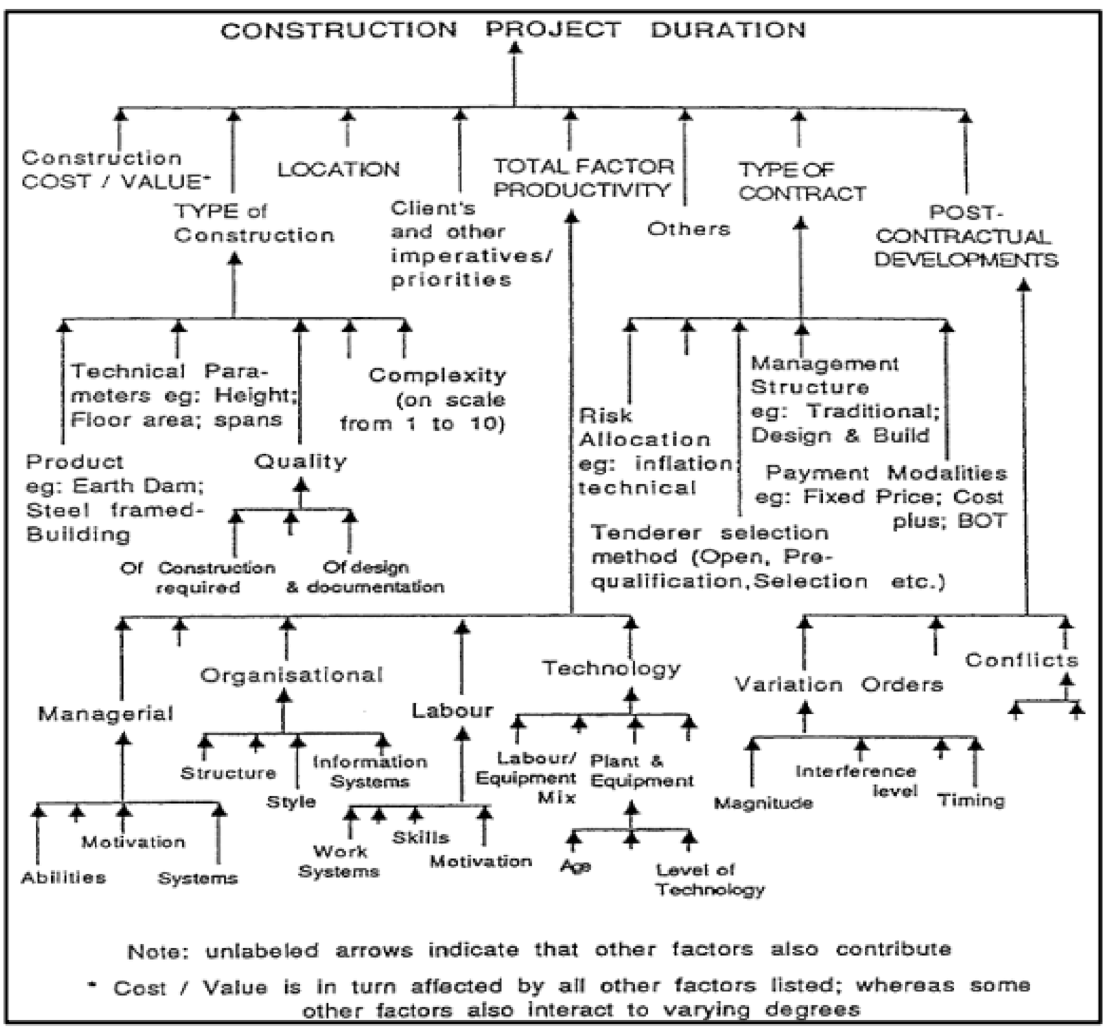

Figure 2-1: Factors affecting construction project duration (Chan and Kumaraswamy 1995)

(Walker, 1995) developed a model to show the significant role the construction management team plays in project duration. After surveying 100 Australian construction project managers, he determined that how quickly a project finished largely depended on the construction manager's skill in handling issues related to project complexity, communication, project scope, and client characteristics. Walker's model (shown in Figure 2-2) positions the management team as a filter for all factors that could affect construction time. Arrows within the model illustrate how different elements influence one another, with the number of positive signs indicating the strength of that influence. The research also pointed out the variability in which factors become important, suggesting that a project manager's unique traits might determine which issues are most critical during construction.

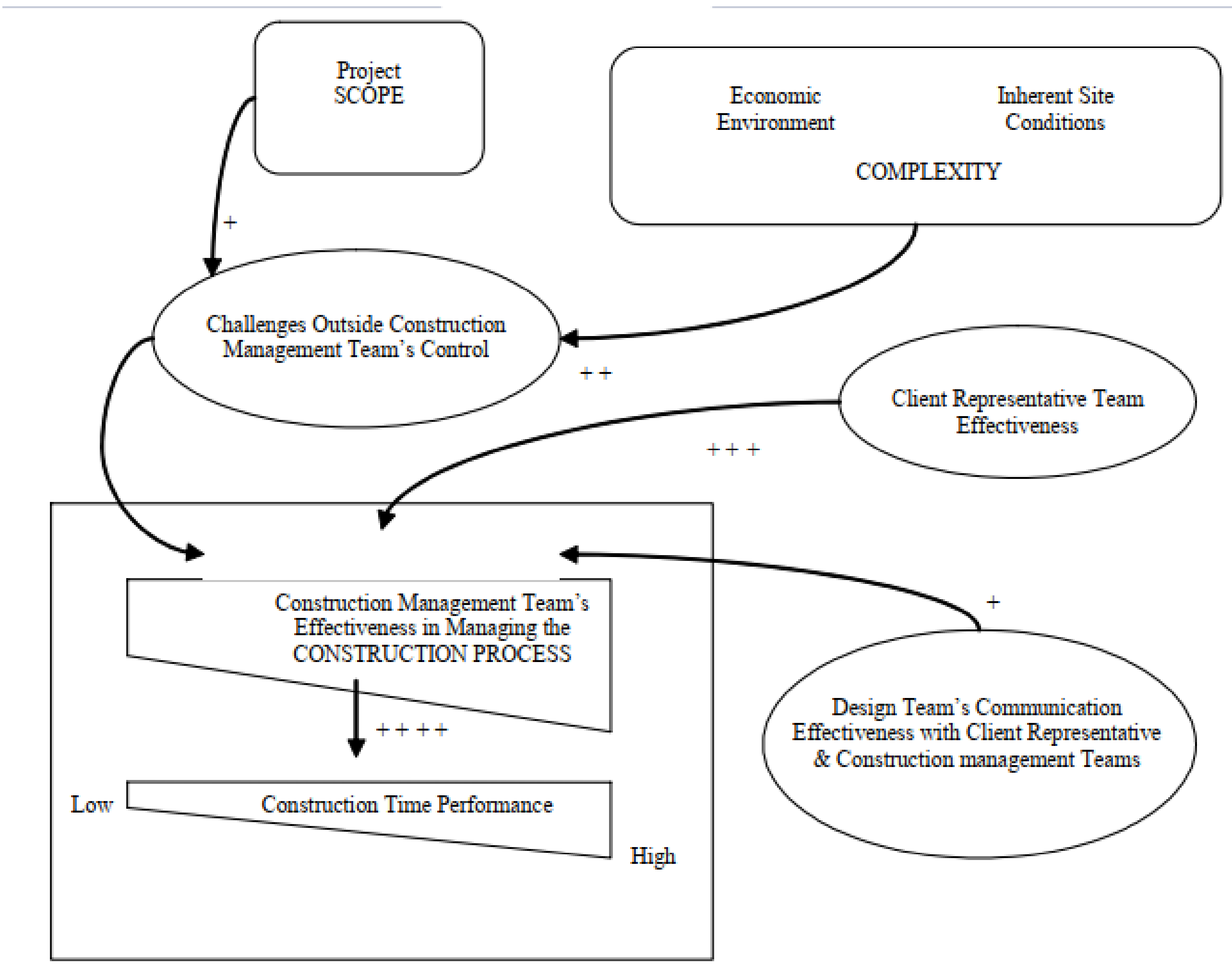

Figure 2-2: Model of Construction Time Performance Causal Factors (D. H. Walker, 1995)

(Chan, 1998)) Classified time influencing factors into four major categories; these include Project scope factors, Project complexity factors, project environment factors, and Management-related attributes. Figure (2-3) shows the major and sub-categories of time-influencing factors identified by (Chan, 1998). The researchers hypothesized this categorization could be a universal model, applicable in Hong Kong and other countries. Currently, it represents the most comprehensive classification of common factors, suitable for diverse project characteristics. This model effectively divides factors into easily identifiable categories, covering both quantitative aspects (such as project scope) and qualitative ones (such as management attributes).

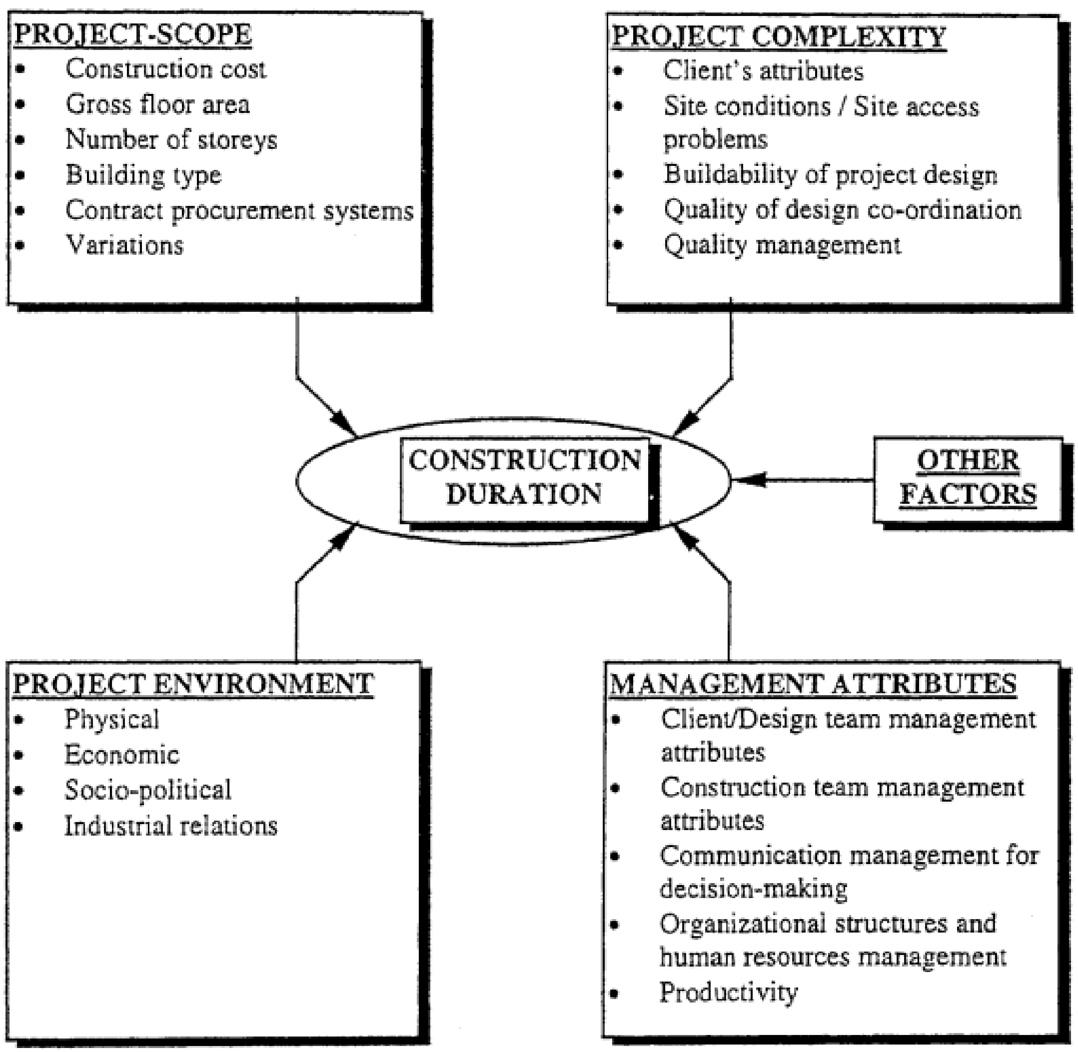

Figure 2-3: Summary of Factors affecting construction project duration (Chan 1998)

There are several common themes identified across the factor-related literature above. A general lack of consensus exists in identifying significant factors influencing construction duration. In spite of this lack of consensus, previous research has been successful in prioritizing these factors in order of relative importance. Previous research has also been successful in subdividing these factors into specific categories, most commonly into categories relating to management attributes, project scope, environment, and design issues. The challenge in the development of a duration estimation model lies in determining which of these factors can be identified and modeled using parameter values (Hoffman, 2005). However, it should be noted that these all factors should not be engaged during time model development. Therefore, selection of among those factors should be done so that one can use those selected factors for model development. The factors must be easily identifiable or known during the planning phase of a project to be useful in predicting construction durations at an early stage. There are various approaches in the selection of variables for time prediction model development. According to (Nkado, 1995) time influencing factors need to be prioritized if used as a basis to model construction time. Such prioritizing the influential factors will assist planners in deciding the factors that will be incorporated in the model development for construction time and hence maintain a practical prediction model development. The other way of selecting variables for the consumption of time prediction purpose is based on the commonly accepted factors from previous literature reviews (Williams, 2008).

2.2.3. Duration Prediction Models for Highway Projects

Several existing models can be found in the literature that attempt to predict the duration of a construction project, which are based on the use of different variables and determining factors. In conceptual estimates, methodologies are usually developed with historical information, which may include basic elements of the projects defined as parameters. Then, the methodology developed to forecast the construction duration becomes a parametric method. The parametric method uses project characteristics to form a model to forecast project duration for future use. The model is developed by establishing relationships between the parameters and the project duration. This section provides a review of Bromilow's Time-Cost modeling and other parametric duration prediction models for highway projects established in prior studies.

2.2.3.1. Time-Cost Models

Time-cost models are primarily used for preliminary estimation and high-level forecasting. They are particularly useful during the early stages of a project when detailed scheduling and resource allocation are not yet available. Bromilow's time-cost model is the widely accepted time-cost relationship for construction time prediction. Many scholars have examined the applicability of the model for their specific project type and country context. However many researchers emphasized building projects as compared to highways and other civil engineering projects. (Car-Pušić & Radujković, 2010) tried to examine the applicability of the BTC model for road projects in Croatia based on a sample consisting of 27 roads that were built in the past ten years in Croatia. Time - Cost relationship for both effectuated and contract values were developed with R^2 values of 0.68 & 0.46 respectively and the BTC formula is given by;

Contract value: $T = 67*C^{0.54}$

Effectuated value: $T = 66*C^{0.45}$

(Tawiah, 2015) sought to develop a model for predicting the duration of feeder road construction projects in Ghana. The study covered 70 road projects completed from 2011 to 2012 in two regions of Ashanti and Brong Ahafo. The result showed that Bromilow's time-cost (BTC) model is applicable to the Department of Feeder Road (DFR) projects and is of the form:

$$T = 64C^{0.134} \qquad (R^2 = 0.187, F = 13.368, p = 0.001)$$

The study by (Czarnigowska & Sobotka, 2013) affirmed Bromilow's time-cost relationship through an analysis of 100 public road projects in Poland. Their research also benchmarked this model against other multifactor approaches, including statistical regression and regression trees. The model's applicability was supported by its adjusted R^2 value of 0.636, a standard error (SEE) of 0.504, and a mean absolute percentage error (MAPE) of 44.84%. The resulting relationship is defined as:

$$\text{Ln } L = 1.2067 + 0.4649*\ln C$$

Where L is the number of working days from the contractor's possession of the building site to the completion of works, C is the actual value of works as paid by the client

(Pokhrel et al., 2021) developed empirical formulae for the determination of contract duration based on the historical data of 78 roads completed within the last five years. The result showed that Bromilow's time-cost (BTC) model is applicable for Road projects in Nepal and is of the form:

$$T = 92.482 \times C^{0.2159} \qquad (adj.R^2 = 0.52503)$$

(Menbere, 2019) also developed early-time prediction models for road construction projects for road construction projects undertaken by the Ethiopian Road Authority. The models were developed based on 30 sets of data collected and it has shown the relationship.

$$TIME = 0.0812 * COST^{0.57} \qquad (R^2 = 0.831)$$

(Menbere, 2022) established Time-Cost relationship based on the BTC model for the 30 road projects in Ethiopia specific to the Amhara region, the model is given;

$$T = 0.14C^{0.43} \qquad (R^2 = 0.824)$$

(Waziri & Nurudeen, 2014) investigated the applicability of Bromilow's Time-Cost (BTC) model for predicting highway construction time in Nigeria, using data from a database of completed projects. The model demonstrated a good fit to this data, evidenced by an R-value of 0.736 and an R^2 value of 0.542. However, the model showed weak prediction efficacy with a Mean Absolute Percent Error (MAPE) of 19% on a test sample, which they considered inadequate for practical application. Based on their analysis, the Time-Cost model for the data was:

$$T = 2.80C^{0.5352}$$

(Adeyemi & Motlakase, 2020) developed a Bromilow's time-cost model for road construction projects in Botswana, utilizing data from 54 projects completed by the Department of Roads. The model demonstrated an R^2 of 0.283266 and an adjusted R^2 of 0.269483. Despite these values indicating a good data fit and falling within previously observed global ranges (0.205-0.850), the researchers suggested that the model might not be sufficiently vigorous for accurately estimating project duration early in its lifecycle based solely on cost. From the analysis, the time-cost model was given by:

$$T = 14.111C^{0.156}$$

(Purnus et al., 2022) developed a mathematical Time-Cost model for road infrastructure projects in Romania. Their study categorized these projects into four types: highways, road rehabilitation, road modernization, and bypasses. Implementation times were estimated using optimistic, most likely, and pessimistic scenarios. The parameters and equations for the Bromilow's Time-Cost model are presented in Table (2-1).

Table 2-2: Bromilow's Time-Cost model for the highway projects (Purnus, Bodea, & Stoian, 2022)

Scenario	Regression equation	R^2	K	B	Bromilow's form
Optimistic	y = 0.3067x + 4.4749	0.6867	87.7858	0.3067	$T = 88C^{0.307}$
Most likely	y = 0.3904x + 4.2948	0.6845	73.3175	0.3904	$T = 73C^{0.390}$
Pessimistic	y = 0.4001x + 4.4874	0.7047	88.8900	0.4001	$T = 89C^{0.400}$

(Gautam & Banstola, 2022) established a Bromilow's time-cost model to estimate contract duration for urban road projects in Pokhara Metropolitan City (PMC). Their study used data from 28 similar urban road projects, each with an estimated cost of ten million Nepalese Rupees. The model, which had a coefficient of determination (R^2) of 0.5093, accounted for 50.93% of the variability in the road project data. The researchers also reported the model's predictive errors: a maximum error of 50.58%, a minimum error of -94.54%, and an average error of 7.76%. The developed Bromilow's time-cost model was:

$$T = 48.202C^{0.3759}$$

2.2.3.2. Simple Regression Models

Simple regression models are characterized by their use of a single independent variable to explain the variation in a dependent variable. This approach is most effective when the relationship is straightforward, and other factors are not considered significant to the analysis. The following models are all considered simple regression models because they rely on a single independent variable (X, representing cost or another key factor) to predict a single dependent variable (Y, representing time). The coefficients $\beta0$and $\beta1$ represent the **y**-intercept of the regression line and the slope of the regression line respectively. ϵ is the error term, representing the difference between the observed value and the predicted value.

- **Linear Regression**: This model posits a direct, straight-line relationship between project cost and duration. It is used when the change in duration is expected to be constant for every unit increase in cost. For instance, a linear model might be used to analyze the relationship between the number of rooms and the construction time for a series of identical housing units.

$$(y = \beta_{0+}\beta_1 * X)$$

- **Logarithmic Regression**: This model is applied when the effect of an independent variable diminishes as its value increases. In a time-cost context, it can model a situation where an initial increase in project cost (e.g., due to more resources) significantly reduces duration, but subsequent, larger cost increases have a progressively smaller impact.

$$(y = \beta_{0+}\beta_{1*}LnX)$$

- **Inverse Regression** (Y=b0+b1X1): This model is suitable for relationships where the dependent variable changes rapidly at first and then levels off. For example, it can be used to model the relationship between the size of a project team and project duration. Adding the first few team members significantly shortens the schedule, but adding more has a marginal effect as coordination complexities increase.

$$(y = \beta_{0+}\beta_1/X)$$

- **Quadratic Regression**: This model captures relationships with a single curve, representing a parabolic trend. It is useful for projects where an increase in a variable, such as complexity, first leads to a decrease in efficiency (and thus longer duration), and then efficiency increases again as specialized resources are brought in.

$$(y = \beta_{0+}\beta_1 * X + \beta_2 X^2)$$

- **Cubic Regression** :As a more flexible polynomial model, cubic regression can capture relationships with two bends. It is used for more complex, non-linear relationships that exhibit multiple phases of acceleration or deceleration, which may occur in large, multi-stage infrastructure projects.

$$(y = \beta_{0+}\beta_1 * X + \beta_{2*}X^2 + \beta_3 * X^3)$$

- **Exponential and Growth Regression**: These models are appropriate for data that exhibits rapid, accelerating growth. In construction, this could be applied to model the cost-time relationship for highly innovative projects where early development is slow, but progress accelerates rapidly after a critical breakthrough.

$$(y = \beta_{0*}\, e^{\beta 1x})$$

- **Power Regression**: This model is used when the dependent variable is proportional to the independent variable raised to a power. The relationship between project cost and duration is often best described by a power function, such as the widely used Bromilow's Time-Cost Model. The model can represent a wide range of curves, including those with accelerating or decelerating trends.

$$(y = \beta_{0*}\, X^{\beta}1)$$

- **S-Curve Regression**: This model is useful for processes that start slowly, accelerate, and then level off as they approach a maximum capacity. It is highly relevant to construction, as project progress often follows an S-shaped curve: slow initial mobilization, followed by a period of rapid construction, and finally a leveling off during the final finishing and closeout stages.

$$(y = e^{(\beta 0 + \frac{\beta 1}{X})}$$

- **Compound Regression**: Similar to exponential growth, this model is used for data that increases or decreases by a constant percentage over time. It can be used to model the growth of a project's cost over its duration due to factors like inflation or compounding interest on a loan.

$$(y = \beta_{0*}\, {\beta_1}^{X})$$

2.2.3.3. Multiple regression model

Multiple regression model is a powerful statistical technique used in highway projects to predict a project's duration by considering multiple influencing factors simultaneously. Instead of relying on a single variable like cost, this method provides a more accurate and realistic model by incorporating several independent variables. Multiple regression model can be expressed as:

$$Y=\beta 0+\beta 1X1+\beta 2X2+...+\beta nXn+\epsilon$$

In this equation, Y represents the dependent variable (e.g., project duration). (X1, X2, Xn) is the independent variable (e.g., length of the road, number of bridges or culverts, terrain complexity) and The β coefficients determine the unique impact of each factor on the project duration, while holding other factors constant.

By using this approach, engineers and project managers can move beyond simple estimations and gain understanding of how different variables contribute to a project's timeline. This allows for more informed decision-making, better resource allocation, and more effective risk management throughout the project lifecycle.

2.2.3.4. Studies on Parametric Duration Modeling for Highway Projects

(Attal, 2010) studied the development of Neural Network Models for Prediction of Highway Construction Cost and Project Duration. Two distinct mathematical techniques Step-Wise Regression Analysis (SRA) and Artificial Neural Networks (ANNs) developed in this effort to come up with reliable models of highway duration prediction. In addition, the given data were classified and analyzed for full-depth section and improvement of the highway.

Finally (Attal, 2010) concluded that the ANN outcome represented higher accuracy and reliability than linear regression analysis as shown in Table (2-2) below.

Table 2-3: Comparison of SRA and ANN highway duration prediction models (Attal 2010)

Project types	Artificial Neural Networks			Regression Analysis		
	R	R^2	Error	R	R^2	R^2 adj.
Highway full-depth section duration prediction models	0.9158	0.8575	0.0013	0.8972	0.8049	0.7854
Highway improvement duration prediction models	0.8355	0.6981	0.01341	0.8036	0.6457	0.55721

(Assefa, 2008), M.Sc. thesis, studied time – cost relationship for public road projects in Ethiopia based on a sample of 33 Federal road projects completed between 1997 and 2008. The study identified the contractor origin and pavement type as a main category to classify the projects. Based on the contractor origin, projects were categorized into two; projects carried out by International Contractors (IC) and Domestic Contractors (DC). Based on the pavement type projects were classified into Asphalt Concrete Surfaced Road Project (AC) and Double Surface Asphalt Treatment Road Project (DBST). Among the 33 projects, 21 were IC projects and 12 were DC projects. Out of the 21 IC projects, 14 were AC Surface projects and 7 were DBST surface projects. All the 12 DC projects were Gravel Surface projects. The result of the study is expressed with the formulas below:

IC - AC Road Projects: $Log\ (T) = -\ 47\ .058 + 8\ .7640\ Log\ (C) -\ 0\ .0394\ (Log(C))^3$

IC - DBST Road Projects: $Log\ (T) = -\ 32\ .759 + 6\ .1661\ Log\ (C) -\ 0\ .0268\ (Log\ (C)\ ^3$

DC – GS Road Projects: $T(C) = 15\ .0615 + 20\ .2224\ (L) -\ 0\ .2631\ (L)^2 + 0\ .0013\ (L)^3$

Where T – Time in calendar days, C – Cost in Ethiopian Birr and L –project length in km

(Jiang & Wu, 2007) analyzed Indiana Department of Transportation (INDOT) data to predict highway construction contract times, exploring the relationship between project cost and duration. They utilized project region, highway type, and weather conditions as key prediction parameters. Their regression analysis indicated that contract time could be predicted with a 95% confidence interval (CI). They further adjusted their model to account for significant factors like road type, location, traffic volume, and execution season on estimated time.

(Williams, 2008), Ph.D. Dissertation, Developed Mathematical Models for Preliminary Prediction of Highway Construction Duration for Full-Depth Section and Highway Improvement project types as illustrated in table (2-3). Nine statistically significant project factors were shown to influence construction duration. Three of these factors are continuous or quantitative variables (Traffic Volume (ADT), New Signal Count, Construction Cost Estimate), and the remaining six factors (District, Curb & Gutter, Median, Area Location Name, State Highway System, Geometric Design Standard) are categorical or qualitative.

Table 2-4: Highway construction duration prediction models (Williams 2008)

Project Type	Section	Models
Full-Depth Section Results	Full Model	$DURATION = y = \beta_0 + \beta_1 \bullet ADT + \beta_2 \bullet SigCnt + \beta_3 \bullet CG_1 + \beta_4 \bullet M_1 + \beta_5 \bullet Dist + \beta_{6i} \bullet GeoStd + \beta_7 \bullet (g_3 \times ADT)$
	Condensed Model	$DURATION = y = \beta_0 + \beta_1 \bullet CostEst + \beta_2 \bullet HWSystem + \beta_3 HWSystem \times CostEst$
Highway Improvement	Full Model	$DURATION = y = \beta_0 + \beta_1 \bullet ADT + \beta_2 \bullet SigCnt + \beta_3 \bullet M_1 + \beta_{4i} GeoStd$
	Condensed Model	$DURATION = y = \beta_0 + \beta_1 \bullet CostEst + \beta_2 \bullet Area$

Where; DURATION = Construction Duration (working-days); ADT = Traffic Volume (vehicles/day); SigCnt = New Signal Count; CG1 = 1 if roadway has curb & gutter, 0 otherwise; GeoStd = 1 if GeoStd is GS-2 through GS-8, 0 otherwise; M1 = 1 if roadway has median, 0 otherwise; Dist = 1 if District is Staunton, 0 otherwise; g3 = 1 if GeoStd is GS-4, 0 otherwise; CostEst = ln(PCES Construction Cost Estimate (2008 $M)); HWSystem = 1 if Interstate, 0 otherwise; Area = 1 if Medium Urban (50,000 – 199,999), 0 otherwise

(Waziri et al., 2014) developed three multiple regression models: linear, semi-log, and log-log transformations. Analysis showed all three models were statistically significant and fit the data well, with R^2 values of 0.546, 0.631, and 0.940, respectively. The log-log model outperformed the others, exhibiting an average percentage Error of -3.64%, a maximum error of 16.2%, and a Mean Absolute Percent Error (MAPE) of 6.87%.

(1) Linear Regression Model: *Duration =283.33 + 0.056Road Length + 0.494No. of Curvets + 0.375Road Thickness + 0.126Cost per Unit length*

(2) Semi-Log Regression: *Ln Duration = 2.212 + 0.0015Road Length + 0.312 No. of Curvets + 0.589Road Thickness + 0.016Cost per Unit length*

(3) Log-Log Regression: *Ln Duration = 1.685 + 0.013*ln Road Length + 0.028*ln No. of Curvets + 0.965*ln Road Thickness + 0.067*ln Cost per Unit length*

(Menbere, 2022) also studied the Construction Duration Prediction Model in the Case of a Highway Construction Project in Amhara Region, Ethiopia. The researcher developed six significant duration models, as illustrated in Table (2-4), which consist of four simple regression models, multiple linear regressions, and an ANN model based on a sample of 30 highway projects. Finally, the researcher (Menbere, 2022) concluded that the Artificial Neural Network (ANN) Model resulted in the best-fit duration model with a high degree of accuracy compared to other models.

(Pokhrel et al., 2021) developed empirical formulae for determining contract duration using historical data from 78 road projects. Beyond project cost, their study considered vehicle accessibility, construction environment, and work complexity as significant factors for predicting road project duration.

$$T = \frac{I}{I_0 \times 0.64} C_1 \times C_2 \times 92.482 \times C^{0.2159}$$

T= Contract Duration in Calendar days, I= Present /Latest NRB index Value, Io= NRB index value of F/Y 2019/20 (Feb), C1= Coefficient for vehicle accessibility and construction environment, C2= Coefficient for complexity of work, C= Cost of Construction work in Lakhs in Nrs.

(Irfan et al., 2011) also developed duration prediction models for highway and bridge construction projects in Indiana, USA. Their models incorporated anticipated project cost and contract type as key variables. Employing log-linear for logistic modeling, they derived distinct models for five specific project types, which are presented in detail in table (2-5).

Table 2-5: Summary of Construction Duration Prediction models (Menbere 2022)

No	variable	R^2	MAPE	MSE	RMSE	Analysis technique	mathematical model equation
BTC	Cost	0.82	21.7%	0.06	0.25	Simple Linear Regression	$T = 0.14C^{0.43}$
SLR 1	Cost	0.82	21.7%	0.06	0.25	Simple Linear Regression	$\mathrm{Ln}\,T = 0.43\mathrm{Ln}C - 1.98$
SLR 2	Financier	0.06	56%	0.33	0.57	Simple Linear Regression	$\mathrm{Ln}\,T = -0.803\,FP + 7.58$
SLR 3	Payment type	0.24	44.7%	0.27	0.52	Simple Linear Regression	$\mathrm{Ln}\,T = -0.612\,PT + 7.23$
SLR 4	Road length	0.82	19%	0.06	0.25	Simple Linear Regression	$\mathrm{Ln}\,T = 0.3\mathrm{Ln}RL + 5.85$
SLR 5	Road width	0.03	57.6%	0.34	0.58	Simple Linear Regression	$\mathrm{Ln}\,T = -0.3\,\mathrm{Ln}RW + 7.4$
SLR 6	Tender type	0.34	40%	0.22	0.47	Simple Linear Regression	$\mathrm{Ln}\,T = -0.704TT + 7.2$
SLR 7	weather condition	0.08	54%	0.32	0.57	Simple Linear Regression	$\mathrm{Ln}\,T = -0.357\,WC + 7.04$
MLR	All variable	0.89	14.7%	0.05	0.22	Multiple Linear Regression	$\mathrm{Ln}T=2.602+0.214\mathrm{Ln}C-0.47FP-0.19PT+0.13\mathrm{Ln}RL+0.03\mathrm{Ln}RW-0.02TT-0.031WC$
ANN	All variable	0.93	16.3%	0.05	0.22	Non-Linear Regression	Output= 0.85*target + 1.1

Where; C-Cost of the project, FP-financer of the projects, PT- pavement type, RL-road length, RW- road width, TT-tender type and WC –Weather condition

Table 2-6: Project duration models (Irfan, Khurshid et al. 2011)

Project types	**Model Equations**
Road maintenance projects	$y = e^{4.87+0.299*COST+0.268*CONTRACT\_TYPE}$
Road construction projects	$y = e^{4.70+0.307*COST+0.237*CONTRACT\_TYPE}$
Road resurfacing project	$y = e^{4.60+0.340*COST + 0.253*CONTRACT\_TYPE}$
Traffic projects	$y = e^{4.57+0.287*COST+0.409*CONTRACT\_TYPE}$
Bridge construction projects	$y = e^{4.43+0.345*COST+0.367*CONTRACT\_TYPE}$

Where Y is the duration in days

The study (Titirla et al., 2021) developed and compared linear regression (LR) and neural network (NN) models for predicting actual project duration for 37 Greek highway projects, using data available at the bidding stage. While the LR model offered simplicity by using only two variables (initial cost and duration), the NN model, though more complex with 12 variables (including initial cost, initial duration, length, lanes, technical projects, bridges, tunnels, geotechnical projects, embankment, landfill, land requirement and tender offer), demonstrated superior predictive accuracy, yielding a Mean Squared Error (MSE) of 1.53×10^{-6} compared to the LR model's MSE of 1.87×10^{-3} and 72% predictability.

(Al-sadi AM et al., 2017) developed an Artificial Neural Network (ANN) model to predict road project duration in Iraq. This model incorporated six factors: road length, number of lanes, and number of intersections, earthworks, pavement type, and furniture level. Their study reported a 90.6% correlation between actual and predicted durations, with an accuracy of 74.27%. However, the model's applicability was limited to projects constructed by the Roads and Bridges Directorate in Iraq between 2011 and 2012.

(Kaleem et al., 2014) developed a Multi-Linear Regression Model (MLRM) to estimate the duration of highway projects based on data from 120 highway projects in Pakistan completed between 2001 and 2012. Their model primarily utilized project type and planned cost, factors available at the project's initial stage. Additionally, the study assessed the influence of project location on time risk using data from four Pakistani regions, concluding a weak correlation between road duration and geographic location. The models exhibited a Mean Absolute Percentage

Error (MAPE) ranging from 20% to 40%, indicating significant overestimation or underestimation in project time.

(Czarnigowska A & A., 2014) utilized data from 100 public road projects completed between 2003 and 2008 in southeastern Poland to predict road duration at the visibility stage. Their study adopted factors like cost, length, culverts, bays, and intersections. They developed three models: simple regression, multifactor regression, and a regression tree. While these models were deemed statistically correct, their precision was noted as low.

(Son et al., 2019) developed a planning-phase construction time estimation model for highway projects. They used linear regression models built on data from 623 projects completed by the Texas Department of Transportation's (TxDOT) Dallas District Office. The developed multiple linear regression model incorporated construction cost, work type, county size, and highway system as predictor variables. The study showed the new model provided a better fit to their data than the established Bromilow's Time-Cost (BTC) model.

(Nevett et al., 2021) used multiple linear regression analysis to estimate the duration of 1,500 highway projects at the early planning stage. Their model incorporated variables including cost data, traffic volume, terrain type, project condition, project size, and construction quantities, achieving a Mean Absolute Percentage Error (MAPE) of 44%.

(Peško et al., 2017) conducted a study to conceptually estimate the duration of 166 urban road projects during the tender offer phase. Their models, which considered factors like construction material quantities, cost categories, and work distribution, achieved a Mean Absolute Percentage Error (MAPE) of 26.26% using an Artificial Neural Network. A slight improvement in accuracy was observed with their Support Vector Machine Model (SVM), which yielded a MAPE of 22.77%.

(Mahamid, 2019) developed five multiple regression models, presented in table (2-6), for preliminary prediction of road construction duration in the West Bank, Palestine, using 112 datasets. Three models incorporated bid quantities, while two used road length and width as predictors. These models showed strong correlations (R^2 values from 0.88 to 0.93) and Mean Absolute Percentage Error (MAPE) ranging from 19.1% to 31.4%.

Table 2-7: Summary of road project duration models (Mahamid 2019)

Model No.	Mathematical model	Variables
1	*Project duration (days)* = $-14.37 + 0.05x_1$	x_1 = Road length (m)
2	*Project duration (days)* = $-73.74 + 0.05x_1 + 10.27x_2$	x_1 = Road length (m). x_2 = Road width (m).
3	*Project duration (days)* = $14.48 + 0.005x_1x_2$	x_1 = Road length (m). x_2 = Road width (m).
4	*Project duration (days)* = $11.26 + 0.0047x_3 + 0.0038x_4 + 0.005x_5$	x_3 = Earthwork quantity (m^3). x_4 = Basecoarse quantity (m^3). x_5 = Asphalt quantity (m^2).
5	*Project duration (days)* = $16.00 + 0.007x_5$	x_5 = Asphalt quantity (m^2)

The research by (Mohamed & Moselhi, 2022) introduced a new method for estimating the conceptual cost and duration of public highway projects using an ensemble of machine learning (ML) models, trained on data from road, bridge, and drainage projects completed between 2004 and 2015. They considered various parameters, including facility type, project scope, highway type, length, width, location, technical complexity, payment, and procurement methods. The method demonstrated its effectiveness by achieving a Mean Absolute Percentage Error (MAPE) of 7.4% for duration and 4.5% for cost.

The study by (HIEP et al., 2023) utilized regression analysis to determine project duration. They employed various models to estimate the duration of road projects, specifically choosing them to explore the relationship between project duration and cost. This analysis also considered project characteristics like geometric parameters (area) and the economic indicator (Gross Domestic Product per capita, GDP). To normalize the data and enhance result interpretability, the researchers transformed the data using a natural logarithm. Ultimately, their analysis revealed a strong correlation between project cost, area, gross national product per capita, and project duration. Finally, the study recommended Model 1 as the most reliable, given its adjusted coefficient of determination (R^2) of 0.76. This model had the following form:

$$Ln(dur) = 0.4Ln(cost) + 0.4Ln(area) - 005Ln(gdp) - 45$$

Or by simplifying:

$$Ln(dur) = 0.4Ln(cost * area) - 005Ln(gdp) - 45$$

(Aswed et al., 2022) developed two mathematical models—the Nonlinear Regression Model (NLRM) and Multi-linear Regression Model (MLRM)—to predict road project execution time in Iraq. Their models incorporated quantities of base layer length, width, earthwork, sub-base and final project cost. The NLRM demonstrated superior accuracy, exhibiting an R^2 of 88.6% compared to MLRM's 76%. Furthermore, NLRM showed lower error rates (MAPE of 11.65% vs. 20.29% for MLRM; RMSE of 45.73 vs. 76.59 for MLRM) and higher average accuracy (88.35% vs. 79.71% for MLRM).

Table (2-7) below summarizes some of the reviews of the literature on the methods, parameters used, and findings of different highway duration prediction models.

Table 2-8: Summary of the reviews of literatures on highway duration prediction models

Author/s	Methods	Predictive parameters	Findings
(Assefa, 2008)	Linear and Multiple regression technique	Project cost, Project length , Type of construction, Contractor origin	Average percentage error of 2 %.
(Attal, 2010)	Step Wise Regression Analysis (SRA) and Artificial Neural Networks (ANNs)	Location, Type of area, Type of highway system, Type of geotechnical design, Construction length, Curb and gutter, Daily average traffic, Geometric design standard, Cross count, New signal count, Contain median, Contain sidewalk, Loops and Ramps	• ANNs resulted in high accuracy. • Highway full depth section with R^2 of 0.8575, error of 0.0013 • Highway improvement with R^2 of 0.6981, error of 0.01341
(Williams, 2008)	Multiple linear regression model	Traffic Volume (ADT), New Signal Count, Construction Cost Estimate, District, Curb & Gutter, Median, Area Location Name, State Highway System, Geometric Design Standard	• R^2 adj = 0.80 and 0.75 for Full-Depth Section full and condensed models, respectively • R^2adj = 0.64 and 0.55 for Highway Improvement full and condensed models, respectively

(Menbere, 2022)	Simple regression models, multiple linear regressions, and ANN model	. Cost, weather condition, tender type, pavement type, financer, road length and road width.	Artificial Neural Network (ANN) Model resulted in the best fit duration model with R^2=0.934, MAPE=16.3%, MSE=0.05, RMSE=0.22
(Waziri et al., 2014)	Multiple regression models: linear, semi-log, and log-log transformations.	Road Length, No. of Curvets, Road Thickness, Cost per Unit length	The log-log model outperformed with MAPE of 6.87%.
(Titirla et al., 2021)	Linear regression (LR) and neural network (NN) models	initial cost, initial duration, length, lanes, technical projects, bridges, tunnels, geotechnical projects, embankment, landfill, land requirement and tender offer	The NN model demonstrated superior predictive accuracy, Mean Squared Error (MSE) of $1.53\times10{-}6$
(Son et al., 2019)	Multiple Linear Regression	Construction cost, work type, county size, and highway system	MAPE of 35.89%
(Nevett et al., 2021)	Multiple Linear Regression	Cost, Project Characteristics, Average Annual Daily Traffic (AADT), Terrain Type, Project Size, Construction Quantities	MAPE of 44%
(Peško et al., 2017)	Artificial Neural Networks, Support Vector Machine	Quantity of construction materials, Distribution of the work across the different work, categories Project type	Artificial Neural Network a MAPE of 26.26%, and the Support Vector Machine Model (SVM) MAPE of 22.77%.
(Mahamid, 2019)	Multiple Linear Regression	Bid quantities, road length, road width	MAPE ranging from 19.1% to 31.4%.
(Mohamed & Moselhi, 2022)	Ensemble of machine learning (ML) models	Facility type, project scope, highway type, length, width, location, technical complexity, payment methods, procurement methods.	MAPE of 7.4%
(Aswed et al., 2022)	Nonlinear regression model (NLRM), Multi linear regression model (MLRM).	Base layer length, width, earthwork, sub base and final cost of the road project.	NLRM resulted in better accuracy with R^2 of 88.6%, MAPE of 11.65%, root means square errors (RMSE) of 45.73, average accuracy percentages (AA) of 79.71%

2.3. Conceptual Framework of the Study

This section outlines the study's conceptual framework, which provides an integrated approach to analyzing the research problem. Based on a comprehensive review of theoretical and empirical literature, a total of sixteen potential construction duration predictive variables were initially identified. These variables comprised five numerical factors (Actual Cost, Road Length, Road

Width, Lane Number, and Culvert Number) and eleven categorical factors (Highway System, Project Scope, Site Accessibility, Geometric Design Standard, Road Classification, Contractor Category, Delivery System, Contract Method, Availability of Bays, Availability of Intersections, and Availability of Bridges). The selection of these variables was guided by their frequent citation in prior research, their anticipated impact on road project duration, and the practical ease of data acquisition during the early project phase. Table (2-8) presented the complete list of variables initially considered for duration prediction model development. However, a subsequent, more refined selection of these variables was performed based on the existence of a statistically significant correlation with the dependent variable, road duration.

Table 2-9: Initially selected road duration predictive variables with respective references

No.	Road duration predictive variables	Parameter Type	Value	References for variables
1	Actual Cost (AC)	Numerical	Ethiopian Birr (ETB)	(Attal, 2010), (Aswed et al., 2022), (Naik & Kumar, 2015), (Pokhrel et al., 2021), (Nevett et al., 2021), (Son et al., 2019), (Titirla et al., 2021), (Menbere, 2022), (Williams, 2008), (Assefa, 2008), (Czarnigowska A & A., 2014)
2	Road Length (RL)	Numerical	Kilo Meter	(Mahamid, 2019), (Aswed et al., 2022), (Waziri et al., 2014), (Al-saadi et al., 2017), (Titirla & Aretoulis, 2019), (Czarnigowska & Sobotka, 2014), (Mahamid, 2019), (Titirla & Aretoulis, 2020), (Mohamed & Moselhi, 2022), (Menbere, 2022), (Attal, 2010), (Assefa, 2008), (Al-sadi AM et al., 2017), (Czarnigowska A & A., 2014)
3	Road Width (RW)	Numerical	Meter	(Mahamid, 2019), (Aswed et al., 2022), (Mohamed & Moselhi, 2022), (Menbere, 2022)
4	Lane Number (LN)	Numerical	Number (Pcs.)	(Al-saadi et al., 2017), (Titirla & Aretoulis, 2019), (Titirla & Aretoulis, 2020), (Titirla et al., 2021), (Al-sadi AM et al., 2017)
5	Culvert Number (CN)	Numerical	Number (Pcs.)	(Aswed et al., 2022), (Czarnigowska & Sobotka, 2014), (Waziri et al., 2014),
6	Highway System (HS)	Categorical	1 if Asphalt concrete, 0 if Double surface treatment	(Jiang & Wu, 2007), (Son et al., 2019), (Velumani et al., 2021), (Kaleem et al., 2014), (Jiang & Wu, 2007), Al-saadi, Zamiem et al. 2017), (Mohamed & Moselhi,

				2022), (Menbere, 2022), (Williams, 2008), (Attal, 2010), (Assefa, 2008), (Al-sadi AM et al., 2017)
7	Project Scope (PS)	Categorical	1 if new construction, 0 if Rehabilitation	(Qiao et al., 2019), (Mohamed & Moselhi, 2022), (Son et al., 2019), (Hoffman et al., 2007)
8	Site Accessibility (SA)	Categorical	1 if Good access, 0 if Poor access	(Attal, 2010), (Mohamed & Moselhi, 2022), (Williams, 2008), (Nevett et al., 2021), (Hoffman et al., 2007)
9	Geometric Design Standard (GDS)	Categorical	1 if DS1, 2 if DS2,…,6 if DS6	(Nevett et al., 2021), (Williams, 2008), (Attal, 2010)
10	Road Classification (RC)	Categorical	0 if Main Road, 1 if Link Road, 2 if Trunk Road	(Qiao et al., 2019), (Hoffman et al., 2007)
11	Contractor Category (CC)	Categorical	1 if Local contractor, 0 if International	(Qiao et al., 2019), (Assefa, 2008), (Jiang & Wu, 2007)
12	Delivery System (DS)	Categorical	1 if Design bid build (DBB), 0 if Design build (DB)	(Menbere, 2022), (Titirla et al., 2021), (Hoffman et al., 2007)
13	Contract Method (CM)	Categorical	1 if Item rate, 0 if Lump sum Contract	(Mohamed & Moselhi, 2022),
14	Availability of Bays (BY)	Categorical	1 if at least one Bay is available, 0 otherwise	(Czarnigowska & Sobotka, 2014),
15	Availability of Intersection (IN)	Categorical	1 if at least one Intersection is available, 0 otherwise	(Czarnigowska & Sobotka, 2014),(Al-saadi et al., 2017), (Al-sadi AM et al., 2017)
16	Availability of Bridge (BG)	Categorical	1 if at least one Bridge is available, 0 otherwise	(Titirla & Aretoulis, 2019), (Titirla & Aretoulis, 2020), (Titirla et al., 2021),
No.	**Dependent Variable**			
1	Actual Duration (AD)	Numerical	Calendar Days	

2.4. Research gap

The need for improved preliminary construction duration estimates has been noted by numerous authors and research has taken place in several construction industry sectors to support the development of these estimates. In the Ethiopian construction industry, according to (Jekale, 2004) due to the lack of early-time planning tools, the government of Ethiopia waived the use of completion time and allowed a low evaluated cost award system for tender evaluation. The current practice of early duration estimation in Ethiopia's construction industry is largely based on individual experience was highly influenced by the skill, experience, and individual intuition of the planning engineer.

Although there is a high interest and need by both researchers and the construction industry as a whole and sub-sector projects, fewer published studies were obtained in the Ethiopian construction context. In building construction projects (Alemu, 2021) studied the Construction time prediction model for public building project cases in Addis Ababa, Ethiopia. (Liben et al., 2024) also conducted on the Comparison of advanced and traditional machine learning algorithms for construction duration prediction focusing on public building projects in Addis Ababa, Ethiopia. In the case of Road/Highway projects; (Assefa, 2008) studied the Time – Cost Relationships for Public Road Construction Projects in Ethiopia, (Menbere, 2019) also examined the Time - Cost Relationships Model in Road Construction in the case of the Ethiopian Road Authority. Besides (Menbere, 2022) studied the Construction Duration Prediction Model in the Case of a Highway Construction Project in Amhara Region, Ethiopia. Therefore, further investigations on the development of a construction duration prediction model for a specific sector and project type are becoming advantageous to supplement the prevailing practice of estimation predominantly based on individual manager's experience.

Chapter 3: Research Methodology

3.1. Study area

Ethiopia is a landlocked country located in Eastern Africa bordered by Djibouti, Eritrea, Kenya, Somalia, South Sudan, and Sudan. The geography of Ethiopia consists of high plateaus with the central mountain range divided by Great Rift Valley. Ethiopia is the second most populous country in Africa after Nigeria, and the fastest-growing economy in the region. However, it is also one of the poorest, and it aims to reach lower-middle-income status by 2025 (IUCN, 2019).

The study utilized archival project data from thirty (30) road construction projects administered by the Ethiopian Roads Authority (ERA) between 2001 and 2016. The case road projects share some similar features, for instance, the tender used is an Open bid tender, and the owner/financial source of the projects is the Federal Democratic Republic of Ethiopia (FDRE). The road network comprises projects originating from diverse regions across Ethiopia. This regional dispersion is evident through examples such as the Hawassa-Bishan Guracha (Tikur Wuha) project in Southern Ethiopia; Fik-Hamero-Imi Bale and Seru-Shenen Wenz-Sheik Hussein in Eastern Ethiopia; Endaselassie-Dejena-Dansha, Sanja-Keraker, and Azezo-Gonder in Northern/Northwestern Ethiopia; and Nejo-Jarso-Begi and Mizan-Dima-Boma in Western Ethiopia. The complete list of case projects is presented in Annex (1). The projects had design parameters, the road lengths (RL) ranged from 5.38 Km to 148.20 Km, road width (RW) varied from 6.00 meters to 23.00 meters, and the number of lanes (LN) ranged from 2 to 6. The case road projects' weather conditions are highly diverse due to Ethiopia's varied topography, ranging from high mountains to low-lying deserts. The climate is broadly categorized into distinct zones, varying from hot and arid to cool and humid, based on altitude and temperature.

3.2. Research design

The study employed a descriptive and exploratory research design, with the ultimate aim of developing a practical duration prediction model for Ethiopian road projects. This approach involved secondary data analysis, drawing on existing archival project documentation. Specifically, the study first described key project parameters such as duration, length, and cost. It then explored the strength and direction of relationships between construction duration and various project characteristics, including actual cost and road length. This exploratory phase also tested the applicability of an existing model (BTC) and identified influencing factors. Finally, these findings were synthesized to propose a new, practical road duration prediction model.

3.3. Selection of Research Method

This research selected survey and historical case study approaches as its research methods for the realization of the proposed research objectives. This was mainly because the study emphasized the development of a construction duration prediction model for ERA road projects based on historical data of recently completed similar road projects administered by ERA.

3.4. The Research Approach

This study used a quantitative research approach, mainly analyzing secondary historical project documentation. The quantitative approach involves the generation of data in quantitative form, which can be subjected to rigorous quantitative analysis in a formal and rigid fashion. This method directly supported the goal of creating a practical construction duration prediction model for Ethiopian Roads Administration (ERA) projects. The study used extensive historical data from many recently completed ERA road projects, which allowed for the systematic extraction and analysis of quantitative data to identify patterns and relationships, ultimately leading to the development of a new, practical road duration prediction model.

3.5. Research population and sampling technique

For this study, a census sampling approach was chosen over traditional sampling methods. This decision was made because of the limited number of completed road projects administered by the Ethiopian Roads Administration (ERA) over the last fifteen years that had full documentation of all necessary project variables. As a result, data from all 30 available projects within this defined population were collected and used for model development, which eliminated any potential sampling error.

3.6. Data source and collection technique

The data for this study were exclusively secondary, and sourced from the extensive archival project documentation of the Ethiopian Roads Administration (ERA). This approach allowed the research to leverage pre-existing, systematically recorded information from past construction projects. The specific archival documents utilized as data sources included: contract documents, project reports, provisional/final acceptance documents, payment certificates, and drawings. The collected data encompassed a range of project parameters, including actual cost, road length, road width, and other relevant numerical and categorical variables as detailed in the conceptual framework.

However, historical cost estimates and final costs at the time of project completion must be appropriately adjusted to maintain their relevance and accuracy. This is because of construction cost estimates are prepared at a specific point in time, and the prices included are valid only for that particular period unless explicitly adjusted. This limitation arises because the costs of construction-related goods and services are continually affected by market dynamics. These dynamics are primarily driven by two factors: inflation (or, less commonly, deflation) and fluctuations in supply and demand within the construction sector (Assefa, 2008). Typically, these adjustments account for inflation and changes in the tender price index. Inflation is commonly addressed by applying the average inflation rate, which is derived from changes in the Consumer Price Index (CPI). The CPI is a key economic indicator that tracks the average variation in the prices of a standardized set of goods and services purchased by urban households. Each item in this "market basket" is weighted based on household spending patterns during the base period, which currently spans from 2001 to 2016.

Accordingly, the cost data used in this study has been adjusted to reflect price levels as of January 2016. The adjustment was made using CPI data obtained from the Central Statistical Agency of the Federal Democratic Republic of Ethiopia as shown in table (3-1). The current birr value (real value) is determined based on historical costs using the formula provided below:

Table 3-1: Year-on-Year Customer Price Index of Ethiopia (source, Ethiopian Statistics Service Reports https://ess.gov.et/price/)

Year	CPI (Year-on-Year)
2001	92.46
2002	100.00
2003	133.25
2004	164.7
2005	176.99
2006	189.19
2007	207.29
2008	221.03
2009	244.65
2010	278.49

2011	322.52
2012	388.17
2013	492.36
2014	659.22
2015	858.42
2016	495.40

$$Real\ Value = \frac{(Nominal\ Value * CPI(current\ year))}{CPI(past\ year)}$$

3.7. Method of data analysis

The analysis of the collected data was systematically conducted through a series of quantitative statistical methods, aligning with the study's descriptive, exploratory, and model development objectives. All statistical computations were performed using SPSS (version 27). Descriptive statistics were employed to characterize the key project parameters derived from the archival data. This involved computing measure such as frequencies and percentages for categorical variables (e.g., Highway System, Project Scope) and central tendency (mean, median) and dispersion (standard deviation, minimum, maximum) for numerical variables (e.g., duration, actual cost, road length, road width). The exploratory and correlational analyses were performed to investigate the relationships between construction duration and the various project characteristics. Pearson correlation coefficients were calculated to assess the strength and direction of linear relationships between numerical independent variables and the dependent variable (i.e. road duration). The proposed model's predictive power and statistical significance were rigorously assessed using various metrics, including the coefficient of determination (R^2 and Adjusted R^2), overall model significance (F-statistic and its p-value), and the significance of individual regression coefficients (t-statistics and p-values). Additionally, assumptions of linear regression, such as normality of residuals, homoscedasticity, and multicollinearity, were checked using appropriate diagnostic plots and the Variance Inflation Factor (VIF). The model's practical utility was further evaluated using Mean Absolute Percentage Error (MAPE) to quantify prediction accuracy.

3.7.1. Simple Regression Analysis

Simple regression, consist only two variables, one variable (defined as independent) is the cause of the behavior of another one (defined as dependent variable) (Kothari, 2004). Regression can only interpret what exists physically i.e., there must be a physical way in which independent variable X can affect dependent variable Y. The basic relationship between X and Y is given by

$$Y = a + bX$$

In this study, three categories of simple regression were conducted: the Bromilow's Time-Cost model, other best-fit curve estimations of Time-Cost relationships, and Time-Road Length relationships.

3.7.1.1. Bromilow’s Time–Cost model Analysis

A literature review disclosed that the first significant recorded detailed assessment of the construction time performance of building projects was initiated in Australia in the late 1960s. A relationship between completed construction cost and the time taken to complete a construction project was first mathematically established by Bromilow (1974) and subsequently updated by Bromilow, et al.(1980) from a survey of 370 Australian building projects, developed a model which predicts construction time known by Bromilow Time Cost Model (BTC) described in the form of the formula:

$$T = KC^B$$

Where: T is the duration of the construction period from the date of site possession to the practical completion, in working days; C is the final cost of building in millions of dollars, adjusted to constant labor and material prices; K is a constant describing the general level of time performance for a $1 million project, and B is a constant describing how the time performance is affected by project size, as measured by cost.

The Bromilow’s equation can be rewritten in the natural logarithmic form. It has the same shape of the linear equation as follows:

$$LnT = LnK + BLnC$$

Where; Ln T = the natural logarithm of TIME, Ln C = the natural logarithm of COST, Ln K = the natural logarithm of K, and B = the coefficient of Ln COST.

3.7.1.2. Other Time-Cost relationship Analysis

Besides the validation of the BTC model, the study tried to examine best-fit regressions in an attempt to find a model that possibly explains a larger portion of the variance in construction duration in terms of actual cost. Thus, this study will examine ten regression equations: Linear equation (LIN), Logarithmic equation (LOG), Inverse equation (INV), Quadric equation (QUA), Cubic equation (CUB), Compound equation (COM), Power equation (POW), S - curve equation (SCU), Growth equation (GRO), and Exponential equation (EXP). Table (3-2) shows the general equation for the above regression models.

Table 3-2: Equation forms of regression models

Regression model	Regression equation
Linear regression (LIN)	$Y = b_0 + b_1 * X$
Logarithmic regression (LOG)	$Y = b_0 + b_1 * \ln X$
Inverse regression (INV)	$Y = b_0 + b_1 / X$
Quadratic regression (QUA)	$Y = b_0 + b_1 * X + b_2 * X^2$
Cubic regression (CUB)	$Y = b_0 + b_1 * X + b_2 * X^2 + b_3 * X^3$
Compound regression (COM)	$Y = b_0 * b_1^X$
Power regression (POW)	$Y = b_0 * X^{b}{}_1$
S-Curve regression (SCU)	$Y = e^{(b0+b1/X)}$
Growth regression (GRO)	$Y = e^{(b0 + b1X)}$
Exponential regression (EXP)	$Y = b_0 * e^{b1X}$

Note: Y denotes the dependent variable; X denotes the independent variable; b_0, b_1, b_2 denote constant

3.7.1.3. Time-Road Length Relationship Analysis

Similarly, the study tried to examine best-fit regressions for time-road length relations besides the previous Bromilow's form. Thus, the ten regression equations; Linear equation (LIN), Logarithmic equation (LOG), Inverse equation (INV), Quadric equation (QUA), Cubic equation (CUB), Compound equation (COM), Power equation (POW), S - curve equation (SCU), Growth equation (GRO) and Exponential equation (EXP); shown in Table (3-2) above have been examined for Time-Road Length relationship.

3.7.2. Multiple Regression Analysis

When there are two or more independent variables, the analysis concerning the relationship is known as multiple correlation, and the equation describing such a relationship is the multiple regression equation (Kothari, 2004). The basic formulation for multiple regression analysis is (Helvacı, 2008):

$$Y = \alpha_0 + \alpha_1 X_1 \dots \dots \dots .. + \alpha_n X_n$$

Where Y = dependent variable; α0 = constant; α1 = partial regression coefficient for X1; X1 = independent variable 1; αn = partial regression coefficient for Xn; Xn = independent variable n; n= number of variable

3.7.3. Models Comparison and Selection

The developed models were compared to each other for the selection of the best-fit duration prediction model for ERA road projects. The study used Pearson correlation coefficient (R) and Model Goodness of fit comprised of significance level (p-value) and coefficient of determination (R^2) for comparisons and selection of the best-fit duration model.

1. Pearson correlation coefficient (R)

Pearson correlation is a measure of association between a dependent variable and an independent variable (Ogunsemi, 2009). The values of the correlation coefficient (R) range from -1 to +1 with negative numbers representing a negative correlation (as one variable increases, the other variable decreases) and positive numbers representing a positive correlation (as one variable increases, the other also increases). The closer the value is to -1 or +1, the stronger the association is between the variables provided that the p-value is less than 0.05. The generally accepted ranges often taught as rules of thumb (Field, 2024) were used for the interpretation of the correlations, shown in table (3-3)

Table 3-3: Rule of thumb for interpretation of correlations

Correlation range	Interpretations of relationships
0.8 to 1.0	Very Strong
0.6 to 0.79	Strong
0.4 to 0.59	Moderate
0.2 to 0.39	Weak/Low
0.0 to 0.19	Very Weak/Negligible

2. **Model Goodness of Fit**

In regression models, the elimination of insignificant variables will be established by using the significance level (p-value) and coefficient of determination (R^2). The p-value indicates the significance of the variables being included in the model. A significance level of 5% ($\rho < 0.05$) is often regarded as appropriate for most statistical tests. Thus, a 5% significance level will be used in this research exercise. The coefficient of determination, R^2, is used to indicate the goodness of fit of the models derived from the empirical data. It ranges from 0 to 1. R^2 of 1 means that all the data fit to regression line; R^2 is 0 if no linear relationship exists between observed and predicted data. Scatter and residual plots are used to detect the validity of the regression model. The scatter plot is a plot that is used to judge how well the regression line fits the data set. The residual plot is a plot of the residuals against the predicted values.

3.7.4. Testing and Validation of Model

Once the model has been developed, the next step is to test the validity of the estimate from the model and compare it with the actual results. The validity of the final model is assessed in terms of predictive accuracy. That is, the predicted duration values are compared with the detailed duration values to verify the predictive efficiency of the duration model. This study employed validations based on regression model assumptions and relative measures of accuracy dealing with errors, i.e. Mean Absolute Percentage Error (MAPE) to prove the validities of the developed duration model. The formulas used for the computation of MAPE are given by;

$$MAPE = \frac{1}{n}\sum_{i=1}^{n} \frac{|(Predicted\ duration)i - (\ Actual\ duration)i|}{|(Actual\ duration)i|} * 100$$

3.7.5. Ethical consideration

Ethics is the practice of doing research in a morally and legally sound manner. They are behavioral norms that draw lines between what is good and unattractive behavior as well as morally right and wrong. The researcher, while gathering the data, will clarify to the participants that this study is in their best interests, that participants will stay anonymous, if necessary, that the researcher will not put participants and his organization in danger, and that the researcher respects vulnerable populations. The data gathered will be used only for this research purpose and confidentially keep the information private. Finally, the researcher ensures that the study will be conducted in an honest, responsible, and ethically accountable way and leads to beneficial outcomes.

The overall research flow is indicated in Figure (3-1) below.

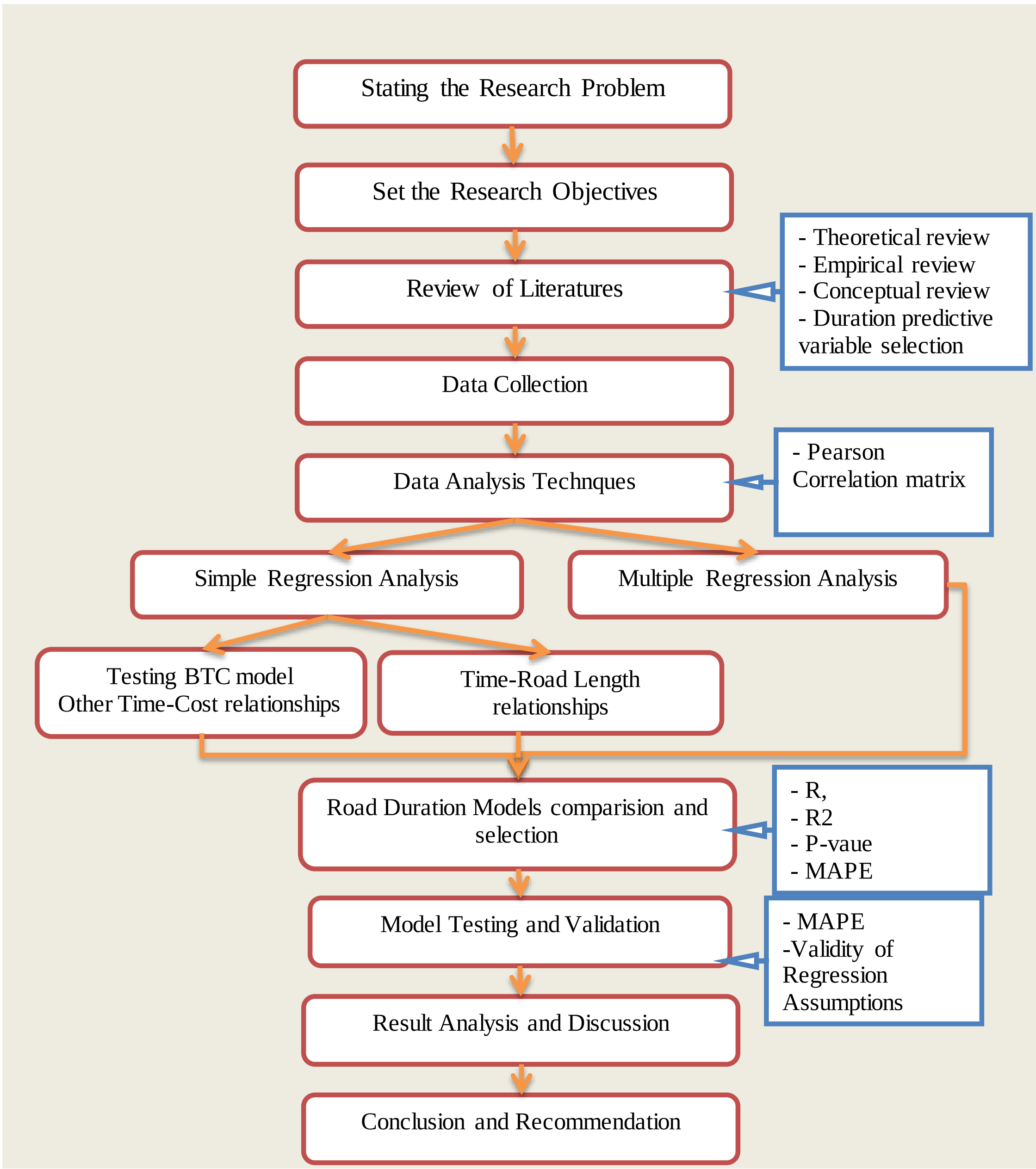

Figure 3-1: Research Methodology flow chart (Source: Own Data, 2025)

Chapter 4: Results and Discussion

4.1. Description of Sample Project Characteristics

The study utilized archival project data from thirty (30) road construction projects administered by the Ethiopian Roads Authority (ERA) between 2001 and 2016. Data were systematically collected from comprehensive project documentation, including contract documents, road designs and drawings, supplementary agreements, progress reports, and final payment certificates.

Descriptive statistics for key continuous variables revealed the scope and characteristics of the sampled projects. The projects had mean contract duration of 1085.13 calendar days and a mean actual duration of 1713.47 calendar days. Financially, the mean contract value stood at approximately 1,045,438,202.23 ETB, with an adjusted final cost averaging 1,093,318,169.48 ETB. Regarding design parameters, the road lengths (RL) in the sample ranged from 5.38 Km to 148.20 Km (Mean = 64.26 Km), road widths (RW) varied from 6.00 meters to 23.00 meters (Mean = 14.80 m), number of culverts (CN) ranged from zero to 282 (Mean = 115.7) and the number of lanes (LN) ranged from 2 to 6 (Mean = 2.13). A more detailed summary of the descriptive statistics for all continuous variables is presented in Table (4-1).

Table 4-1: Descriptive Statistics of continuous variables

	N	Minimum	Maximum	Mean	Std. Deviation
CD (Cal. days)	30	548.00	1464.00	1085.13	211.49
CC (ETB)	30	2,153,822.50	1,967,496,759.60	1,045,438,202.23	625,992,416.07
AD (Cal. days)	30	1004.00	2895.00	1713.47	503.55
AC adjusted (ETB)	30	2,754,635.25	3,011,114,991.00	1,093,318,169.48	696,342,308.86
RL (Km)	30	5.38	148.20	64.26	34.22
RW (m)	30	6.00	23.00	14.80	5.22
LN (No.)	30	2.00	6.00	2.13	.73
CN (No.)	30	.00	282.00	115.07	88.13
Valid N (list wise)	30				

The case road projects had some similar features for instance the tender method used is Open bid tender, and the owner/financial source of the projects is the Federal Democratic Republic of Ethiopia (FDRE). The descriptive statistics of categorical project variables of the study were computed based on frequencies and percentages. Table (4-2) below summarizes the categorical project variables of the project.

- Highway System (HS): The majority of projects (56.7%) were Asphalt Concrete roads, with the remaining 43.3% being Double Surface Treatment roads.
- Project Scope (PS): New Construction projects constituted 56.7% of the sample, while Rehabilitation projects accounted for 43.3%.
- Site Accessibility (SA): A larger proportion of the projects (60.0%) had good site accessibility, compared to 40.0% with poor access.
- Geometric Design Standard (GDS): The sample showed a distribution across different design standards, with DS4 being the most common at 40.0%, followed by DS5 at 36.7%, and DS6 at 20.0%. Only a small fraction (3.3%) was classified as DS3.
- Road Classification (RC): Link roads were the predominant type, representing 73.3% of the projects. Main Access Roads and Trunk Roads each comprised 13.3% of the sample.
- Contractor Category (CON): International contractors executed the majority of the projects (63.3%), with local contractors undertaking 36.7%.
- Delivery System (DS): The Design-Bid-Build (DBB) system was used in 60.0% of the projects, while the Design-Build (DB) system was employed in 40.0%.
- Contract Method (CM): The Item-Rate contract method was overwhelmingly prevalent, used in 86.7% of the projects, with Lump Sum contracts accounting for the remaining 13.3%.
- Availability of Bays (BY): Slightly more than half of the projects (53.3%) did not have at least one bay in the road section, whereas 46.7% did.
- Availability of Intersection (IN): A large majority of the projects (76.7%) did not include at least one intersection, while 23.3% did.
- Availability of Bridge (BG): Most projects (90.0%) did not feature at least one bridge structure along the road section, with only 10.0% including one or more bridges.

Table 4-2: The frequency and percentage of categorical road project parameters

I. No.	Categorical Variables		Frequency	Percent	Valid Percent
1	HS	Double Surface Treatment	13	43.3	43.3
		Asphalt Concrete	17	56.7	56.7
2	PS	Rehabilitation	13	43.3	43.3
		New Construction	17	56.7	56.7
3	SA	Poor Access	12	40.0	40.0
		Good Access	18	60.0	60.0
4	GDS	DS3	1	3.3	3.3
		DS4	12	40.0	40.0
		DS5	11	36.7	36.7
		DS6	6	20.0	20.0
5	RC	Main Access	4	13.3	13.3
		Link Road	22	73.3	73.3
		Trunk Road	4	13.3	13.3
6	CON	International	19	63.3	63.3
		Local	11	36.7	36.7
7	DS	DB	12	40.0	40.0
		DBB	18	60.0	60.0
8	CM	Lump sum	4	13.3	13.3
		Item-rate	26	86.7	86.7
9	BY	Not Available	16	53.3	53.3
		Available at least one	14	46.7	46.7
10	IN	Not Available	23	76.7	76.7
		Available at least one	7	23.3	23.3
11	BG	Not Available	27	90.0	90.0
		Available at least one	3	10.0	10.0

4.2. Bromilow's Time-Cost model (BTC) to predict the construction durations for Ethiopian road project

This section focused on validating and applying Bromilow's Time-Cost (BTC) model to Ethiopian road projects, specifically examining the relationship between construction duration (TIME/AD) and actual cost (COST/AC). Additionally, the researcher investigated various curve-fitting analyses for these time-cost relationships.

The BTC form of time - cost model is represented with the form:

$$TIME = K\,COST^{B} \qquad or \qquad LnTIME = LnK + B^{*}\,LnCOST$$

Where, TIME (AD) is the actual construction duration of construction period from the possession of site to the practical completion, measured in calendar days;

COST (AC) is the actual cost in ETB, adjusted to average price adjustment;

K is a constant describing the general level of time performance for ETB project;

B is a constant describing how the time performance is affected by project size as measured by cost.

The model summary, ANOVA, and Coefficients of the analysis is shown in the table (4-3) below. Accordingly, the linear regression model predicting LnAD from LnAC was statistically significant ($p=0.034$). This model explained 15.1% of the variance in LnAD ($R^2=0.151$). Both the intercept ($p<0.001$) and the slope coefficient for LnAC ($p=0.034$) were statistically significant. The full regression results and coefficients are presented in table (4-3), and a scatter plot of the relationship is displayed in Figure (4-1). Visual inspection of the scatter plot reveals a positive linear relationship between LnAC and LnAD. Nevertheless, the data points are visibly dispersed around the fitted regression line, which indicates a moderate strength for this observed relationship.

Table 4-3: Model summary, ANOVA, and Coefficients of the analysis for the BTC model

Model Summary

R	R Square	Adjusted R Square	Std. Error of the Estimate
.389	.151	.121	.272

The independent variable is LnAC.

ANOVA

	Sum of Squares	Df	Mean Square	F	Sig.
Regression	.369	1	.369	4.988	.034
Residual	2.073	28	.074		
Total	2.442	29			

The independent variable is LnAC.

Coefficients

	Unstandardized Coefficients		Standardized Coefficients		
	B	Std. Error	Beta	T	Sig.
LnAC	.075	.034	.389	2.233	.034
(Constant)	5.874	.688		8.542	.000

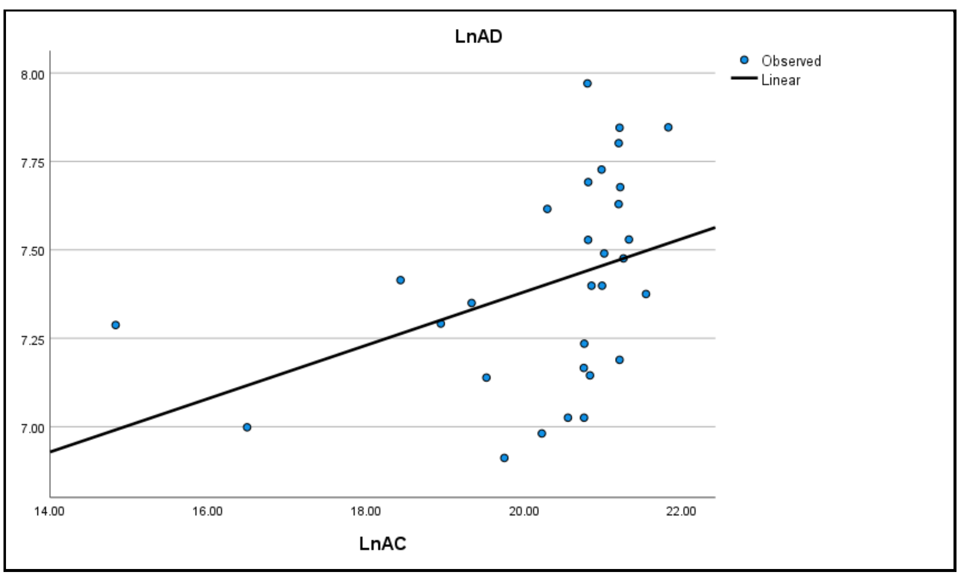

Figure 4-1: Scatter plot diagram for BTC model

Thus, applying the BTC model of time – cost relationship for the case study projects resulted in the following relationship:

$$LnAD = 5.87 + 0.08 * LnAC$$

Therefore, the equation rearranged in terms of the original variables is:

$$AD = 355.67 * (AC)^{0.075}$$

Where; AD-Actual Duration (Cal. days); AC-Actual Cost Adjusted to January 2016 (ETB)

This result suggests the wider applicability of Bromilow's time-cost concepts, showing that more expensive projects typically take longer to complete, even if the rate of increase is nonlinear and decreasing. Although the model is statistically significant, the moderate R^2 value of 0.151 indicates that the actual cost only explains a small amount of the variation in project duration. This implies that there were other significant variables at play that were not specifically included in this single-variable BTC model.

The researcher, besides validation of Bromilow's Time-Cost principle, tried to investigate different forms of curve fitting analysis for time-cost relationships for road projects in Ethiopia. Accordingly, the study examined ten regression equations: Linear equation (LIN), Logarithmic equation (LOG), Inverse equation (INV), Quadric equation (QUA), Cubic equation (CUB), Compound equation (COM), Power equation (POW), S - curve equation (SCU), Growth equation (GRO) and Exponential equation (EXP). Both normal time-cost relationships (i.e. untransformed) and Ln transformed (LnTime-LnCost) relationships were investigated.

Table 4-4: Model Summary and Parameter Estimates of Time-Cost Relationship (Untransformed)

	Model Summary					Parameter Estimates			
Equation	R Square	F	df1	df2	Sig.	Constant	b1	b2	b3
Linear	.277	10.740	1	28	.003	1297.185	3.808E-7		
Logarithmic	.150	4.954	1	28	.034	-936.648	130.390		
Inverse	.020	.583	1	28	.452	1730.587	-1077720226.842		
Quadratic	.277	5.182	2	27	.012	1288.569	4.012E-7	-8.254E-18	
Cubic	.278	3.333	3	26	.035	1304.270	3.130E-7	7.464E-17	-1.919E-26
Compound	**.285**	11.160	1	28	**.002**	1289.731	1.000		
Power	.151	4.988	1	28	.034	355.541	.075		
S	.017	.487	1	28	.491	7.414	-568798.568		
Growth	**.285**	11.160	1	28	**.002**	7.162	2.225E-10		
Exponential	**.285**	11.160	1	28	**.002**	1289.731	2.225E-10		

The independent variable is AC adjusted.

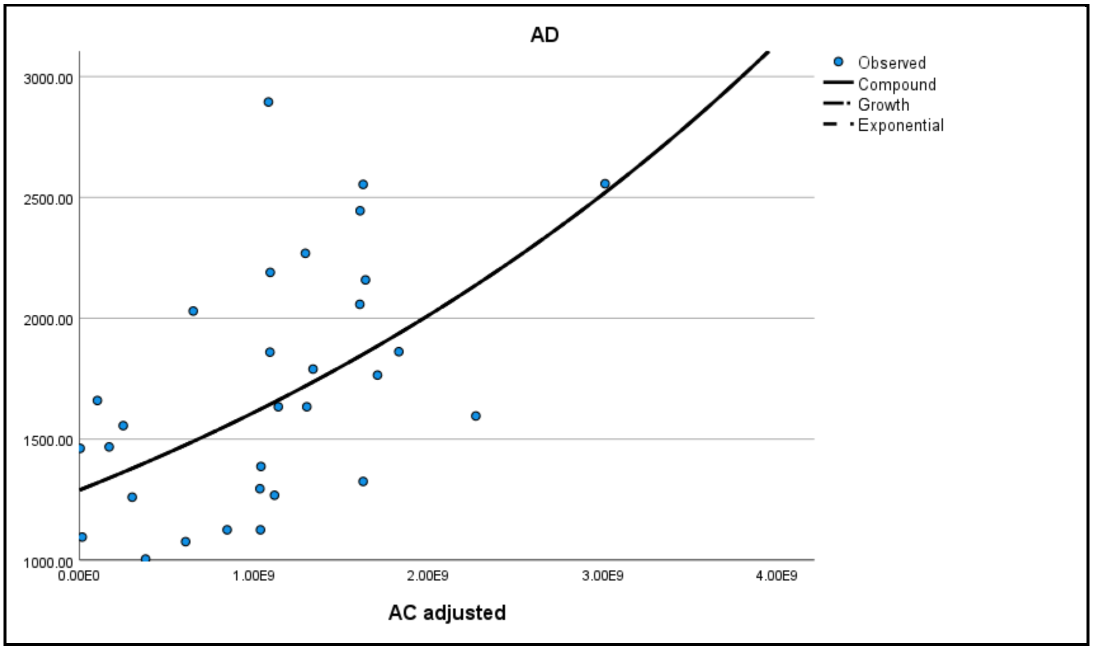

Figure 4-2: Scatter plot for AC vs. AD

All ten-regression models were examined to find the best-fit model of actual time and actual cost relationship as illustrated in Table (4-4). Several models demonstrated a statistically significant fit to the data at a significance level of p=0.05. The Compound (COM), Growth (GRO), and Exponential (EXP) models exhibited the highest R-squared value, explaining 28.5% of the variance in AD (R^2=0.285), and were highly statistically significant (p=0.002). Figure (4-2) showed the scatter plot of the Compound, Growth, and Exponential models of time-cost relation for ERA road projects. The plot visually demonstrates a positive and non-linear relationship between Actual Cost and Actual Duration. As project cost increases, the duration tend to increase, but at an accelerating rate, which these curved models attempt to capture. The model equations are given as follows:

$$COM\ Equation: AD = 1289.731 * 1.00^{AC}$$

$$GRO\ Equation: AD = e^{(7.162+2.225e-10*AC)}$$

$$EXP\ Equation: AD = 1289.731 * e^{2.225e-10*AC}$$

Where: AD-Actual road project duration (Cal. days), AC-Adjusted Actual Cost (ETB)

Furthermore, the study also extended the investigation of best-fit curve estimation of the Time-cost relationship for Ln transformed data, i.e. LnAD-LnAC model. The result shown in Table (4-5) revealed that the Cubic equation (CUB) produced a relatively moderate result than other regression forms with an R^2 value of 0.272 at a significant level of 0.014. This means that 27.2% of the variance in construction duration could be explained in terms of actual cost with the Cubic equation (CUB). The scatter plot in Figure (4-3), depicting the cubic model for LnAD versus LnAC, revealed that the relationship between duration and cost remained non-linear even after logarithmic transformation. This suggests a more complex, potentially U-shaped pattern, where project duration initially decreases or stabilizes, then significantly increases with rising costs. The model equation is given by:

$$LnAD = 10.202 - 0.030 * LnAC^{\,2} + 0.001 * LnAC^3$$

Where: LnAD-Ln transformed Actual road project duration (Cal. days), LnAC-Ln transformed Adjusted Actual Cost (ETB)

Table 4-5: Model Summary and Parameter Estimates of LnTime-LnCost relationship (Ln transformed)

Dependent Variable: LnAD

	Model Summary					Parameter Estimates			
Equation	R Square	F	df1	df2	Sig.	Constant	b1	b2	b3
Linear	.151	4.988	1	28	.034	5.874	.075		
Logarithmic	.137	4.434	1	28	.044	3.433	1.320		
Inverse	.122	3.901	1	28	.058	8.530	-22.715		
Quadratic	.268	4.943	2	27	.015	17.064	-1.144	.033	
Cubic	**.272**	5.054	2	27	**.014**	10.202	.000	-.030	.001
Compound	.151	4.965	1	28	.034	6.021	1.010		
Power	.136	4.410	1	28	.045	4.335	.178		
S	.122	3.877	1	28	.059	2.153	-3.057		
Growth	.151	4.965	1	28	.034	1.795	.010		
Exponential	.151	4.965	1	28	.034	6.021	.010		

The independent variable is LnAC.

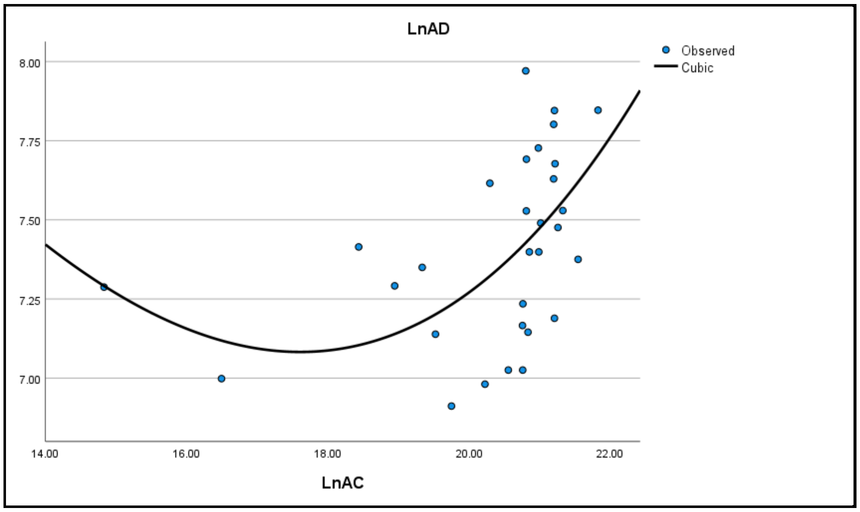

Figure 4-3: Scatter plot for LnAD vs. LnAC

Despite the statistical significance of the Cubic model in the logarithm-transformed analysis of the time-cost relationship (LnAD vs. LnAC, R^2 = 0.272, p = 0.014), a comparison with the untransformed variables (AD vs. AC) indicated that the Compound, Growth, and Exponential models provided a marginally stronger explanation of the dependent variable. These untransformed models achieved a slightly higher R-squared value of 0.285 (p=0.002), suggesting a comparatively better fit for describing the time-cost relationship in road projects in Ethiopia.

4.3. Empirical relationships between the construction duration-road length for Ethiopian road construction projects

In an analogous way to the above time-cost relationship, the ten regression forms were tested for searching best-fit time-road length relationships for the case study road projects. Similarly, to the above time-cost relationship both normal (without transformation) and Ln-transformed Time-Length relationships were investigated.

The Model Summary section in Table (4-6) presented the fit of each model, indicated by the R-squared value and the significance of the overall model F-test. The model with the highest R-squared value was the Cubic model, explaining 18.2% of the variance in AD (R^2=0.182). However, the overall Cubic model was not statistically significant (p=0.151). Several other models demonstrated a statistically significant fit to the data at a significance level of p=0.05. However, the Compound, Growth, and Exponential models, while explaining slightly less variance than the Cubic model, all demonstrated a statistically significant relationship, suggesting they might offer more parsimonious and robust explanations of the relationship between AD and RL compared to the Cubic model in this specific dataset.

Table 4-6: Model Summary and Parameter Estimates for Time-Length relationship (Untransformed)

Dependent Variable: AD

Equation	Model Summary					Parameter Estimates			
	R Square	F	df1	df2	Sig.	Constant	b1	b2	b3
Linear	.158	5.262	1	28	.030	1337.394	5.852		
Logarithmic	.124	3.962	1	28	.056	860.610	216.876		
Inverse	.044	1.274	1	28	.269	1795.550	-2729.906		
Quadratic	.166	2.694	2	27	.086	1245.387	9.784	-.031	
Cubic	**.182**	1.922	3	26	**.151**	1424.600	-4.525	.207	-.001
Compound	**.162**	5.419	1	28	**.027**	1320.806	1.003		
Power	.119	3.772	1	28	.062	1016.848	.122		
S	.035	1.009	1	28	.324	7.448	-1.406		
Growth	**.162**	5.419	1	28	**.027**	7.186	.003		
Exponential	**.162**	5.419	1	28	**.027**	1320.806	.003		

The independent variable is RL.

The scatter plot for the AD-RL relationship (Figure 4-4) indicated a positive, non-linear trend where Actual Duration generally increased with Road Length. However, the wide scatter of data points around the Compound, Growth, and Exponential curves suggests that Road Length alone, even with these non-linear transformations, is a weak predictor for untransformed Actual Duration.

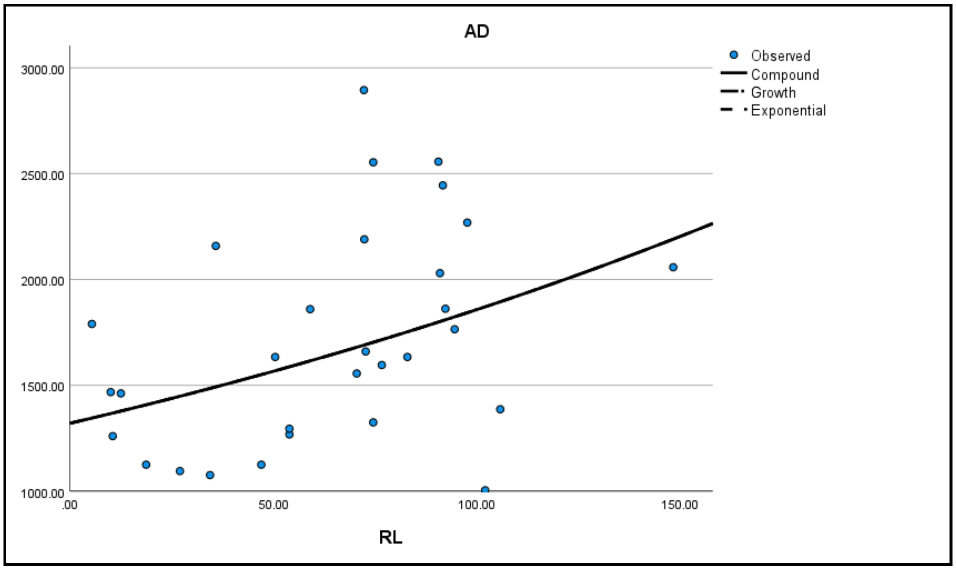

Figure 4-4: Scatter plot for AD vs. RL

In addition, the study investigated the relationship between time and road length based on Ln transformed data set, i.e. Ln Duration-Ln Road Length. Table (4-7) showed the model summary and parameter estimates for the Ln Duration-Ln Road Length relationship. In this transformed dataset, the Cubic model yielded the highest R-squared value, explaining 24.7% of the variance in LnAD (R^2=0.247). The overall Cubic model, however, was not statistically significant at the conventional 5% level (p=0.057). The Quadratic model emerged as the only statistically significant model at 5%, explaining 20.9% of the variance in LnAD (R^2=0.209, p=0.042).

Table 4-7: Model Summary and Parameter Estimates for Ln Time-Ln Road Length relationship (Ln transformed)

Dependent Variable: LnAD

	Model Summary					Parameter Estimates			
Equation	R Square	F	df1	df2	Sig.	Constant	b1	b2	b3
Linear	.119	3.772	1	28	.062	6.924	.122		
Logarithmic	.088	2.709	1	28	.111	6.952	.338		
Inverse	.057	1.682	1	28	.205	7.622	-.798		
Quadratic	**.209**	3.569	2	27	**.042**	8.304	-.755	.129	
Cubic	**.247**	2.849	3	26	**.057**	11.116	-3.540	.990	-.085
Compound	.117	3.721	1	28	.064	6.938	1.017		
Power	.087	2.657	1	28	.114	6.965	.045		
S	.055	1.635	1	28	.211	2.030	-.106		

Growth	.117	3.721	1	28	.064	1.937	.016		
Exponential	.117	3.721	1	28	.064	6.938	.016		

The independent variable is LnRL.

Figure (4-5) displayed the scatter plot of the Quadratic model for the LnAD-LnRL relationship. The plot indicated a non-linear (U-shaped) relationship where, as LnRL increased, LnAD initially decreased or leveled off before significantly increasing. This suggested that even after log-transformation, the impact of road length on duration followed a curvilinear, rather than a simple linear, pattern.

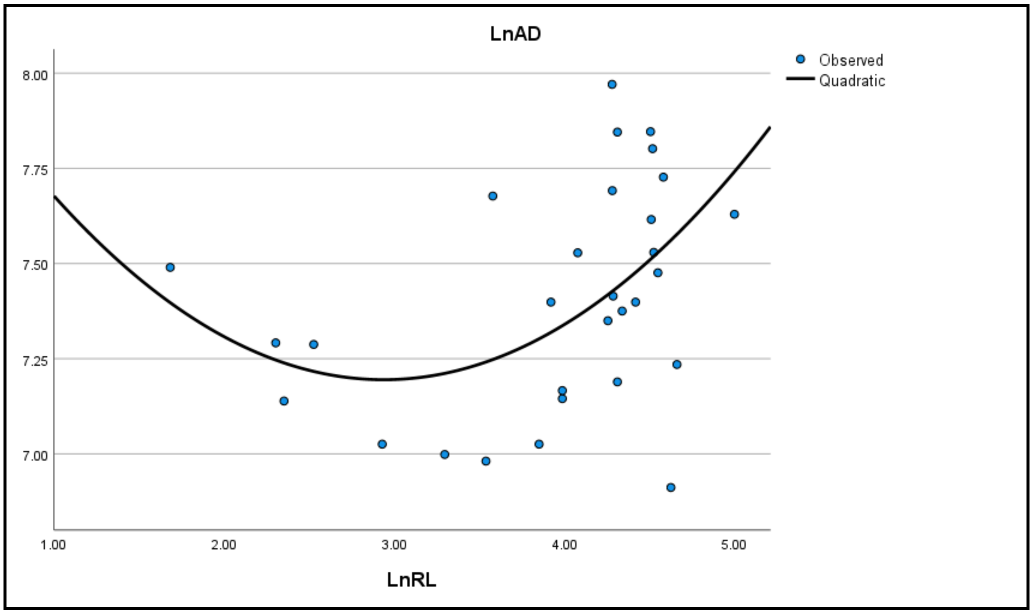

Figure 4-5: Scatter plot for LnAD vs. LnRL

Comparing the Ln-transformed variables with the previous untransformed variables, it appears that the natural logarithm transformation of the variables has altered the best-fitting model and its statistical significance. The change in the best-fitting and significant models after transformation suggests that the original relationship between road length and actual duration might not be purely linear or simple exponential. The fact that the Quadratic model became significant in the ln-ln space indicates a more complex, and curvilinear relationship. The equation for the Quadratic relationship in the transformed data is given by:

$$\boldsymbol{LnAD = 8.304 - 0.755 \times LnRL - 0.129 \times (LnRL)^2}$$

4.4. Conceptual construction duration prediction model for Ethiopian road construction projects based on the gaps of the BTC model

Multiple regression analysis has consistently proven its effectiveness as a powerful approach for modeling construction time, surpassing the constraints of single-variable forecasts for road project duration, such as those that are exclusively based on cost or length. By making it easier to evaluate the combined impact of several independent variables on a single result, this statistical method provides a more complete and nuanced picture of project duration. This integrated approach greatly improves the accuracy and dependability of predictions for the duration of construction. In the literature review section of this article, section (2.2.4), some of the prior studies were discussed. Similarly, to the previous models, both the normal data set and Ln transformed data were tested. The independent variables included for the multiple regression method were extracted from the correlation Table (4-8) & Table (4-9), which had shown significant correlation with the actual project duration.

The Pearson correlation matrix shown in Table (4-8) below revealed that six (6) variables out of the total sixteen (16) have shown a significant correlation (p-value < 0.05) with the actual construction duration. The correlation matrix shows that the Actual construction duration has a significant positive correlation with Actual Cost (r=0.527, p=0.003). The positive correlation between project duration with actual cost indicates that the duration of the project increases as the project cost increases. The Projects with higher budgets are generally more complex, requiring more time. However, the relationship is not linear; Bromilow's time-cost principles also prove this relationship. The Highway System categorical variable, consisting of Double Surface Treatment and Asphalt Concrete, has shown a strong positive correlation with the project duration for the ERA road project (r=0.736, p=0.000). This implies that road projects involving Asphalt Concrete tend to have significantly longer durations compared to those involving Double Surface Treatment, and this association is very strong and highly reliable. This could be due to the inherent complexity, material requirements, construction processes, or quality standards associated with Asphalt Concrete pavements compared to Double Surface Treatment. The Project Scope, which consists new construction and rehabilitation category, also has shown a positive correlation (r=0.487, p=0.006) with the duration of the project. This correlation suggested that projects categorized as "new construction" tend to have moderately longer durations compared to "rehabilitation" projects, and this association is statistically significant. This is intuitively plausible, as new construction often involves more work that is extensive, material sourcing, and site preparation than rehabilitation efforts. Road Length measured in unit meters has a moderated positive correlation with road duration for ERA project ($r=0.398$, $p=0.030$). The statistical significance ($p=.030$) indicated that road length is not just randomly associated with duration, but it is a genuine factor influencing how long a project will take. The positive correlation means that as the physical length of a road project increases, the time required to complete it also increases. In addition, the Contractor Category (a categorical variable distinguishing between international and local firms) demonstrated a moderate positive correlation with project duration ($r=0.451$, $p=0.012$). This statistically significant association implies that, on average, projects assigned to local contractors are associated with moderately longer completion times than those handled by international contractors. This disparity may be attributed to several factors, including the perceived higher capacity, established reputation, or greater access to larger financial guarantees often associated with international firms, which can influence project execution speed. Conversely, the actual project duration demonstrated a significant negative correlation with Site Accessibility ($r=-0.704$, $p < 0.001$). This strong negative relationship implies that projects with good site accessibility tend to be completed earlier than those with poorer accessibility. The remaining variables in the correlation matrix did not show a statistically significant linear relationships with the actual construction duration (AD) at the $p<.05$ level.

In addition to the analysis of untransformed variables, the Pearson correlation matrix was also established for the natural logarithm-transformed continuous variables. The correlation matrix, presented in Table (4-9), revealed that five of the sixteen variables examined exhibited a statistically significant linear correlation ($p<0.05$) with the natural logarithm of actual project duration (LnAD). Specifically, LnAD showed statistically significant positive correlations with the natural logarithm of Actual Cost (LnAC) ($r=0.389$, $p=0.034$), Highway System (HS) ($r=0.764$, $p<0.001$), Project Scope (PS) ($r=0.493$, $p=0.006$), and Contractor Category (CON) ($r=0.473$, $p=0.008$). Conversely, a significant negative correlation was found between LnAD and Site Accessibility (SA) ($r=-0.693$, $p<0.001$). The remaining variables in the correlation matrix did not demonstrate a statistically significant linear relationship with the natural logarithm of actual project duration.

Table 4-8: Pearson correlation matrix for Actual duration with duration predictive variables (Untransformed)

		AD	AC adjusted	HS	PS	SA	GDS	RC	RL	RW	LN	CON	DS	CM	CN	BY	IN	BG
AD	Pearson Correlation	1	.527**	.736**	.487**	-.704**	.041	-.091	.398*	.145	.029	.451*	-.278	.261	.080	-.076	.023	.000
	Sig. (2-tailed)		.003	.000	.006	.000	.828	.631	.030	.445	.880	.012	.137	.163	.673	.691	.903	.998
AC adjusted	Pearson Correlation		1	.289	.457*	-.531**	-.140	-.116	.447*	-.028	.066	.192	-.110	.103	.219	-.222	-.055	.271
	Sig. (2-tailed)			.122	.011	.003	.461	.542	.013	.884	.728	.311	.562	.589	.244	.239	.772	.148
HS	Pearson Correlation			1	.321	-.577**	.044	.000	.272	.222	.162	.247	-.165	.251	-.024	-.126	.005	-.157
	Sig. (2-tailed)				.083	.001	.817	1.000	.145	.239	.391	.189	.384	.182	.899	.508	.978	.407
PS	Pearson Correlation				1	-.439*	.044	-.261	.069	.169	.162	.107	-.165	.053	.118	.009	.005	.067
	Sig. (2-tailed)					.015	.817	.164	.718	.371	.391	.574	.384	.782	.535	.962	.978	.724
SA	Pearson Correlation					1	.067	.000	-.336	-.105	.152	-.226	.167	-.120	-.056	.082	-.032	.045
	Sig. (2-tailed)						.725	1.000	.070	.582	.424	.230	.379	.527	.769	.667	.866	.812
GDS	Pearson Correlation						1	.000	-.350	-.089	.289	-.091	.151	-.129	-.365*	-.104	.084	.109
	Sig. (2-tailed)							1.000	.058	.641	.121	.634	.427	.498	.047	.584	.659	.566
RC	Pearson Correlation							1	.080	.101	-.360	-.134	.264	.380*	.042	.129	.153	.000
	Sig. (2-tailed)								.673	.597	.051	.480	.159	.038	.827	.496	.421	1.000
RL	Pearson Correlation								1	-.229	-.325	.137	-.007	.090	.797**	-.178	.067	.084
	Sig. (2-tailed)									.224	.080	.471	.969	.636	.000	.347	.724	.660
RW	Pearson Correlation									1	.224	-.058	.028	.405*	-.121	.401*	.091	-.160
	Sig. (2-tailed)										.233	.761	.884	.026	.526	.028	.634	.397
LN	Pearson Correlation										1	-.141	-.227	.073	-.247	-.174	-.102	-.062
	Sig. (2-tailed)											.456	.227	.702	.189	.359	.590	.745
CON	Pearson Correlation											1	-.085	.095	.003	-.018	-.093	.208
	Sig. (2-tailed)												.656	.618	.986	.923	.626	.271
DS	Pearson Correlation												1	.280	.069	.355	.129	.045
	Sig. (2-tailed)													.134	.717	.055	.498	.812
CM	Pearson Correlation													1	-.042	.367*	.216	-.196
	Sig. (2-tailed)														.827	.046	.251	.299
CN	Pearson Correlation														1	.009	.183	.163
	Sig. (2-tailed)															.961	.332	.391
BY	Pearson Correlation															1	.432*	-.089
	Sig. (2-tailed)																.017	.640
IN	Pearson Correlation																1	.079
	Sig. (2-tailed)																	.679
BG	Pearson Correlation																	1
	Sig. (2-tailed)																	

Table 4-9: Pearson correlation matrix for Actual construction duration with duration predictive variables (Ln transformed)

		LnAD	LnAC	LnRL	LnRW	LnCN	LnLN	HS	PS	SA	GDS	RC	CON	DS	CM	BG	BY	IN
LnAD	Pearson Correlation	1	.389*	.345	.191	.249	.055	.764**	.493**	-.693**	.011	-.070	.473**	-.273	.304	.028	-.061	-.003
	Sig. (2-tailed)		.034	.062	.313	.192	.773	.000	.006	.000	.955	.714	.008	.144	.102	.884	.749	.988
LnAC	Pearson Correlation		1	.458*	.006	.622**	.087	.306	.313	-.426*	-.315	-.016	-.055	-.291	-.043	.192	-.210	.111
	Sig. (2-tailed)			.011	.973	.000	.648	.099	.092	.019	.090	.933	.774	.118	.821	.309	.265	.558
LnRL	Pearson Correlation			1	-.215	.854**	-.520**	.197	.168	-.396*	-.350	.147	.075	-.025	.000	.108	-.152	.123
	Sig. (2-tailed)				.254	.000	.003	.297	.374	.030	.058	.439	.696	.894	.999	.570	.424	.518
LnRW	Pearson Correlation				1	.060	.207	.206	.117	-.089	-.176	.117	-.047	.039	.430*	-.199	.408*	.093
	Sig. (2-tailed)					.757	.273	.275	.537	.641	.353	.540	.806	.837	.018	.293	.025	.624
LnCN	Pearson Correlation					1	.[c]	.175	.212	-.171	-.356	-.025	-.093	-.136	-.059	.088	-.064	.190
	Sig. (2-tailed)						.	.364	.269	.375	.058	.899	.630	.481	.759	.650	.742	.323
LnLN	Pearson Correlation						1	.162	.162	.152	.289	-.360	-.141	-.227	.073	-.062	-.174	-.102
	Sig. (2-tailed)					.		.391	.391	.424	.121	.051	.456	.227	.702	.745	.359	.590
HS	Pearson Correlation							1	.321	-.577**	.044	.000	.247	-.165	.251	-.157	-.126	.005
	Sig. (2-tailed)								.083	.001	.817	1.000	.189	.384	.182	.407	.508	.978
PS	Pearson Correlation								1	-.439*	.044	-.261	.107	-.165	.053	.067	.009	.005
	Sig. (2-tailed)									.015	.817	.164	.574	.384	.782	.724	.962	.978
SA	Pearson Correlation									1	.067	.000	-.226	.167	-.120	.045	.082	-.032
	Sig. (2-tailed)										.725	1.000	.230	.379	.527	.812	.667	.866
GDS	Pearson Correlation										1	.000	-.091	.151	-.129	.109	-.104	.084
	Sig. (2-tailed)											1.000	.634	.427	.498	.566	.584	.659
RC	Pearson Correlation											1	-.134	.264	.380*	.000	.129	.153
	Sig. (2-tailed)												.480	.159	.038	1.000	.496	.421
CON	Pearson Correlation												1	-.085	.095	.208	-.018	-.093
	Sig. (2-tailed)													.656	.618	.271	.923	.626
DS	Pearson Correlation													1	.280	.045	.355	.129
	Sig. (2-tailed)														.134	.812	.055	.498
CM	Pearson Correlation														1	-.196	.367*	.216
	Sig. (2-tailed)															.299	.046	.251
BG	Pearson Correlation															1	-.089	.079
	Sig. (2-tailed)																.640	.679
BY	Pearson Correlation																1	.432*
	Sig. (2-tailed)																	.017
IN	Pearson Correlation																	1
	Sig. (2-tailed)																	

Accordingly, the independent variables for a normal data set consisted of the adjusted Actual Cost (AC), Highway System (HS), Project Scope (PS), Site Accessibility (SA), Road Length (RL), and Contractor Category (CON). For the Ln-transformed duration model, however, a minor modification to the above-listed variables was made based on the correlation matrix in Table (4-9); namely, the road length variable was excluded due to its insignificant correlation with the transformed project duration. The enter method of multiple regression analysis was utilized. Tables (4-10) & (4-11) show the model summary, ANOVA, and coefficients of the regression models for the normal data set and Ln transformed data respectively.

(1) <u>Multiple regression analysis for normal data (Untransformed)</u>

A multiple linear regression analysis was conducted to assess the predictive power of six independent variables – AC (adjusted actual cost), CON (contractor category), PS (project scope), RL (road length), HS (highway System), and SA (site accessibility) – on the dependent variable, Actual duration (AD).

Table 4-10: Multiple regression analysis output for the normal data (Untransformed)

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.877[a]	.769	.709	271.79278	1.828

a. Predictors: (Constant), CON, PS, RL, HS, AC adjusted, SA

b. Dependent Variable: AD

ANOVA

Model		Sum of Squares	Df	Mean Square	F	Sig.
1	Regression	5654351.167	6	942391.861	12.757	.000[b]
	Residual	1699040.299	23	73871.317		
	Total	7353391.467	29			

a. Dependent Variable: AD

b. Predictors: (Constant), CON, PS, RL, HS, AC adjusted, SA

Coefficients

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	T	Sig.	Collinearity Statistics Tolerance	Collinearity Statistics VIF
1	(Constant)	1246.526	215.018		5.797	.000		
	AC	7.755E-8	.000	.107	.794	.436	.550	1.818
	HS	425.167	125.985	.426	3.375	.003	.632	1.583
	PS	163.272	120.504	.163	1.355	.189	.691	1.448
	SA	-241.717	143.127	-.239	-1.689	.105	.501	1.997
	RL	1.615	1.720	.110	.939	.357	.736	1.360
	CON	245.374	107.376	.239	2.285	.032	.920	1.087

a. Dependent Variable: AD

The Model Summary (Table 4-10) revealed that the independent variables collectively accounted for a substantial portion of the actual project duration (AD). The R-squared value of 0.769 indicates that approximately 76.9% of the variability in Actual Duration can be explained by the predictor variables included in the model. The ANOVA results (implicit in the overall model summary) showed that the regression model was highly statistically significant ($p< 0.001$). This strong overall significance confirms that the model as a whole is a reliable predictor of Actual road duration.

The Analysis of the coefficients Table (4-10) revealed that Highway System (HS) (p=0.003), and Contractor Category (CON) (p=0.032) demonstrated a statistically significant effect on Actual duration. Conversely, although conceptually important, AC adjusted (p=0.436), Project Scope (PS) (p=0.189), Site Accessibility (SA) (p=0.105), and Road Length (RL) (p=0.357) variables did not show statistically significant unique contributions to the model at the conventional 5% significance level.

Additionally, the Collinearity diagnostics (Tolerance and VIF) showed that all VIF values were well below 10 (ranging from 1.087 to 1.997), and tolerance values were above 0.1 (ranging from 0.501 to 0.920). These values indicate that multicollinearity is not a significant concern among the independent variables in this regression model. The relationship is shown by the equation:

$$\boldsymbol{AD = 1246.526 + 7.755e^{-8}(AC) + 425.167(HS) + 163.272(PS) - 241.717(SA) + 1.615(RL) + 245.374(CON)}$$

Where; AD-Actual road duration (Cal. days); AC (Adjusted actual project cost in ETB; HS (Highway System) 1 for Asphalt concrete & 0 for Double Surface treatment; PS (Project Scope) 1 for new construction & 0 for rehabilitation project; SA (Site Accessibility) 1 for Good site access & 0 for poor site access; RL (Road length in Km); and CON (Contractor category) 1 for Local Contactor & 0 for International Contractor;

(2) Multiple regression analysis for Ln transformed data

A multiple linear regression was performed to investigate the relationship between the natural logarithm of Actual Duration (LnAD), as the dependent variable, and five independent variables: CON (Contractor Category), LnAC (Adjusted Actual Cost), PS (Project Scope), HS (Highway System), and SA (Site Accessibility).

Table 4-11: Multiple regression analysis output for Ln Transformed data

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.886[a]	.786	.741	.14763	1.687

a. Predictors: (Constant), CON, LnAC, PS, HS, SA

b. Dependent Variable: LnAD

ANOVA

Model		Sum of Squares	Df	Mean Square	F	Sig.
1	Regression	1.919	5	.384	17.611	.000[b]
	Residual	.523	24	.022		
	Total	2.442	29			

Coefficients

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	6.794	.442		15.374	.000		
	LnAC	.021	.021	.107	.989	.333	.767	1.304
	HS	.269	.068	.467	3.962	.001	.642	1.559
	PS	.101	.062	.176	1.643	.113	.782	1.279
	SA	-.137	.075	-.235	-1.827	.080	.541	1.847
	CON	.173	.059	.292	2.928	.007	.896	1.116

a. Dependent Variable: LnAD

The Model Summary (Table 4-11) indicates that the chosen set of independent variables explains a substantial proportion of the variance in LnAD. The R-squared value of 0.786 suggests that approximately 78.6% of the variability in LnAD can be accounted for by the combined influence of CON, LnAC, PS, HS, and SA. The ANOVA indicated that the regression model was statistically significant ($p<0.001$). This overall significance confirms that at least one of the independent variables contributes significantly to explaining the variance in LnAD.

Further examination of the Coefficients table (Table 4-11) revealed that Highway System (HS) ($p<0.001$), and Contractor category (CON) ($p=0.007$) demonstrated a statistically significant positive effect on LnAD. Conversely, Adjusted cost (LnAC) ($p=0.333$), Project Scope (PS) ($p=0.113$), and Site Accessibility (SA) ($p=0.080$) did not show statistically significant unique contributions to the model at the conventional 5% significance level.

Collinearity diagnostics revealed that all Variance Inflation Factor (VIF) values were well below 10 (ranging from 1.116 to 1.847), and Tolerance values were above 0.1 (ranging from 0.541 to 0.896), indicating that multicollinearity among the independent variables is not a significant issue in this regression model, ensuring the reliability of the individual coefficient estimates. The relationship is expressed in the form of;

$$\mathbf{LnAD} = \mathbf{6.794} + \mathbf{0.021}(\mathbf{LnAC}) + \mathbf{0.269}(\mathbf{HS}) + \mathbf{0.101}(\mathbf{PS}) - \mathbf{0.137}(\mathbf{SA}) + \mathbf{0.173}(\mathbf{CON})$$

Where; LnAD-Ln transformed Actual road duration (Cal. days), LnAC- Ln transformed adjusted actual cost in ETB, HS (Highway System) 1 for Asphalt concrete & 0 for Double Surface treatment, SA (Site Accessibility) 1 for Good site access & 0 for poor site access, PS (Project Scope) 1 for new construction & 0 for the rehabilitation project, CON (Contractor category) 1 for Local Contactor & 0 for International Contractor

4.5. Comparisons and final road duration model selection

To identify the most precise model for predicting road project duration, various simple and multiple regression models were developed. These models were then rigorously compared based on their explanatory power (R^2) and predictive accuracy (Mean Absolute Percentage Error - MAPE), with results detailed in Table (4-12).

The simple regression models consistently demonstrated limited explanatory power for project duration, as indicated by their low R^2 values, ranging from 0.151 for Bromilow's Time-Cost Model to 0.285 for the untransformed Compound/Growth/Exponential Time-Cost model. These R^2 values fall at or even below the lower end of the previously observed global ranges for similar simple construction duration models (typically cited between 0.205 and 0.850). Furthermore, these simple models, especially the untransformed ones, exhibited unacceptably high MAPE values (20.56% for untransformed Time-Cost and 21.59% for untransformed Time-Length). Even with logarithmic transformation, the simple models' MAPEs remained comparatively higher (e.g., 2.97% for Bromilow's and 2.79% for LnTime-LnCost Cubic), indicating that their predictions, while perhaps explaining some variance, still carry significant average prediction errors. This collective performance underscores that a single cost or road length variable alone is insufficient to accurately capture the multifaceted nature of road project duration in the Ethiopian context, rendering these simple models unsuitable for robust estimation and policy setting.

Table 4-12: Road duration prediction model comparisons

Models	**Equation**	**R^2**	**MAPE**
1. Simple Regression models			
Bromilow's Time-Cost Model		0.151	2.97
Time-Cost (Untransformed)	Compound/Growth/Exponential	0.285	20.56
LnTime-LnCost (Ln transformed)	Cubic	0.272	2.79
Time-Length (Untransformed)	Compound/Growth/Exponential	0.162	21.59
Ln Time-Ln Length (Ln transformed)	Quadratic	0.209	2.72
2. Multiple Regression Models			
MLR (untransformed)		0.769	10.23
MLR (Ln transformed)		***0.786***	***1.43***

In contrast, the multiple regression analyses yielded significantly superior performance. The untransformed Multiple Linear Regression (MLR) model showed a substantial improvement in explanatory power (R^2=0.769), but its MAPE of 10.23% still suggested considerable average prediction error. Critically, the Ln-transformed Multiple Linear Regression (MLR) model emerged as the optimal choice, demonstrating both high explanatory power and remarkable precision. With an R^2 of 0.786, this model explained a substantial 78.6% of the variance in project duration, placing it firmly within the higher end of globally observed R^2 ranges for comprehensive construction duration prediction models. More remarkably, its Mean Absolute Percentage Error (MAPE) was a mere 1.43%; the low MAPE signifies a very high level of predictive accuracy.

The substantial improvement in both R^2 and MAPE observed in the Ln-transformed MLR model is attributable to its inclusion of multiple, relevant explanatory variables (Actual Cost, Contractor Category, Project Scope, Highway System, and Site Accessibility), and the use of a logarithmic transformation, which often better captures the non-linear relationships prevalent in construction data. This finding strongly supports the argument that accurate duration estimation in complex domains like road construction necessitates a multivariate approach that accounts for diverse project characteristics. The developed Ln-transformed MLR model, with its high R^2 and low MAPE, provides a robust, data-driven tool for front-end duration predictions in the Ethiopian road construction industry, offering a significant improvement over current subjective methods and poorly performing simple models. The final selected model is represented by the equation;

$$\mathbf{LnAD} = \mathbf{6.794} + \mathbf{0.021}(\mathbf{LnAC}) + \mathbf{0.269}(\mathbf{HS}) + \mathbf{0.101}(\mathbf{PS}) - \mathbf{0.137}(\mathbf{SA}) + \mathbf{0.173}(\mathbf{CON})$$

The relationship revealed that the road construction duration in Ethiopia is largely affected by the type of highway system whether it's Asphalt Concrete or Double Surface Treatment. It was visible that Asphalt Concrete road construction required a longer duration than other road systems as the scope of the project increased, along with factors like material resource availability and transportation to the construction site.

In addition, it has been noted that project site accessibility has a significant impact on the duration of road projects. Road projects with poor site accessibility tend to take a longer duration than those with easily accessible construction sites.

Furthermore, contractor category also had a significant impact on the project duration. The analysis indicated that local contractors required a longer time for construction than international contractors did. This might be attributed to the superior resource availability, financial capacity, and technical expertise of international contractors, enabling them to implement projects in a shorter period compared to local contractors.

4.6. Validity of Regression assumptions

The validity of regression assumptions refers to a set of conditions that must be met for the results of a linear regression model to be considered reliable, unbiased, and statistically valid. Therefore, a rigorous validation of these assumptions is paramount to ensure the statistical validity and predictive trustworthiness of the developed models. For the multiple linear regression (MLR) models developed in this study, particular attention was paid to validating the following key assumptions: linearity, homoscedasticity, normality of residuals, and absence of multicollinearity.

4.6.1. Linearity

The assumption of linearity posits a direct, straight-line relationship between the dependent variable (outcome) and the independent variable(s) (predictors). Visual inspection of the scatterplot depicting standardized residuals against standardized predicted values for Ln Duration (Figure 4-6) revealed a random distribution of points around the zero line, devoid of any clear curvilinear trend. This outcome indicates that the linearity assumption is substantially satisfied for the chosen regression model.

4.6.2. Homoscedasticity

Homoscedasticity, a key assumption of linear regression, dictates that the variance of the residuals (or errors) remains constant across all levels of the independent variable(s). As depicted in Figure 4-6, the scatterplot of standardized residuals versus standardized predicted values confirms a consistent and random spread of points around the zero line, devoid of any widening or narrowing patterns. This observation largely indicates that the homoscedasticity assumption is met, thereby supporting the reliability of the standard errors and coefficient estimates.

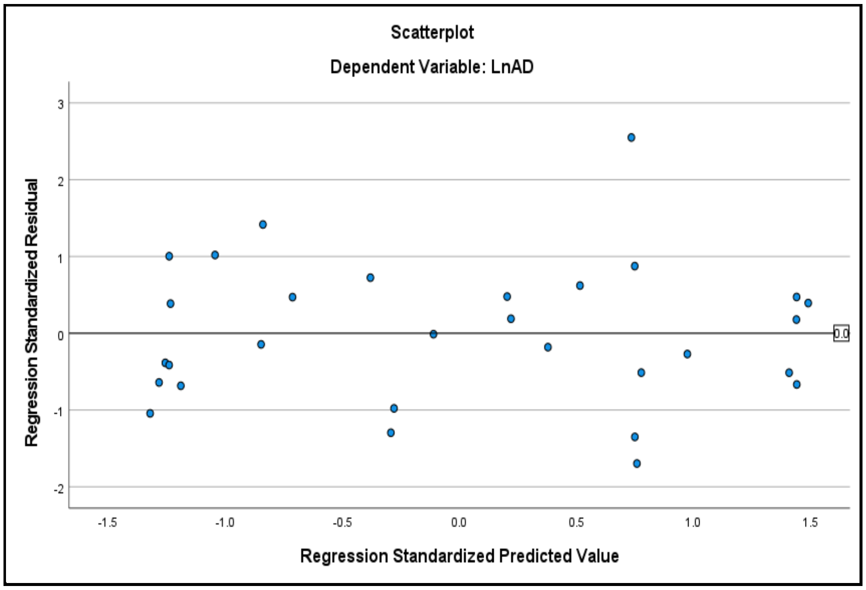

Figure 4-6: Scatter plot for standardized regression residual

4.6.3. Normality of Residuals

The assumption of Normality of Residuals in linear regression suggests that the model's errors (or residuals) follow a normal distribution. Figure 4-7, a histogram of these residuals, demonstrates an approximately bell-shaped distribution centered near zero, thereby confirming the normality assumption.

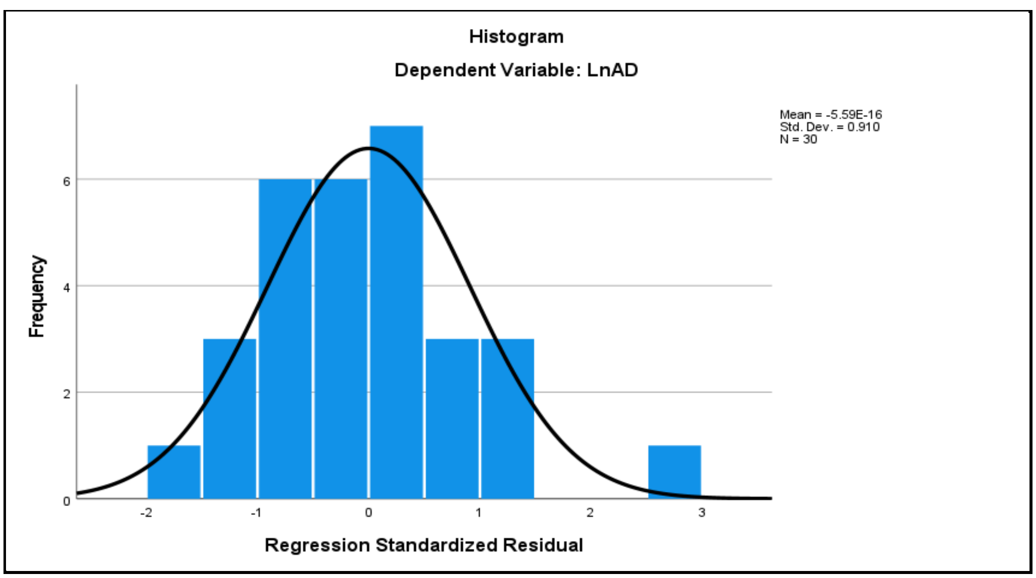

Figure 4-7: Histogram for regression standardized residual

Furthermore, the Normal P-P Plot of standardized residuals (Figure 4-8) visually confirmed that the residuals generally clustered along the diagonal line, indicating that the assumption of normality of residuals was largely met.

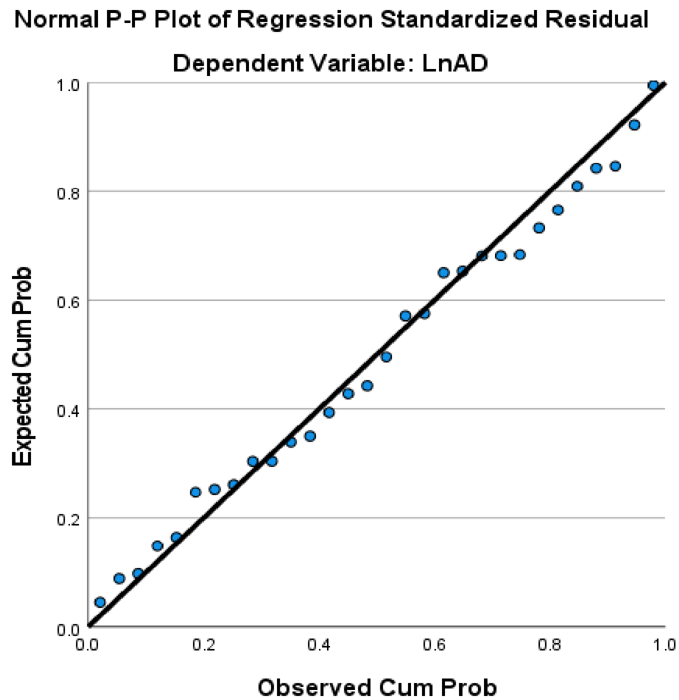

Figure 4-8: Normal P-P plot for regression standardized residual

4.6.4. Multicollinearity

Multicollinearity is a phenomenon occurring in multiple linear regressions where two or more independent variables are highly correlated with each other. To ensure the reliability of coefficient estimates, independent variables within a multiple regression model should ideally not exhibit high inter-correlation. This assumption is commonly assessed by inspecting the correlation matrix between independent variables and, more formally, by calculating Variance Inflation Factor (VIF) and Tolerance statistics. As presented in Table 4-11 (Collinearity diagnostics), all VIF values were well below the common threshold of 10, and Tolerance values were above 0.1. These results indicate that multicollinearity is not a significant issue in the selected regression model, thereby ensuring the reliability of the individual coefficient estimates.

Chapter 5: Conclusion and Recommendation

5.1. Conclusions

The investigation into the applicability of Bromilow's Time-Cost (BTC) principle for Ethiopian Roads Administration (ERA) road projects revealed that while the principle itself holds general applicability, the basic BTC model demonstrated limited predictive efficiency. This was quantitatively evidenced by an R^2 value of 0.151, indicating that the model accounted for only approximately 15.1% of the variance in construction duration. Consequently, this suggests the basic BTC model lacks sufficient predictive power for precise estimations within the context of Ethiopian road construction projects.

Beyond the basic BTC model, this study explored ten additional regression models, utilizing both untransformed and natural logarithm (Ln) transformed data, to identify more robust time cost relationships for Ethiopian road projects. Notably, the untransformed models (specifically, Compound, Growth, and Exponential equations) achieved a slightly higher R^2 of 0.285 (p=0.002) compared to the basic BTC model (R^2 =0.151) and the LnTime-LnCost cubic equation (R^2=0.272). While this indicates a marginal improvement in fit for describing time-cost relationships, the overall explanatory power of these alternative simple regression models remained limited.

In parallel with the time-cost analysis, this study also investigated the empirical relationships between construction duration and road length for Ethiopian road construction projects. Ten different regression forms were tested using both untransformed and natural logarithm (Ln) transformed data. For untransformed data, the Compound, Growth, and Exponential equations yielded an R^2 value of 0.162 (significant at the 5% level). The Ln-transformed Time-Length relationship, specifically the Quadratic equation for Ln Time – Ln Length, showed a relatively moderate fit with an R^2 value of 0.209. Overall, these simple regression models consistently demonstrated limited explanatory power, falling short of the precision required for reliable duration prediction based solely on road length.

Addressing the limitations of simpler regression approaches, this study successfully developed a robust Multiple Linear Regression (MLR) model for conceptual construction duration prediction in Ethiopian road projects. Comprehensive analyses, utilizing both untransformed and natural logarithm (Ln)-transformed data, demonstrated a significant improvement in predictive capabilities. The Ln-transformed MLR model notably outperformed its untransformed counterpart, achieving a superior R^2 value of 0.786 and a remarkably low Mean Absolute Percentage Error (MAPE) of 1.43%. In contrast, the untransformed MLR model, while showing a substantial R^2 of 0.769, exhibited a considerably higher MAPE of 10.23%.

Consequently, the Ln-transformed MLR model was selected as the optimal duration prediction tool for ERA road projects. This model integrates five independent variables: Actual Cost (AC), Highway System (HS), Project Scope (PS), Site Accessibility (SA), and Contractor Category (CON). Its validated performance metrics confirm its strong capability to provide accurate road duration predictions, offering a significantly enhanced data-driven tool for the early planning and evaluation phases of road construction projects in Ethiopia. This model is valuable for front-end predictions, serving as a policy-setting instrument or a basis for assessing planner's estimates and construction performance, though it is not intended to substitute detailed project duration estimates derived from comprehensive scheduling techniques.

5.2. Recommendations

Based on the findings of this study, the following recommendations are proposed:

- The main parties involved in Ethiopian road projects should apply the selected MLR model for benchmarking a precise construction time estimates during the early planning phase (e.g., for feasibility checks), for making construction time decisions during the bid preparation phase, and to evaluate whether the contract periods stipulated by the clients are reasonable.
- The model should be continuously updated with additional and recent project data because the regression model is based on project performance.

5.3. Recommendation for Future Research

The following recommendations are offered for future studies:

- Further research is recommended to incorporate a larger number of similar projects to enhance the reliability of the models.
- For a more comprehensive and practical tool, researchers are advised to consider additional qualitative factors such as terrain type, district/region, political stability, workforce productivity, site conditions, weather conditions, and financial and management aspects provided these are known or easily quantifiable early in the project lifecycle.
- Future researchers are encouraged to explore more advanced machine learning algorithms (e.g., Artificial Neural Networks, Support Vector Machines, Random Forests, and Gradient Boosting) to capture complex non-linear relationships, aiming for improved road duration prediction model development.

References

Abdel Aziz, A. (2009). *Time prediction for highway pavement projects using regression analysis.* Paper presented at the Construction Research Congress 2009: Building a sustainable future.

Adeyemi, A. Y., & Motlakase, B. T. (2020). Assessing the Suitability of Bromilow Time-Cost (BTC) Model to Predict Project Time on Road Construction Projects in Botswana *Civil and Environmental Research, 12*(12).

Ahmadu, H. A., Ibrahim, Y. M., Ibrahim, A. D., & Abdullahi, M. (2015). Modelling building construction durations. *Journal of Financial Management of Property and Construction, 20*(1), 65-84. doi:10.1108/jfmpc-02-2014-0004

Al-saadi, A., Zamiem, S. K., Al-jumaili, L., Jameeljubair, M., & Hashemi, H. (2017). Estimating the optimum duration of road projects using neural network model. *International Journal of Engineering and Technology, 9*(5), 3458-3469.

Al-sadi AM, Zamiem SK, & LA, A.-j. (2017). Estimating the optimum duration of road projects using neural network model. . *International Journal of Engineering and Technology, 9*(5), 3458-3469.

Al-Sultan, A. S. (2004). *Determination of Construction Contract Durationfor Public Projects in Saudi Arabia.*

Alemu, S. K. (2021). Construction time prediction model for public building projects. *Engineering, Construction and Architectural Management, 29*(5), 2183-2206. doi:10.1108/ecam-11-2020-0975

Ameyaw, C., Mensah, S., & Arthur, Y. D. (2012). *Applicability of Bromilow's time-cost model on building projects in Ghana.* Paper presented at the Proc 4th West Africa Environment Research Conference, Abuja, Nigeria.

Assefa, A. (2008). Time-Cost Relationships for Public Road Construction Projects in Ethiopia. *Unpublished Master's thesis, Addis Ababa University*.

Aswed, G. K., Ahmed, M. N., & Mohammed, H. A. (2022). Predicting initial duration of project using linear and nonlinear regression models. *International Journal of Advanced Technology and Engineering Exploration, 9*(97), 1730.

Attal, A. (2010). *Development of neural network models for prediction of highway construction cost and project duration.* Ohio University,

Ayele, T. (2019). *CONSTRUCTION PERFORMANCE MEASUREMENT PRACTICE IN THE ROAD SECTOR: IN CASE OF ADDIS ABABA CITY ROAD AUTHORITY PROJECTS.* Addis Ababa University Press Addis Ababa, Ethiopia,

Bhokha, S., & Ogunlana, S. O. (1999). Application of artificial neural network to forecast construction duration of buildings at the predesign stage. *Engineering, Construction and Architectural Management*.

Car-Pušić, D., & Radujković, M. (2010). Model for Proposed Construction Time for Transportation Facilities. *Suvremeni promet, 30*(1/2), 127.

Chan, A. P. (1999). Modelling building durations in Hong Kong. *Construction Management & Economics, 17*(2), 189-196.

Chan, A. P. (2001). Time–cost relationship of public sector projects in Malaysia. *International journal of project management, 19*(4), 223-229.

Chan, D. W., & Kumaraswamy, M. M. (1995). A study of the factors affecting construction durations in Hong Kong. *Construction Management and Economics, 13*(4), 319-333.

Chan, W.-m. (1998). Modelling construction durations for public housing projects in Hong Kong. *香港大學學位論文*, 1-0.

Choudhury, I. (2013). Development of a Prediction Model for Construction Time of Educational Projects in Texas. *Professional Constructor: The Journal of the American Institute of Constructors, 37*(1).

Choudhury, I., Khan, M., & Matin, A. (2002). *Relationship between construction time and project cost of health sector construction works in Bangladesh.* Paper presented at the Proceedings of the 38th Annual Conference of the Associated Schools of Construction.

Czarnigowska A, & A., S. (2014). Estimating construction duration for public roads during the preplanning phase. . *Journal of Engineering, Project, and Production Management, 4*(1), 26-35.

Czarnigowska, A., & Sobotka, A. (2013). Time-cost relationship for predicting construction duration. *Archives of Civil and Mechanical Engineering, 13*, 518-526.

Czarnigowska, A., & Sobotka, A. (2014). Estimating construction duration for public roads during the preplanning phase. *Journal of Engineering, Project, and Production Management, 4*(1), 26-35.

Endut, I. R., Akintoye, A., & Kelly, J. (2006). *Relationship between duration and cost of Malaysian construction projects.* Paper presented at the Proceedings of the International Conference in the Built Environment in the 21st Century (ICiBE 2006).

Field, A. (2024). *Discovering statistics using IBM SPSS statistics*: Sage publications limited.

Finkel, G. (2015). *The economics of the construction industry*: Routledge.

Gautam, N., & Banstola, R. (2022). Estimation of Contract Duration for Construction on Urban Road Projects within Pokhara Metropolitan City.

Girma, T. (2011). The History of Road Construction in Ethiopia. In: Ethiopian Roads Authority.

Golizadeh, H., Sadeghifam, A. N., Aadal, H., & Majid, M. Z. A. (2015). Automated tool for predicting duration of construction activities in tropical countries. *KSCE Journal of Civil Engineering, 20*(1), 12-22. doi:10.1007/s12205-015-0263-x

Helvacı, A. (2008). *Comparison of parametric models for conceptual duration estimation of building projects.*

HIEP, D. V. H., SODIKOV, J., SADIKOV, I., & MUSULMANOV, Q. (2023). Project Duration Estimation for Road Construction in Developing Countries: An Empirical Analysis *Proceedings of the Eastern Asia Society for Transportation Studies, 14*.

Hoffman, G. J. (2005). *Estimating performance time for air force military construction projects*. Retrieved from

Hoffman, G. J., Thal Jr, A. E., Webb, T. S., & Weir, J. D. (2007). Estimating performance time for construction projects. *Journal of Management in Engineering, 23*(4), 193-199.

Irfan, M., Khurshid, M. B., Anastasopoulos, P., Labi, S., & Moavenzadeh, F. (2011). Planning-stage estimation of highway project duration on the basis of anticipated project cost, project type, and contract type. *International journal of project management, 29*(1), 78-92.

Jekale, W. (2004). *Performances for public construction projects in developing countries: Federal road and building projects in Ethiopia.* Doctoral Dissertation, Norwegian university of science and technology, Norway,

Jiang, Y., & Wu, H. (2007). A method for highway agency to estimate highway construction durations and set contract times. *International Journal of Construction Education and Research, 3*(3), 199-216.

Jin, R., Han, S., Hyun, C., & Cha, Y. (2016). Application of Case-Based Reasoning for Estimating Preliminary Duration of Building Projects. *Journal of Construction Engineering and Management, 142*(2). doi:10.1061/(asce)co.1943-7862.0001072

Kaka, A., & Price, A. D. (1991). Relationship between value and duration of construction projects. *Construction Management and Economics, 9*(4), 383-400.

Kaleem, S., Irfan, M., & Gabriel, H. F. (2014). *Estimation of highway project duration at the planning stage and analysis of risk factors leading to time overrun.* Paper presented at the T&DI Congress 2014: Planes, Trains, and Automobiles.

Kassa, Y. F. (2020). Determinants of infrastructure project delays and cost escalations: the cases of federal road and railway construction projects in Ethiopia. *American Scientific Research Journal for Engineering, Technology, and Sciences, 63*(1), 102-136.

Kothari, C. (2004). Research Methodology 2 rev. In: NEW AGE INTERNATIONAL (P) LIMITED, PUBLISHERS.

Le-Hoai, L., Dai Lee, Y., & Nguyen, A. T. (2013). Estimating time performance for building construction projects in Vietnam. *KSCE Journal of Civil Engineering, 17*(1), 1-8.

Le- Hoai, L., & Dai Lee, Y. (2009). Time- cost relationships of building construction project in Korea. *Facilities*.

Liben, S. M., Belachew, D. A., & Elsaigh, W. A. (2024). Comparing advanced and traditional machine learning algorithms for construction duration prediction: a case study of Addis Ababa's public sector. *Engineering Research Express, 6*(4). doi:10.1088/2631-8695/ad979f

Long, L.-H., Lee, Y.-D., & Cho, J.-W. (2009). Construction of time-cost model for building projects in Vietnam. *Korean Journal of Construction Engineering and Management, 10*(3), 130-138.

Mahamid, I. (2019). The development of regression models for preliminary prediction of road construction duration. *Int. J. Eng. Inf. Sys, 3*, 14-20.

Mak, M., Ng, S., Chen, S., & Varnam, M. (2000). *The relationship between economic indicators and Bromilow's time-cost model: a pilot study*. Paper presented at the 16th Annual ARCOM Conference"(A. Akintoye, ed.).

Menbere, L. (2019). *TIME-COST RELATIONSHIPS MODEL IN ROAD CONSTRUCTION*. ADDIS ABABA SCIENCE AND TECHNOLOGY UNIVERSITY,

Menbere, Y. (2022). *CONSTRUCTION DURATION PREDICTION MODEL THE CASE OF HIGHWAY CONSTRUCTION PROJECT IN AMHARA REGION, ETHIOPA.*

Mensah, I., Adjei-Kumi, T., & Nani, G. (2016). Duration determination for rural roads using the principal component analysis and artificial neural network. *Engineering, Construction and Architectural Management, 23*(5), 638-656.

Mensah, I., Nani, G., & Adjei-Kumi, T. (2016). Development of a model for estimating the duration of bridge construction projects in Ghana. *International Journal of Construction Engineering and Management, 5*(2), 55-64.

Mohamed, B., & Moselhi, O. (2022). Conceptual estimation of construction duration and cost of public highway projects. *J. Inf. Technol. Constr., 27*(29), 595-618.

Naik, M. G., & Kumar, D. R. (2015). *Construction Project Cost and Duration Optimization Using Artificial Neural Network*. Paper presented at the Aei 2015.

Nevett, G., Goodrum, P. M., & Littlejohn, R. L. (2021). Understanding the effect of bid quantities, project characteristics, and project locations on the duration of road transportation construction projects during early stages. *Transportation Research Record, 2675*(2), 121-134.

Njeem, W. (2012). *Computer Integrated Model to Estimate the Construction Cost and Duration of Building Projects at Their Feasibility Stage.* Université d'Ottawa/University of Ottawa,

Nkado, R. N. (1995). Construction time-influencing factors: the contractor's perspective. *Construction Management and Economics, 13*(1), 81-89.

Odabasi, E. (2009). Models for estimating construction duration: An application for selected buildings on the METU campus. *Unpublished MSc Thesis). School of Natural and Applied Sciences of Middle East Technical University. Available at:* https://etd*. lib. metu. edu. tr/upload/12610696/index. pdf [Accessed on 21 November 2018]*.

Ogunsemi, D., & Jagboro, G. (2006). Time- cost model for building projects in Nigeria. *Construction Management and Economics, 24*(3), 253-258.

Peško, I., Mučenski, V., Šešlija, M., Radović, N., Vujkov, A., Bibić, D., & Krklješ, M. (2017). Estimation of costs and durations of construction of urban roads using ANN and SVM. *Complexity, 2017*.

Pokhrel, P., Misra, J., Nepal, D. B., & Adhikari, P. (2021). Estimation of construction duration for roads and bridges in Nepal. *Nepal Journal of Civil Engineering, 1*(1), 41-49.

Purnus, A., Bodea, C., & Stoian, M. (2022). *The Bromilow Time-Cost model: a case study for the infrastructure projects in Romania.* Paper presented at the IOP Conference Series: Materials Science and Engineering.

Qiao, Y., Labi, S., & Fricker, J. D. (2019). Hazard-based duration models for predicting actual duration of highway projects using nonparametric and parametric survival analysis. *Journal of Management in Engineering, 35*(6), 04019024.

Rahman, M., Lee, Y. D., Ha, D. K., & Chun, Y. H. (2014). Development of time-cost models for building construction projects in Bangladesh. *Journal of Construction Engineering and Project Management, 4*(3), 13-20.

Son, J., Khwaja, N., & Milligan, D. S. (2019). Planning-phase estimation of construction time for a large portfolio of highway projects. *Journal of Construction Engineering and Management, 145*(4), 04019018.

Tadewos, S. G., & Patel, D. (2018). Factors influencing time and cost overruns in road construction projects: Addis Ababa, Ethiopian scenario. *International Research Journal of Engineering and Technology (IRJET), 5*, 177-180.

Tawiah, M. (2015). *Using Bromilow's model and other regression model to predict the duration of feeder road projects in Ghana.*

Tegegne, M., & Mengistu, G. (2019). An Overview of Road Infrastructure Development in Ethiopia. *Journal of Infrastructure Development, 11*(1), 1-15.

Titirla, M., & Aretoulis, G. (2019). Neural network models for actual duration of Greek highway projects. *Journal of Engineering, Design and Technology, 17*(6), 1323-1339.

Titirla, M., & Aretoulis, G. (2020). *Comparison of linear regression and neural network models to estimate the actual duration of Greek highway projects.* Paper presented at the XIV Balkan Conference on Operational Research, Virtual BALCOR 2020.

Titirla, M., Larbi, W., & Aretoulis, G. (2021). Prediction methods for the actual duration of Greek highway projects. *WSEAS Transactions on Business and Economics, 18*, 1389-1396.

Velumani, P., Nampoothiri, N., & Urbański, M. (2021). A comparative study of models for the construction duration prediction in highway road projects of India. *Sustainability, 13*(8), 4552.

Walker, D. H. (1995). An investigation into construction time performance. *Construction Management and Economics, 13*(3), 263-274.

Walker, D. H. T. (1994). *AN INVESTIGATION INTO FACTORS THAT DETERMINE BUIWING.* Royal Melbourne Institute of Technology,

Waziri, B., & Nurudeen, Y. (2014). Application of Bromilow's Time-Cost (BTC) Model to Highway Projects in Nigeria. *Arid Zone Journal of Engineering, Technology and Environment, 10*(1), 1-11.

Waziri, B. S., Jibrin, A. T., & Kadai, B. (2014). Functional duration models for highway construction projects in Nigeria. *Arid Zone Journal of Engineering, Technology and Environment, 10*, 13-23.

Williams, R. C. (2008). *The development of mathematical models for preliminary prediction of highway construction duration.* Virginia Tech,

Annex-1: Road project parameters, coding, and measuring value

Table 0-1: Road project parameters

No.	CD	CC	AD	AC	HS	PS	SA	GDS	RC	RL	RW	LN	CON	DS	CM	CN	BY	IN	BG
P1	548	592,084,400.00	1076	603,551,142.09	0	0	1	5	0	34.4	10	2	0	0	0	73	0	0	0
P2	1095	1,558,150,706.39	1004	897,005,974.61	0	0	1	5	1	102	12	2	0	1	0	234	0	0	0
P3	1095	1,133,472,329.44	2445	1,259,227,607.20	1	1	0	4	1	91.6	14	2	1	1	1	154	0	0	0
P4	913	1,337,718,925.93	1634	1,549,906,444.13	0	1	1	4	1	82.89	12	2	0	0	1	187	1	0	0
P5	1096	382,559,000.41	1268	467,316,188.47	0	0	1	3	2	53.9	14	2	0	0	1	58	0	0	0
P6	1095	1,466,787,322.35	1295	1,399,064,061.64	0	1	1	4	0	53.9	10	2	0	0	0	168	0	0	0
P7	1095	754,212,552.61	1556	1,050,079,578.00	1	1	1	4	1	70.45	12	2	1	1	1	233	1	1	0
P8	731	661,753,584.55	1125	945,236,784.00	0	0	1	6	2	18.7	14	2	0	1	1	22	1	1	0
P9	1095	1,520,365,013.02	2554	1,624,715,707.00	1	1	0	5	1	74.5	14	2	1	0	1	173	0	0	0
P10	1095	1,404,717,193.38	1325	1,624,715,707.00	0	1	0	4	1	74.5	19	2	0	1	1	255	1	0	0
P11	1093	786,796,548.12	1125	890,255,467.23	0	0	1	5	1	47	8	2	0	1	1	61	1	1	0
P12	1095	1,544,036,614.16	1387	1,914,915,705.74	0	0	1	5	1	105.7	12	2	0	1	1	282	0	1	1
P13	1464	1,530,020,366.78	2269	1,721,325,325.00	1	1	0	5	1	97.6	14	2	0	1	1	165	1	0	0
P14	1095	150,023,688.11	1468	890,255,467.23	0	0	1	4	1	10	16.5	2	1	1	1	8	1	0	0
P15	1461	1,907,769,988.45	2190	1,892,544,315.23	1	1	0	5	1	72.3	21	2	1	0	1	123	1	1	0
P16	1093	1,565,456,877.21	1462	1,754,635.25	0	0	1	6	1	12.5	23	2	1	1	1	6	1	0	0
P17	1462	1,652,857,321.54	2895	1,981,935,668.20	1	1	0	5	1	72.2	21.5	2	0	0	1	110	1	1	0
P18	770	304,063,258.15	1260	301,180,117.00	1	0	1	4	1	10.5	23	2	0	1	1	13	1	0	0
P19	913	1,925,451,264.41	1634	1,998,422,035.29	1	1	0	4	1	50.4	23	2	0	0	1	28	0	0	0
P20	1275	1,540,235,699.85	1860	1,590,546,235.58	1	1	1	6	2	59	21	2	0	1	1	109	0	0	0
P21	1461	1,434,007,400.00	1790	1,437,500,058.00	1	1	1	6	0	5.38	21	6	0	0	1	0	0	0	0
P22	1095	1,967,496,759.60	1660	201,521,012.33	1	0	1	5	1	72.64	14	2	0	1	1	96	1	0	0
P23	1095	801,212,552.61	1596	950,059,444.60	0	1	1	4	1	76.6	19	2	1	1	1	168	1	0	1
P24	1095	1,704,769,988.45	1862	1,829,354,752.22	1	1	0	4	1	92.23	21	2	0	1	1	176	1	1	0
P25	913	100,332,597.22	2159	267,079,421.58	1	1	0	6	1	35.82	6	2	1	0	0	22	0	0	1
P26	1095	740,685,321.21	1765	761,601,163.24	1	0	0	5	2	94.5	10	2	0	1	1	140	0	0	0
P27	935	6,153,822.50	1095	7,184,539.55	0	1	1	6	1	27	6	2	0	1	1	5	0	0	0
P28	913	210,943,656.53	2030	216,400,339.53	1	0	0	4	1	90.89	12	2	1	0	1	74	0	0	0
P29	1278	204,626,157.84	2058	216,400,339.53	1	0	1	4	1	148.2	11	2	1	0	1	277	0	0	0
P30	1095	478,385,156.18	2557	500,011,114,991.0	1	1	0	5	0	90.5	10	2	1	1	1	32	0	0	0

Table 0-2: Road project parameters, label & measuring value

S.No.	Project Parameters	Label	Parameter Type	Value
1	Actual Cost	AC	Numerical	Ethiopian Birr
2	Highway System	HS	Categorical	1 if Asphalt concrete, 0 if Double surface treatment
3	Project Scope	PS	Categorical	1 if new construction, 0 if Rehabilitation
4	Site Accessibility	SA	Categorical	1 if Good access, 0 if Poor access
5	Geometric Design Standard	GDS	Categorical	1 if DS1, 2 if DS2,…,6 if DS6
6	Road Classification	RC	Categorical	0 if Main Road, 1 if Link Road, 2 if Trunk Road
7	Road Length	RL	Numerical	Kilo Meter
8	Road Width	RW	Numerical	Meter
9	Lane Number	LN	Numerical	Number (Pcs.)
10	Contractor Category	CC	Categorical	1 if Local contractor, 0 if International
11	Delivery System	DS	Categorical	1 if DBB, 0 if DB
12	Contract Method	CM	Categorical	1 if Item rate, 0 if Lump sum Contract
13	Culvert Number	CN	Numerical	Number (Pcs.)
14	Availability of Bays	BY	Categorical	1 if at least one Bay is available, 0 otherwise
15	Availability of Intersection	IN	Categorical	1 if at least one Intersection is available, 0 otherwise
16	Availability of Bridge	BG	Categorical	1 if at least one Bridge is available, 0 otherwise
17	Actual Duration	AD	Numerical	Cal. Days

Table 0-3: Project code and name

Project code and name			
P1	Hawassa-Bishan Guracha (Tikur Wuha) Project	P16	Azezo-Gonder Road Construction Project
P2	Chereti-Goroboksa-Gorodamole Road Project; Lot Il: Hageremekor-Kundi	P17	Gambela-Abobo-Gog-Dima,Lot-2 Pugnwido-Gog-Gilo River D.B Road Project
P3	Mizan – Dima–Boma Road Upgrading Project Contract 1: Mizan – Dima Contract	P18	Yabelo- Bypass Road Project
P4	Injibara-Chagni Pawi Junction-Fendika-Ayma Road Project, Contract 2: Pawi Junction-Fendkia-Ayma project	P19	Iste-Simada Road Construction Project
P5	Ambo — Gedo overlay road project.	P20	Muketuri-Kokeb Mesk Road Project
P6	Design and Construction of F6 Junction – F4 Junction Road Project	P21	Abay River Brigde And Pavment Works
P7	Contract 1: Endaselassie-Dedebit	P22	Morka-Gircha-Chercha Road Project
P8	Construction works of Metu bypass road project	P23	Endaselassie-Dejena-Dansha Road project Contract 2: Dedebit-Adiremet
P9	Jinka-hana DB project-lot 1 jinka-mendir road project	P24	Bilalo-Kursa-Arsi Negele Road Project
P10	MESLE-KORI-TERU road project ,Contruct 2	P25	Seru – Shenen Wenz – Sheik Hussein Road Project, Lot II
P11	SANJA-KERAKER ROAD PROJECT	P26	Ageremariam – Yabelo Road
P12	Fik-Hamero-Imi Bale road project lot 2	P27	Seru-ShenenWonz –Shekussien, Lot I
P13	GODE-KELNFO-FERER-LOT 1	P28	Ziway – Butajira- Gubre Road Upgrading project
P14	Halaba-ALEMGEBEYA-WILBARG	P29	Nejo-Jarso-Begi Road upgrading project
P15	Atat Mazoria-Gunchire-Lera Road Project	P30	Awash – Kulubi –Dire Dawa/Harar Overlay Road Project, Contract 1: Awash – Meiso

Annex-2: Data collection tool

Table 0-4: Data collection tool

Parameters	Value	Description
Project Name		*The name of the project as per the contract document*
Projects Location/Region		*Specific location (region)of the project*
Contract Duration		*Contract duration of the project in calendar days*
Contract Cost		*Contract cost of the project in ETB*
Actual Duration		*Actual completion duration of the project in calendar days*
Actual Cost		*Actual cost of the project in ETB*
Type of Highway System		*Type of Highway system: (i) Asphalt Concrete road; (ii) Double Surface Treatment road; (iii) Gravel Surface road*
Project Scope		*Project Scope categories: (i) New construction; (ii) Rehabilitation/reconstruction; (iii) Resurfacing/renewal,*
Site Accessibility		*Project site accessibility category: (i) Good Access; (ii) Poor access*
Geometric Design Standard		*Geometric Design Standard categories as per ERA manual: (i) DC1;(ii) DC2; (iii) DC3; (iv) DC4;(v) DC5; (vi) DC6*
Road Functional Classification		*Road functional classification options: (i) Trunk road; (ii) Link road; (iii) Main Access road; (iv)if other please specify*
Construction Period		*The commencement and completion period of the project (dd/mm/yy)*
Length of the Road		*The length of the road in Km*
Road Width		*The width of the road in m*
Number of Lanes		*The number of lanes of the road*
Contractor Category		*The contractor category options: (i) Domestic (Local) Contractor; (ii) International Contractor*
Project Delivery System		*The project delivery methods options: (i) Design Bid Build (DBB); (ii) Design Build (DB); (iii) Construction Management Consultancy; (iv) if other please specify*
Tender Type		*The tender type options: (i) Open bid; (ii) Limited/selective bid; (iii) Negotiated; (iv) if other please specify*
Contract Method		*The contract method options: (i) Lump-sum; (ii) Item rate; (iii) cost plus types; (iv) if other please specify*
Number of Culverts		*The number of culverts available in the road segment*

Availability of Bays		*The number of bays available in the road segment*
Availability of Intersections		*The number of intersections available in the road segment*
Availability of Bridges		*The number of bridges available in the road segment*