ADDIS ABABA UNIVERSITY

ADDIS ABABA INSTITUTE OF TECHNOLOGY

AFRICAN RAILWAY CENTER OF EXCELLENCE

BALLAST DEGRADATION MODELING

CASE STUDY OF ADDIS ABABA-LIGHT RAIL TRANSIT

A Thesis in Railway Engineering (Civil Infrastructure)
By

Alemwork Eshetie

Advisor: Dr. Tezera Firew Azmatch

Co Advisor: Mr. Anteneh Zewdu

A thesis submitted to Graduate Studies of Addis Ababa University in the Partial Fulfillment for Degree of Master of Science in Railway Engineering (Civil Infrastructure)

August, 2025

The undersigned members of the Addis Ababa Institute of Technology under the department of Railway Engineering have examined the thesis entitled **'Ballast Degradation Modelling: Case study of Addis Ababa Light Rail Transit'** presented by **Alemwork Eshetie**, a candidate of Master of Science in Railway Engineering (Civil Infrastructure) and hereby certify that is worthy of acceptance.

THE BOARD OF EXAMINERS

Dr.Tezera Firew Azmatch		
Advisor	____________________	____________________
	Signature	Date Date
Mr. Anteneh Zewdu	____________________	____________________
Co Advisor	Signature	Date
Dr. Tensay G/Medhin	____________________	____________________
Internal Examiner	Signature	Date
Dr. Henok Fikr	____________________	____________________
External Examiner	Signature	Date
Mr. Binyam Ayalew	____________________	____________________
Chairperson of Department	Signature	Date

DECLARATION

I hereby declare that the work, which is being presented in this thesis, entitled "**Ballast Degradation Modelling: Case study of Addis Ababa Light Rail Transit**", is original work of my own, and has not been presented for a degree in any other university; and that all the sources of the material used for the thesis have been duly acknowledged.

______________________ ________________

Alemwork Eshetie Date

This is to certify that the above declared made by the candidate is correct to the best of my knowledge.

__________________________ ________________

Dr. Tezera Firew Azmatch (Advisor) Date

______________________ ________________

Mr. Anteneh Zewdu (Co Advisor) Date

ACKNOWLEDGEMENTS

Selecting this research title and completing the study was not an easy task. However, these challenges were passed with participation of many people in one or the other way. I am very grateful to express my deepest gratitude to my advisor Dr.Tezera Firew Azmatch for his unreserved assistance and constructive comments at all stages of this thesis. I would like to extend my deepest gratitude to my co-Advisor Mr. Anteneh Zewdu for his support, and all his unreserved assistance in giving me supportive ideas, helpful comments and sharing materials throughout the study. Furthermore, I would like to express my gratitude to Mr. Tadewos Worku and other staffs from Civil Infrastructure department of Addis Ababa Light Rail Transit for providing me with necessary data for my study.

ABSTRACT

Railway infrastructure should be maintained regularly to improve its performance. This operation requires a large investment and shall be planned well to avoid human and financial resource waste. Various studies found that most of the track geometrical irregularities that seek maintenance are related to the degradation of the underlying ballast. This implies that for planning track geometry maintenance the understanding of ballast degradation mechanism is crucial. In this research, a ballast degradation model that characterizes ballast quality degradation and predicts ballast degradation from different parameters is developed for Addis Ababa Light Rail Transit (AALRT). This paper uses a statistical approach for ballast degradation modeling. Fractal dimensioning of track longitudinal level was applied to determine ballast quality index and to characterize ballast quality degradation. Previously recorded inspection data of track longitudinal profile were collected from Ethiopian Railway Corporation (ERC) and used to develop the model. The irregularities of the vertical track profile are quantified as a ballast quality index. This ballast quality index is used as a basis for developing ballast degradation model. Two ballast regression models were developed for straight and curved sections of AALRT. This degradation model is useful for AALRT infrastructure managers for understanding and predicting ballast quality and it will upgrade the performance of the railway maintenance operation.

Keywords: Railway ballast; ballast quality index; fractal analysis; ballast degradation model

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACRONYMS AND ABBREVIATIONS

Term	Explanation / Meaning / Definition
AALRT	Addis Ababa Light Rail Transit
AI	Artificial Intelligence
ANFIS	Adaptive Network-based Fuzzy Inference System
ANN	Artificial Neural Network
ANOVA	Analysis of Variance
BBT	Ballast Box Test
BDR	Ballast Degradation Rate
BQI	Ballast quality index
CWR	Continuous Welded Rail
EMGT	Equivalent Million Gross Tone
ERC	Ethiopian Railway Corporation
IBQI	improved ballast quality index
LAA	Loss Angeles Abrasion value
LLA	Loss Angeles Abrasion test
LRVs	Light Rail Vehicles
MA	Micro-Deval Abrasion value
MGT	Million Gross Ton
OCC	Operation Control Centre
OLS	Ordinary Least Squares
TGI	Track Geometry Index
WLS	Weighted Least Squares ,
2SLS	two-stage least squares

CHAPTER ONE

1. INTRODUCTION

1.1. Background of the Study

Railroad is the most efficient transportation mode for moving goods and passengers on the earth’s surface(1). Due to low construction cost, recognized construction technology and ease of maintenance and repair Conventional ballasted track (Fig 1-1) has been developed worldwide as the main type of rail structure. The Ballast is a component of railway sub structure that provide functions of resilience against the accumulation of differential track settlements, used as water drainage pathways and storage space for fine-grained particles (fines). Further, it provides easy adjustment of track geometry, preserving the safety and riding comfort of train operation and diffuses the dynamic loading of the train to maintain the long-term stability of railway sub grade.

Railway track deteriorates over time and these leads to massive failures with an enormous financial lose. As a result, detection and rectification of track defects are major issues in the railway industry. It is very important to decide when and how to perform maintenance operations and how to allocate the resources such as labor, materials, machines and funds to the parts of the system with highest need. Railway track defects can be categorized into one of two groups: track structural defects and track geometry defects. Track structural defects are the defects on the track constituents that include the condition of the rail, sleeper, fastening systems and the sub grade. Track geometry defects are the deterioration of the track geometry parameters such as profile, alignment, gage etc. The two categories of the track defects may be caused one by the other, such that the track geometry defects are resulted from the structural conditions of the constituents such as the ballast(2).

Railway track geometry degrades and deteriorates under traffic loads that can lead to failure, resulting in massive casualties and loss of life and damage to infrastructure assets. To show the degradation pattern of the track (3) develop a curve that depicts different phases of the track deterioration . From this curve, the track goes through three phases namely; youth phase, intermediate phase and old phase. The deterioration of the track in its early life after completion of construction or completion of major renewal works denotes the youth phase. In this phase, deterioration is observed due to initial settlement of track and the duration is quite short. Intermediate phase occurs during majority of track lifetime and characterized by linear deterioration pattern. The third phase is characterized by increasingly rapid deterioration due to old condition of track and Quasi-exponential deterioration pattern is observed in the phase. Ballast degradation is often a primary factor contributing to the development of track geometry degradation and irregularity. From (3) it is shown that the three phases of the track deterioration is resulting from the degradation of the ballast. The youth phase is due to re arrangement of ballast particle and the intermediate phase is because of linear degradation of the ballast due to traffic loading further the Quasi-exponential deterioration pattern is due to highly degraded ballast resulting from the repetition of tamping with increased life of ballast.

Hence, ballast degradation is the basic cause of track geometry degradation and it provides adjustment of the track geometry, the study of degradation of ballast needs special attention to overcome severe losses due to track deterioration. Ballast degrades due to several factors such as traffic passage, loading and other environmental factors. Ballast degradation erodes the interlocking between particles, resulting in insufficient bearing capacity, differential settlements in rail tracks, and obstructing the free drainage of the ballast layer, which results in the loss of track quality and necessitating maintenance work.

Crushing of ballast materials due to repetitive rail loading significantly affects the performance of the ballast layer, which may lead to major accidents. From previous studies, it was found that the major amount of maintenance fund of railway track goes to the substructure part of the rail track and most of which is devoted to ballast maintenance and its replacement(4).

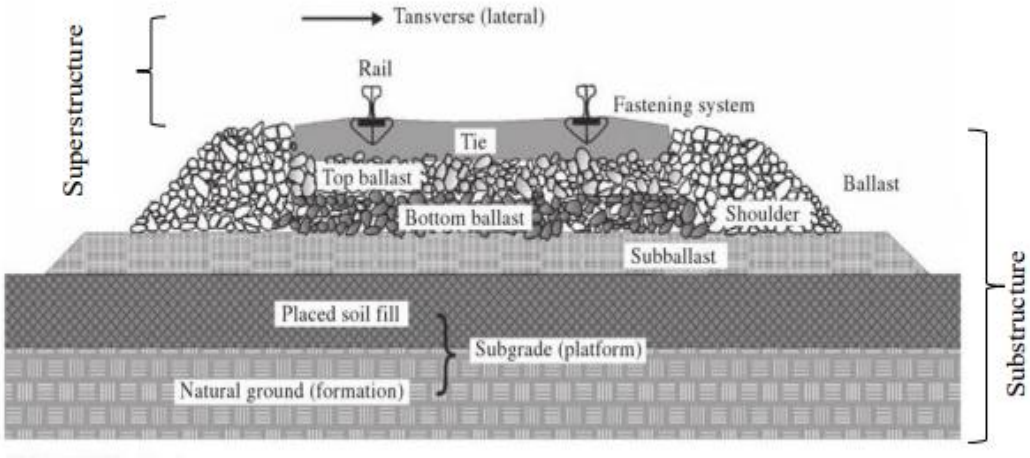


Figure 1- 1: conventional ballasted track

1.2. Problem Statement

Rail lines shall be inspected and maintained periodically to meet planned safety and quality performance. These maintenance activities require large amount of time, machines as well as human and financial resources. Due to this, railway infrastructure managers need to plan these operations in advance. To do this models that can predict ballast degradation and its maintenance requirement shall be developed by considering appropriate parameters. As the main track component that can adjust the track geometry is the ballast, Ballast degradation model that can predict the ballast quality of the future for planning maintenance activities is to be developed. Most of the research done in the previous time on ballast quality investigation were focused on laboratory experiments. Such studies include (5),(6),(7),(8).

Due to lack of suitable laboratory equipment and others reasons, this method cannot be always applicable and sufficient to study ballast degradation. Hence, ballast particles are larger aggregates they require large-scale laboratory tests to investigate their degradation pattern. The other reason is that Laboratory experiments cannot consider all the factors that affect ballast degradation in the field; for instance environmental factors, the cut or fill condition of the sub grade, seasonal variation and the gradient of the track are some of those factors. To fill the above stated gaps in the study area this study uses data recorded on the actual track to investigate and model ballast degradation pattern. In return, the developed model can be used to predict the ballast degradation from different influencing parameter considerations to plan maintenance activities. The author did not see ballast degradation model developed for ballast quality prediction used in maintenance planning purpose in AALRT.

1.3. Objective

Main Objective

Developing a ballast degradation model for AALRT that can predict the ballast quality for different sections at different times and conditions

Specific Objectives

- ✓ To identify factors that influence Ballast quality
- ✓ To develop Ballast quality index (BQI) by fractal analysis
- ✓ To investigate the effects of different contributing factors that may influence the ballast quality index
- ✓ To develop a regression model that can predict ballast quality from the influencing Parameters

1.4. Significance of the Research

This research work is an original work and the results of the study will be of great benefit

- ✓ For Ethiopian railway infrastructure managers to best plan track geometry maintenance schedule
- ✓ For maintenance crew to reduce the time and effort they invest for inspection to check whether maintenance is needed or not
- ✓ For the railway industry as a general to reduce the cost of time, human resource and machinery that imparts the overall operational efficiency of the industry
- ✓ For students to understand ballast degradation pattern and the relations of ballast quality degradation with different variables and give them the research insight towards the area. Further it gives an insight to dig new methodologies of investigating ballast degradation in addition to the laboratory techniques

1.5. Scope and Limitations of the Research

This research covers different aspects of railway Ballast Quality degradation. In this research study, Fractal analysis of Track longitudinal level is used as Ballast quality index for ballast degradation modeling and the usual method of studying ballast quality by Laboratory experiments is out of the scope of this study. The study is focused on the straight and curved section of Addis Ababa Light Rail Transit. The limitation of the research is that the model predicts the ballast quality of the future, but it cannot tell the exact time when the ballast will be recycled and the track will be maintained but by seeing the ballast quality degradation pattern it is possible to schedule maintenance activities. Data for ballast recycling is not incorporated in the study due that the ballast has not been recycled in the study area.

CHAPTER TWO

2. LITERATURE REVIEW

2.1 Introduction

Modeling track degradation is the necessary part of a track maintenance support system and is essential for predictive maintenance. Different studies have been conducted on track degradation prediction models in railway transportation system. The degradation models were developed for the whole railway track as well as for some of the track components such as the ballast. Hence, the ballast degradation model developed in this study is developed from the track longitudinal profile (track geometry) the review of track geometry degradation models and track structure (component) degradation models is essential.

2.1 Track Structure

There are two types of railway track, ballasted track and slab track. Ballasted track is the most widely used type of track due to low construction cost, recognized construction technology, ease of maintenance and repair. It consists of two parts; track Super structure (Rail, Sleeper, Rail pads, fastenings) and Sub structure (Ballast, Sub ballast, fill material and sub grade). On this type of track, Rails are mounted on sleeper, which can be wooden, steel or concrete and rail’s function is to transfer load to the sleepers. Sleepers are laid on the ballast, which provides lateral and longitudinal stability and the Ballast acts as shock absorber by reducing the vibrations. Geometry deviations of track in a range of millimeters occur through the life of the track and due to this routine maintenance such as tamping of ballast is applied to restore the track to its appropriate condition (9).

2.2. Ballast

Ballast is a selected crushed granular aggregates ranging in size from 10 to 60 millimeters that are placed as the top layer of the substructure of a ballasted track. It is a key structural part of the railroad and its main function is to transfer the induced loads from the superstructure to the subgrade without failure and provide good drainage (10). Ballast is ideally composed of coarse aggregates sourced from either blasted igneous rock, metamorphic or, well-cemented sedimentary rock quarries (11).The ballast used to support railway tracks is critical in developing and sustaining the track response characteristics and, consequently the riding quality. For ballasted tracks, the primary criterion for low maintenance expenditures is an elastic, non-cemented, stable, and weather-resistant ballast bed that is adequately laid and compacted on a stable, compacted sub-ballast and sub grade(12).

2.3. Track Geometry Degradation

Track irregularities (geometry degradation) is considered as deviations from normal projected track geometry and the exceed tolerance levels (13). The condition of the railway track has an important function in the operation of the railway system (i.e., maintenance, operational safety, and passenger comfort) and it is important to consider any irregularities and discover the phenomena that cause track deterioration and to forecast irregularities (14). Railway track degradation is commonly quantified by assessing the deterioration of geometrical parameters (15).

The geometry parameters that are widely used to represent the track condition and to plan maintenance activities can be divided into five classes: longitudinal level/profile, alignment, gauge, cant and twist (16). Longitudinal level/profile is the measure of non-uniformity of top surface of the rail measured in a short distance and the vertical deviation of this centerline is termed as longitudinal level/ Profile defect. Alignment is the track geometry of the track centerline projected onto the longitudinal horizontal plane and horizontal deviation of this centerline is termed as alignment defect. Gauge is the distance between the inner sides of the railheads and Gauge defect is the change in distance between the inner railheads. Cant (cross-level) is the difference in elevation of top surfaces of the two rails measured at specific point and the extra difference to the designed value is termed as cross level defect. Twist is the difference between two cross-levels taken at a defined distance apart and the extra difference between two cross-level measurements at a certain distance compared to the designed value is termed as twist defect(17).

From the above-described geometrical parameters, the quality of a particular track component can be determined based on one geometric parameter. For instance, a gauge signal has useful information about the condition of the fastening system, cant is indicative of rail-pads and track longitudinal level is indicative of ballast quality and that track longitudinal irregularities are highly correlated with ballast differential settlement (18). It was also shown that longitudinal level measurements are more indicative of ballast quality if analyzed with the fractal analysis method (19).

2.3. Track Geometry and Track Structure Degradation Models

degradation models are mathematical relationships between a dependent variable (deterioration or change in condition) and a set of casual variables, including design attributes, traffic loading, environmental factors, age and maintenance history(9). By analyzing the track degradation and developing a suitable degradation model a company can make proper decisions to optimize inspection times, estimate the residual life of the track components, calculate the life cycle cost, and predict a suitable time for renewal activities. According to previous studies, physical (Mechanistic), statistical and artificial intelligence approaches have been used for modeling track geometry and track structure degradation (20).

2.3.1. Physical/ Mechanistic Models

Mechanistic or physical models are based on prior physical information. They are developed by testing the mechanical properties of all the elements that make up the track. Mechanistic models of ballast degradation were based on Laboratory investigations and usually make no use of geometry data. In this approach Laboratory investigations of ballast cyclic loading using large scale triaxial tests, ballast tamping experiments and other laboratory tests have been a dominant approach to investigate the changes in mechanical characteristics of the ballast including ballast strength, fouling rate, stiffness and ballast settlement. By embedding, the results of the laboratory investigation and by considering some degradation laws Mechanistic models are developed to study the long-term degradation behavior of the ballast layer (18). The mechanistic approach is described as the modeling of mechanical reactions in the track, which result in track degradation. The most important advantage of mechanistic modeling is that by using this type of model the relationship between track responses and parameters of traffic can be properly clarified (20).

(5) Elaborated that ballast deterioration is the main factor affecting rail track degradation. The paper studied the six mechanisms comprising track degradation including dynamic forces, rail shape, sleeper spacing, sleeper support, ballast settlement, and substructure. It was argued that ballast settlement plays the most

important role in track degradation and mechanistic track degradation model was developed by considering axle load, number of axle loads, ballast type, sleeper type and size, and sub grade condition. It was found that axle load is the most important of these factors in ballast settlement and that ballast settlement is related to the fifth root of the number of axle loads. The study developed Equation 2.1 as a track degradation model based on ballast settlement.

$$S=k_s \frac{Ae}{20}((0.69+0.028L)N^{0.2}+2.7x10^{-6}N \quad \text{(2.1)}$$

Where;

A_e: passing tonnage in Million Gross Tone (MGT),

N: total number of passing axles,

Ks: a factor corresponding to the type and size of sleeper, ballast type and the condition of the sub grade

L: lift given by tamping machines

(6)studied the degradation of Japanese railway tracks and developed a mechanistic degradation model for growth of track irregularities. He evaluated track deterioration due to ballast settlement under repeated loading passage. From the study Equation 2.2 were developed to estimate the settlement of tamped tracks under frequent loading by train passage.

$$y= \gamma\ (1-e^{-\alpha x} + \beta x). \quad \text{(2. 2)}$$

Where;

y: ballast settlement (mm),

x :repeated number of loading or tonnage carried by track

γ: initial rapid settlement; it is cause by rearranging of ballast particles.

α: acceleration in which stone without spring force initiate slide.

β: lateral movement of ballast particles under sleeper, α, β and γ are coefficients.

The effect of traffic parameters (speed, load and repetition) on the long-term settlement of ballast was studied in (7). A 8.08 m three-dimensional finite element track model loaded with a moving axel was modeled by Using finite element modeling software and the model were analyzed for different values of speed, load and number of repetitions. Drucker Prager plastic model was applied in the study to represent the granular behavior of ballast and sub-ballast and Parameters for the Drucker Prager plastic model are generated from a triaxial test model using PFC (3D) discrete element modeling software.

On the model the sleeper spacing, rail gauge and characteristics of ballast material, sleeper and rail were taken from CREC Addis Ababa-Mieso railway project, testing reports and Chinese Rail Standard. From the study, it was found that the total plastic strain of ballast was affected more by the variation of speed and the variation of a number of cycles than axel loads. Increase in the speed of train movement by 20 km/hr will result in increasing the stress transferred to the sub grade by up to about 1000 kPa and increasing the number of cycles of train movement by four affects the total plastic strain of ballast more than increasing of tonnage by 5 tons.

(21) Modeled Ballast, sub-ballast, and subgrade settlement behaviors. The study found that track settlement is composed of three parts: ballast, sub-ballast, and subgrade settlement. However, the effect of ballast on track settlement is significantly higher than that of the other two factors, especially after performing a number of tamping interventions which shares the finding of(5). In this study, an exponential function was used to model ballast settlement based on wheel loads, abrasion number, fouling index, and number of load cycles. Linear functions were applied to model sub-ballast settlement and power functions were applied to model sub grade settlement.

(8) developed a new mechanistic ballast degradation model for life cycle assessment of railway ballast materials. Ballast box test (BBT) was used in the study and equation 2.3 was developed from the study.

$$IBQI\ (N) = 20.5 - 0.007\ x\ e(-3.2x10^{-7}N + 7.56) \quad \ldots\ldots\ldots\ldots.. (2.\ 3)$$

Where;

IBQI (N): Improved ballast quality index at a given number of load cycles

N: number of load cycles

(22) developed another model in 2017 from Los Angeles Abrasion and Micro-Deval Abrasion tests. From this study, lifespan of the ballast expressed in accumulated traffic flow (MGT) was defined by equation 2.4.

$$Life = 10^{6}\ e\ (8.08 - 0.0382N_{A}) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots.. (2.\ 4)$$

Where;

Life: lifespan of the ballast, expressed in accumulated traffic flow MGT

N_a: LAA + 5 MA

LAA: Los Angeles Abrasion value

MA: Mill Abrasion value

The practical application of mechanistic models requires high level of details to predict track degradation condition. In the face of uncertainty of the track behavior, missing only one influencing factor can give rise to a model generating invalid results (23)

The effect of tamping operations on the mechanical qualities of ballast bed is thoroughly analyzed by developing a tamp-sleeper-ballasts coupling model by using Discrete Element Model (DEM) and Multi body Dynamics (MBD) (24). The model was verified by field tests and was applied to analyze the macro-mesoscopic motion of ballasts and the effects of tamping parameters on the mechanical qualities of ballast bed. The results show that tamping destroys the uniformity of ballast bed, and the top layer under sleeper and the crib are the most severely affected spaces of the ballast layer.

2.3.2. Statistical models

This approach involves the analysis of many observations of track measurements collected for a long period and corresponding casual parameters including traffic, environmental and maintenance variables(9).Statistical approaches can address uncertainties more efficiently by exploring the patterns of degradation from historical recordings of track geometry data.

A statistical model is constructed from a set of input and output variables and it requires sufficient data samples to construct a relationship among the output variable and the influencing factors (18). Statistical methods such as correlation analysis, regression analysis, and stochastic processes may be applied to design a statistical degradation model (25). Since real field data are used to construct the degradation model, an accurate estimation of track degradation can be derived from statistical modeling but it does not consider the mechanical background for track components and their interactions with influencing factors. (26) Suggested that when sufficient data are available, a statistical model is preferable to a mechanistic model.

The objective of statistical based degradation models is to find a general pattern for the statistical distribution of the track geometry degradation using inspected data of track condition. For developing an accurate degradation model the basic data required are infrastructure inventory, layout and operating work history and condition measurements (27).

A generic track degradation model that can be used to predict the future condition of track was developed in (27). The study stated that by identifying the condition parameters, essential and temporary activities, suitable curves for degradation behavior and the restoration model, a quite accurate long-term prediction of track condition can be achieved. The approach employed in the study provides an overview about track degradation modeling. From the study, it was found that although different curves can be fitted on measurement data in a maintenance cycle, it is accepted that the track degradation follows approximately linear behavior in a maintenance cycle.

(28) Develop linear regression model to predict the track degradation of Melbourne tram system. Gauge was the selected parameter to develop the model. Degradation models were developed by considering track gauge deviation in previous year, curve radii, MGT, rail support, track surface, rail profile and rail type as input variables and track gauge in current year as an output variable. To determine which variables are significant in prediction of the output variable Pearson Correlation analysis were used for continuous parameters and one-way ANOVA analysis was applied for categorical variables.

From the study it was found that track gauge deviation in previous year, track surface and rail type are significant in prediction of the output variable. Four different linear regression models were developed based on the condition of track segments (repaired or unrepaired), and the type of tracks (curve or straight) and the developed models were capable of predicting the gauge values with R^2 from 0.71 to 0.91.

(18) develop a ballast degradation model from track car recording data for railway turnouts. Fractal analysis was used to determine the ballast quality index and ballast degradation rate was determined as the slope of the best-fit line to the time series of the

ballast quality index. Moreover, a multiple linear regression modeling was used to estimate the effects of different contributing variables on the ballast degradation rate. These variables include the number of past tamping, maximum permissible train speed; passing million gross tons (MGT), the weather condition (seasonality, temperature and precipitation), the current degraded state of the track geometry, the current degraded shape of the turnout, the type of the turnout and also the section of the turnout.

The model was estimated with the ordinary least squares method and all the variables in the model are significant in 99.9% confidence level. The results of estimated coefficients show that increasing speed limit, tamping, being in the crossing section, MGT, initial degradation level, warm season, higher ratio of the front to the mid-section have increasing effect on ballast degradation rate in turnouts.

The deterioration of two tangent light rail tracks found in a southern German city were evaluated and compared in (14). The standard deviation of the irregularity of the track parameters fond in DIN EN 13848-5 standard (equation 2.5) and TGI (track geometry index) developed by the Indian railways (equation 2.6) were used to evaluate geometrical quality of the tracks under study.

$$SD=\sqrt{\frac{\sum_{i=1}^{N}(X_1-\bar{X})^2}{N-1}} \quad \text{.. (2. 5)}$$

Where:

SD: standard deviation of track parameters such as gauge, twist and alignment

N: the number of signals measured on a track section

X_i: irregularity at point i [mm]

$\bar{X}$: the average value of track regularity per the given segment.

On the other hand, in the TGI system TGI were calculated based on the average of weighted indices of the different geometry parameters including unevenness index, twist index, gauge index and alignment index to quantify the quality of the track segment

$$TGI=\frac{2UI+TI+6AI}{10} \quad \text{.. (2. 6)}$$

Where:

TGI: Track Geometric Index

UI: unevenness index

TI: twist index

GI: gauge index

AI: alignment index.

These indices result from exponential functions that use calculated standard deviations of each parameter at different states and conditions of the infrastructure.

Multistage linear regression equations were developed in (29) to predict Future longitudinal level irregularity. A one-year longitudinal level data and passing tonnage data from the Beijing-Jiulong Railway Line were used to develop the model. The line was divided to sections of 200m length and the model was developed for the sections separately.

Standard deviation of the longitudinal level data was calculated for each of the inspection time in the year and the value obtained is categorized in to two stages by dividing the change curve of the irregularity of the longitudinal level into different stages. For each stage, linear regression equations were fitted to the standard deviation data and corresponding cumulative tonnage. Equation 2.7 was developed for stage one and Equation 2.7 was developed for stage two. The equations determined were used for predicting future longitudinal level data of any time with its corresponding cumulative tonnage.

$$\sigma_T = 2x10^{-8}xT_{0n} + 2.220 \quad (2.7)$$

$$\sigma_T = 1x10^{-7}xT_{1n} + 1.875 \quad (2.8)$$

Where:

σ_T: longitudinal level data of any time

T_{0n} and T_{1n} are the accumulative passing tonnage from the first inspection date in Stage one and two respectively.

(30)Developed Models used to quantify the effects of tamping and renewal on track geometry and ballast degradation of turnouts. Longitudinal level data from loaded geometry recordings were used in the study to determine track quality index and ballast quality index. The track was segmented in to sections of the equal length and Track quality index is determined as the standard deviation of the track longitudinal level for each of the sections.

The ballast quality index was determined from the fractal analysis of the track longitudinal level of medium wavelength of the segmented track sections. The track quality index and ballast quality index were determined for several consecutive periods and the effect of tamping was analyzed. The effect of tamping was represented by the amount of recovery after tamping in the track quality ($\Delta\ sd_{LL}$) and ballast quality ($\Delta\ Q_{bl}$). Two linear regression models called tamping recovery model for track quality (equation 2.9) and tamping recovery model for ballast quality (equation 2.10) were developed from the study.

$$\Delta\ sd_{LL} = 0.23 - 0.44\ sd_{LL}^{[b]} + 0.05v \quad (2.9)$$

$$\Delta\ Q_{bl} = 0.0008 - 0.36\ Q_{bl}^{[b]} + 0.0011v \quad (2.10)$$

Where:

$\Delta\ sd_{LL}$: amount of recovery in the track quality after tamping

$\Delta\ Q_{bl}$: amount of recovery in ballast quality after tamping

$sd_{LL}^{[b]}$: track quality before tamping

$Q_{bl}^{[b]}$: ballast quality before tamping

v : line speed

In the regression analysis, the computed R^2 was 46% in equation 2.9 and 38% in equation 2.9.

From the study, ballast renewal recovery model (equation 2.9) was also developed to study the effect of renewal on track quality and ballast quality recovery.

$$\Delta\ Q_{bl} = 0.0079 - 0.89\ Q_{bl}^{[b]} \quad (2.11)$$

$\Delta\ Q_{bl}$: amount of recovery in ballast quality after Renewal

$Q_{bl}^{[b]}$: ballast quality before Renewal

The results from the model showed that renewal restores 89% of degradation in the ballast layer and 67% of degradation in longitudinal profile. Tamping has minor effect on ballast quality restoration (36%), but it can restore 44% of longitudinal level degradation. From the results of the study It was concluded that when the condition of ballast are quantified by the fractal method, the chance of having low quality after renewal is near zero but it was different for the standard deviation of the longitudinal level. This difference in the estimated effect of renewal shows the applicability and usefulness of the fractal method for ballast condition evaluation (Galeazzi, 2021).

Tamping recovery model was developed for open track in (31). In the study, Tamping recovery of track was modeled by considering the standard deviation of the track longitudinal level. The methodology used was similar to (Galeazzi, 2021) but the tamping recovery was determined in two ways and the smaller value was adopted. First tamping was described as a linear relationship between the track quality recovery and track quality before tamping and secondly the recovery is dependent on the previous tamping operations. From the study, it was concluded that Track quality of the present tamping could never be better than the quality of the previous tamping.

(32) Studied the effects of maintenance on track geometry deterioration. Longitudinal level data and maintenance and renewal records were collected from UK rail and used in the study. The standard deviation for 200m track section was used as the representative of the track quality index. A linear function was used to model the underlying trends for the track geometry degradation. The track geometry degradation was described by the relationship between the track quality index and the time taken to achieve its state. The distribution of times to achieve any level of track geometry quality was determined by the line speed and the maintenance history to determine how track across the network degrades with respect to time.

From the study, it was found that the distributions of times to achieve a specified track quality following maintenance intervention were found to be highly dependent on line speed and maintenance history. For higher line speeds the track will require more maintenance to keep it at the acceptable standards than tracks with lower speed and the probability of achieving the required level of quality performance following tamping is less for the higher line speeds and this shows that higher performance levels become increasingly difficult to attain. From the study, it was concluded that tamping decreases the time it takes to deteriorate a track quality to the specified condition and it increases ballast deterioration rate.

(33) developed a time-series stochastic model and regression model to analyze and predict track degradation**.** An automated inspection data, asset data and operation data were used as an input. In this study the gauge parameter was used to represent the track quality index (dependent variable) and variables such as rail profile, rail type, track surface, curve radius, number of trips and annual rail usage (in million gross tons, MGTs) were used as an explanatory variables.

A correlation analysis and analysis of variance (ANOVA) were carried out for continuous variables and categorical variables, respectively by using SPSS software to evaluate their relationship with the changes in gauge value. After identifying the most influencing variables in the above way two degradation models were developed, a linear regression and a time series model, for curve and straight sections separately in order to predict the degradation of tram tracks. Linear regression and Time-series models were developed by using SPSS and MATLAB software respectively. In the study, 70% of the data were used for training and 30% for testing. From the correlation analysis and analysis of variance MGT was the common variable influencing the degradation of rails based on the changes in gauge value

The linear regression models developed by using SPSS software to the predict track degradation were equation 2.12 for curved sections and equation 2.13 for straight sections.

$$\text{Gauge}(t) = 1.075\ \text{Gauge}(t-1) + 1.462e\text{-}8\ \text{MGT}(t-1) \quad \text{(2. 12)}$$

$$\text{Gauge}(t) = 1.186\ \text{Gauge}(t-1) + 2.53e\text{-}8\ \text{MGT}(t-1) \quad \text{(2. 13)}$$

Where:

Gauge (t): gauge value at time t

Gauge (t-1): gauge value at time t-1

MGT (t-1): Tonnage at time t-1

The developed linear regression models estimated track degradation with R-squared values of 99.7% and 98.7% for curved and straight sections respectively while the time series models estimated track degradation with R-squared values of 89% and 97.9% for curved and straight sections respectively. A comparison of the time series and linear regression model in the study showed that both models can predict the gauge value with a very low error percentage and it was concluded that the application of linear regression is more suitable for the study as it was simpler and less complex than the application of the time series model.

Track geometrical degradation model, track geometrical improvement model and cost model of the mechanized tamping operations were developed in (34) to conduct optimized maintenance planning by using a stochastic geometrical degradation model. The track degradation model was developed based upon real data collected from the railway using regression methods and the other two models were developed according to combination of expert knowledge and related literature. To develop the track degradation model, the line was divided in to a section of 1 Km and each point of time has a series of data related to various track sections. To consider this variability, a distribution function was fitted to each point of time. Genetic algorithm was used in the study to carry out maintenance planning. The track geometrical degradation model was developed as Equation 2.14 and the track Geometrical Improvement Model was found to be equation 2.15.

$$TQI = 0.27T + TQI_0 \qquad (2.14)$$

Where:

TQI : track quality index;

T: time in month;

$TQI_0 = 6.51$

$$TQI_a = 0.768TQI_b \qquad (2.15)$$

Where:

TQI_a: track-quality index after improvement

TQI_b: track-quality index before improvement.

To obtain a cost model for tamping, the costs considered were the expenditure for renting tamping machines, the cost of providing one tone of ballast and the cost paid for machine operators and labor and cost to run tamping for a section to be tamped were calculated. Once maintenance operation assigned for a certain section, the associated track-quality index (after tamping) is modified using the improvement model and the amount of track-quality index for the next year is calculated using the track geometrical degradation model and by integrating those values with the cost model the optimal 6-year maintenance program was scheduled in the paper.

Ballast degradation model was developed for railway turnouts by considering ballast quality index as the output variable. Ballast quality index was estimated by fractal analysis of the irregularities of the track vertical profile and this ballast quality index was the basis for developing ballast degradation models in different sections of the turnout based on a segmentation scheme. Using track geometry data of 88 turnouts in the Danish railway network for the period 2012-2017, the study develops and compares ballast degradation models based on regression analysis and stochastic processes (lognormal and Gamma processes).

A multiple linear regression modeling was used to estimate the effects of different influencing variables on the ballast degradation rate. The variables considered in the study include the number of past tamping, maximum permissible train speed, passing million gross tons (MGT),weather condition (seasonality, temperature and precipitation), current degraded state of the track geometry, current degraded shape of the turnout, type of the turnout and section of the turnout . From the study, it was found that all the variables in the model were significant in 99.9% confidence level.

From the study it was found that increasing speed limit, tamping, being in the crossing section, MGT, initial degradation level, warm season, higher ratio of the front to the mid-section, have an increasing effect on ballast degradation rate in turnouts (23).

Similar study was done in (18) to develop a universal ballast degradation model for turnouts that can forecast the expected ballast degradation rate over a period of three months. Probability distribution and Bayesian inference was used in the study to develop a ballast degradation model that can predict ballast degradation rate from the influencing variables. From the study, it was concluded that BDR (ballast degradation rate) in turnouts is strongly affected by both geometry-related and weather-related factors.

The impact of undercutting of ballast on geometry roughness trend were evaluated and the effect of softer sub grades on the effectiveness of undercutting were investigated in (35)**.** Five years of track geometry data and two seasons of ballast undercutting records were used in the study. Sections in which undercutting operations were carried out were divided in to 60 m segments and standard deviation of the geometric parameters such as cross-level, alignment and surface were determined for each segment and for each track geometry survey. Furthermore, Sub grade conditions under the track were determined using the Agricultural Region of Alberta Soil Inventory Database (AGRASID) for each of the segments and the impact of sub grade condition on the effectiveness of undercutting were determined.

From the study, it was found that there exists variability in the spatial relationship between ballast conditions and roughness and the value of standard deviation of geometric parameters is consistently reduced after improving ballast conditions of some segments, while it gets unaffected or increased in other sections. It was also stated that Ballast renewal does not affect track roughness in the three geometric parameters in the same manner. Cross level and surface roughness trends agree with one another more than either of one does with alignment. It was also depicted that the proportion of segments constructed on mineral sub grades show a greater sustained reduction in roughness after ballast renewal than those constructed on organic sub grades.

The results of the study shown that 60% of the track segment exhibited a sustained decrease in roughness after ballast renewal which supports the idea that degraded ballast is a primary cause of increased track geometry roughness. However, the remaining 40% of segments suggest that other factors, beyond ballast conditions, were also influencing track roughness at a significant scale. From the study, it was concluded that for tracks constructed over soft organic sub grades undercutting of degraded ballast might not be sufficient to correct track geometry roughness. To manage track roughness efficiently soft sub grade related contributions to track geometry variability must be taken in to account by railway operators(35).

Degradation pattern of isolated longitudinal level defects was modeled by a defect-based model and a section based model in (36) to predict the probability of occurrence of isolated defects within the track sections. To model the track geometry degradation and predict track geometry defects, the longitudinal level measurements were used.

The evolution of the amplitude of the longitudinal level defects within a maintenance cycle was modeled using simple linear regression model. In addition to gradual degradation, an abrupt change in the degradation path (shock event) that can increase the amplitude of the defect over time was also considered in the degradation modeling. The effectiveness of tamping in correcting the longitudinal level defects was also analyzed and considered in the study. The section-based model was developed using binary logistic regression and the standard deviation of the longitudinal level, kurtosis of the longitudinal level, the time interval and the presence of defects in the latest measurement were selected as the explanatory variables in the model.

From the study it was found that isolated longitudinal level defects have a linear degradation pattern. Tamping intervention rectified defects that have exceeded upper bound of the intervention limit (UH2) and all the measurements are within the planning limits but after one year defects that have exceeded the lower bound of the intervention limit (UH1) occurred at the same position. Around 35% of the sections with a UH1 defect or UH2 defect after a tamping intervention the defects occurred again at the same position.

From the developed section based model it was found that the probability of the occurrence of UH2 defects in a track section is linked to the standard deviation and kurtosis of the track section and the developed binary logistic regression model shows a satisfactory performance in predicting the occurrence of UH2 defects.

(37) developed Ballast degradation model used for scheduling Ballast tamping activities (ballast maintenance operations). The standard deviation of the longitudinal level defect of 200m track segments were used to develop the model. First regression analysis was carried out between standard deviation of longitudinal level defect and cumulative tonnage for each of 200m segments to determine the degradation rate of the SD of the longitudinal-level defect (mm/MGT). Parallel to this the tamping efficiency was determined. Next the standard deviation of the of longitudinal level defect after the last tamping were calculated by the developed model and a tamping operation was triggered when the value of σ_{LL} reaches the tolerances defined by the infrastructure manager. In this way, the models can schedule the tamping operation.The model developed to forecast ballast maintenance operations was equation 2.16

$$\sigma_{LL} = C_{0,LL}\ \beta_{LL} + C_{1,LL}\ L \quad \text{(2. 16)}$$

Where:

σ_{LL} :standard deviation of the of longitudinal level defect

L : the accumulated axle tonnage/loads in million gross tones (MGT) for a 200m track segment since the last renewal or tamping

$C_{0,LL}$:the initial values of the SD of the longitudinal-level defects (in mm)

$C_{1,LL}$: the degradation rate of the SD of the longitudinal-level defects (in mm/MGT)

β_{LL} : the loss of efficiency of tamping for the SD of the longitudinal-level defect

(38) developed a predictive maintenance framework for predicting future changes in rail track geometry measurements. The development of the framework only focuses on linear variables and linear regression was divided in to two types to make the framework more robust and effective. A variable based linear regression and time based linear regression was used. The variables considered were dipped left (these longitudinal rails may be joined together by a fishplate and when the joint wears out it dips on inner sides), Cant Deficiency, Cross Level , Curvature ,gauge, twist, alignment and dipped right.

The geometry data was recorded for four periods. Each dependent variable was predicted by other independent variables for all the four time periods in variable based linear regression and each variable in last fourth was predicted from rest of the three previous data in the time based linear regression. From the variable based linear regression it was found that dipped left can be predicted with less confidence by the independent variables (R^2 = 0.13) and cant deficiency and cross-level were well predicted by the independent variables (R^2 = 0.86 and 0.73 respectively) and curvature was almost perfectly predicted by the independent variables(R^2 =0.97). From the time based linear regression cross- level, curve and cant deficiency were perfectly predicted by the dependent variables (R^2 =0.99, 0.94 and 0.83 respectively) and gauge was predicted with some confidence by the dependent variables (R^2 =0.41).

Regression models were developed for the north to south line of AALRT to predict the future longitudinal level from previous readings. Number of trips, curve radius, track surface and tonnage were the influencing variables (independent variables) and the longitudinal level was considered as the dependent variable to be determined. In the study it was found that, there was linear relationship between the independent variables (i.e., number of trips, tonnage, Speed, Grade) and dependent variable (longitudinal leveling) for the straight and curved sections of N-S line of AALRT(9). Equation 2.17 and equation 2.18 were the developed models for straight and curved sections of N-S line of AALRT.

$$LL(t) = -4.971 + 7.58e\text{-}4\ Tonnage(t-1) + 0.057\ Number\ of\ trips(t-1) + 0.012\ Speed + 0.131\ Grade \quad (2.17)$$

$$LL(t) = -0.543 + 7.25\ e\text{-}4\ Tonnage(t-1) + 0.078\ Number\ of\ trips(t-1) + 0.055\ Speed + 0.162\ Grade - 1.09\ e\text{-}4\ Radius \quad (2.18)$$

Where:

LL (t) : change in longitudinal leveling at a time t

(39) suggested a methodology for evaluating track degradation in terms of track geometry irregularities and developed a multivariate regression model to demonstrate the relationship between the track degradation measurement variable and other influencing variables. Since different sections of track are not identical, the track was split into sections with similar influencing variables and two exponential functions were applied to deal with the initial and second phases of degradation. From the study, it was concluded that axle load has a nonlinear relationship with degradation and degradation after tamping is dependent on the number of previous tamping. It was also concluded that soil consisting of clay material would settle sooner than other types of soil and light rail tracks degrade faster than heavy rail tracks. Furthermore, it was found that harsh rainfall increases degradation rate.

2.3.3. Mechanical-Empirical Models

Mechanical-empirical models are models developed to improve the performance of both mechanical and Statistical model by using the strong features of each. By coupling the mechanical properties of the system and the actual measurement and operation features, mechanical-empirical models produce better degradation model. These models are applied to overcome the shortcomings of mechanical and statistical models and improve degradation predictions.

Mechanistic- Empirical approach was used in (40) to develop A railway track condition index by considering Passenger Ride Comfort (PRC) and maintenance approach. The track condition index was developed by performing field investigation and statistical analyses of the data obtained from the field investigation. The field measurements were made in two phases.

The geometry data were recorded using a track-recording car in the first phase and train accelerations at passenger positions were measured in the second phase. Next to this, geometry parameter indexes (SD of track geometry data) were obtained from the first phase and ride comfort index was obtained from the second phase data. After that by investigating the correlations between the track geometry parameter indexes and PRC, the new index was established. Finally, by using both the new index and the conventional index, an algorithm was developed for prioritizing and scheduling maintenance activities and the advantages of the new algorithm over the conventional algorithm (current maintenance approach) was demonstrated by making comparisons between the results of the new algorithm and those of the current approach.

The effectiveness of the new index was verified by applying the conventional index and the new index to a railway line. It was found that the required maintenance actions obtained from the existing approach were considerably different from those of the new approach and on verification; it was shown that the new index effectively reflects the PRC whereas the conventionally used index does not necessarily reflect it. It was also shown that the new algorithm not only guarantees track safety but also ensures the required level of PRC.

(25) used Mechanical Empirical models to develop a track degradation model which incorporate the track structural condition. The Mechanical model was used to investigate the mechanical properties of track structure and railroad vehicle and their contribution on track degradation. While the Empirical model was applied to relate the effective parameters and degradation ratio which were used to predict the future track quality. Iran railway line was used as case study and inspection data collected for a

period of 2 years of the line was used in the analysis.

Track quality, traffic and maintenance history were the main parameters taken in to account in the study as factors affecting rate of track deterioration. Two track quality indices were considered, which are namely Track Geometry Index (TGI), which is condition of track geometry parameters such as gauge, profile, twist and alignment, and Track Structure Index (TSI), which involves condition of track structures such as rail, sleepers, ballast, fastening system, sub grade and drainage system. Most effective traffic parameters considered are Equivalent million gross tons (EMGT) and average running speed (V).

Two models were developed in the study, the first was based on the relationship between the track geometry conditions and time and the second was based on structural visual inspections of the mechanical conditions of the track components over time. The two models developed were equation 2.19 and equation 2.20 for the first and second model respectively.

$$\frac{TGI2}{TGI1} = \alpha_4 \exp(\beta_1 V + \beta_2 EMGT + \beta_3 TGI_1) \times (\lambda_1 T^4 + \lambda_2 T^3 + \lambda_3 T^2 + +\lambda_4 T + 1) \ldots \quad (2.\ 19)$$

$$\frac{TSI2}{TSI1} = \alpha_4 \exp(\beta_1 V + \beta_2 EMGT + \beta_3 TSI_1) \times (\lambda_1 T^4 + \lambda_2 T^3 + \lambda_3 T^2 + +\lambda_4 T + 1) \ldots \quad (2.\ 20)$$

Where:

TGI_2: future track geometry index,

TGI_1: present track geometry index,

TSI_2: future track structure index,

TSI_1: present structure index,

T: the time (in seconds).

Furthermore, a model was formulated from the study providing the correlation between TGI2 and TSI2 to limit the study to fewer inspections as defined in equation 2.22.

$$TSI_2 = \eta_1 \eta_2 \eta_3 \eta_4\ TGI_2 \ldots \quad (2.\ 21)$$

Where:

$\eta_1 = k_1 V + k_2$

$\eta_2 = k_3 EMGT + k_4$

$\eta_3 = k_5 TGI_1 + k_6$

$\eta_4 = k_7 T + k_8$

K_1 to k_8 are constant coefficients for line characteristics such as Straight line, curved line, tunnel, turnout and bridge

2.3.4. Artificial Intelligence Models

Artificial Intelligence models are type of machine learning models used in the prediction of the rail track degradation. AI (artificial intelligence) models involve activities and developments relating to human-like intelligence reproduced by computer applications. For this purpose, they exploit computer techniques or reasoning algorithms that attempt to automate intelligent functions (28).

(41) Applied ANN (Artificial Neural Network) to evaluate railway track quality condition of Iran railway network. Track geometry data including gauge, profile, alignment and twist were obtained from track-recording car and an artificial intelligence model were developed for the establishment of correlations between the track structural conditions and the data obtained. Multilayered feed forward networks with one hidden layer were utilized in the study. The back propagation-learning algorithm using the real-valued sample data trained the networks and the mean squared error technique was used to assess the network performance in the training and validation procedures.
(42) developed ANN Model that can predict track degradation rate. Tack geometry measurements, asset information, and maintenance history for five line sections from the Swedish railway network were collected and standard deviation of the longitudinal level of different track sections were used to characterize the quality of the track geometry.

Linear model of track quality with time as explanatory variable was used to model the temporal degradation of the track geometry within a tamping cycle and furthermore to consider the section-to-section variation of each track section, the degradation rate of each track section was determined using ANN and based on a set of features related to each track section.

In order to develop an ANN model, which can accurately predict the track geometry degradation the following features of the track sections, were considered. Ballast age , maintenance history (the number of tamping interventions), rail type , sleeper type , sleeper age, existence of a drum , existence of level crossing , existence of bridge, track speed class ,average annual frequency of the trains passing along the track section

,degradation level after tamping or renewal , curvature and ballast type feature. Garson method was applied to explore the relative importance of the variables affecting geometry degradation rate.

In developing the model the dataset was split into a training and a testing set and the data division were random. The performance of the ANN model was assessed using the coefficient of determination (R^2) and the mean squared error (MSE) by assigning training sample sizes ranging from 20% to 90% of the whole dataset. From the explanatory variables affecting the degradation rate, maintenance history was found to be the most significant one and the degradation level after tamping was found to have the second highest relative importance. The frequency of the trains passing along the track is ranked as the third most important factor and other variables such as track speed class, ballast age, sleeper age, sleeper type, and the existence of a level crossing made a roughly an equal contribution to the prediction of the track geometry degradation rate and ranked as last contributors. From the predicted versus the measured degradation rates, the calculated R^2 value was found to be 0.850, indicating that the developed ANN model has an acceptable performance. From the study, it was concluded that ANN could properly be used to explain the section-to-section variation in degradation rates in the studied line sections.

(43) developed an artificial intelligence model to predict conditions of track in the future. The data set used in the study was collected from Yarra Trams and gauge parameter was used to represent the track quality index. MATLAB software was utilized in matching, extracting and conducting statistical analysis to determine the most influential variables that affect the tack quality index. Following this an Adaptive Network-based Fuzzy Inference System (ANFIS) and Artificial Neural Network (ANN) models were developed for straight and curve sections.

In the study, the total data set was randomly divided into two separate sets, which were used as the training and testing sets**.** From the statistical analysis, it was found that the variables such as MGT and history of gauge values were the most common variables which impacts gauge degradation and MGT was used as an independent variable in developing the model. The ANFIS model developed from the study was capable of predicting future gauge values with an R-squared value of 0.60 and 0.78 for curved and that of the straight section respectively. However, the R-squared value for ANN model reduced to 0.4587 and 0.5813 for curved and that of the straight section respectively. From the study, it was concluded that developing a model, which is capable of predicting the degradation of light rail tracks is necessary due to lack of sufficient studies done on the degradation of LRT.

ANN and regression models were developed by (28) to predict Track degradation for tram system by considering gauge deviation as a representative of track geometry irregularities. The models were developed by considering track gauge deviation in previous year ,curve radii, MGT, rail support ,track surface ,rail profile and rail type as input variables and track gauge in current year as an output variable. To determine which variables are significant in prediction of the output variable Pearson Correlation analysis has been done for continuous parameters and one-way ANOVA analysis has been applied for categorical variables. It was found that track gauge deviation in previous year, track surface and rail type are significant in prediction of the output variable.

According to the results of the degradation model for the repaired straight segments, previous gauge deviation has a clear correlation with current gauge deviation and the rate of degradation for those surfaced with asphalt were lower compared to those surfaced with concrete. However, for repaired curve segments, those segments surfaced with asphalt have higher degradation rate than those covered by concrete, which is similar for unrepaired straight segments. For the unrepaired curve segments previous gauge deviation have similar effects with other segments but Rail type is identified to have impact on track degradation in which Grooved rails have greater resistance to degradation than T-shapes rails. The developed models are capable of predicting the gauge values with R^2 from 0.71 to 0.91. From the study, it was concluded that the performances of both regression and ANN models in predication of rail degradation are approximately similar and acceptable.

(44) developed ANFIS model for estimating rail track degradation for the curved and straight sections of Melbourne tram track system**.** Gauge value considered as a paramount factor for track degradation and it was used as a parameter to predict the future track degradation. Data including gauge value, curve radius, annual tonnage in MGT, track surface (asphalt and concrete surfaces), rail profile , rail type (grooved and T-shapes), rail support (concrete and steel sleepers), location of routes, and track

installation date have been collected for 6 years from the operator of the Melbourne tram network and used to develop the model.

The dataset is divided into training and testing set. From the total data set, 70% of the data was used for training purpose while the rest is used for testing. The ANFIS model consists of three important inputs including the two previous values of gauge (the previous two readings of the gauge from the past two years) and the MGT value to estimate the gauge value for the current year. Results show that the model can predict the gauge values for the coming year with R^2 value of 0.60 and 0.78 for the curves and straight sections respectively.

From the developed models, it was found that the deviation in the values of gauge value is higher in the straight section than the curved ones. The key reason for this was considered that straight section covers most frequent stop-starts and it was also assumed that switches and bridges which are more in danger of degradation are mostly quantified as straight sections in the study. It was indicated that there exist a higher probability of degradation on the straight sections over curves and this section type needs more inspections.

ANN model were developed in (45) for track geometry deterioration prediction by using a comprehensive field investigation data gathered over a period of 2 years on approximately 180 km of Turkish state railways**.** The line was divided into analytical segments (AS) of uniform characteristics and this enables to obtain AS of uniform properties. The segmentation considers the variables including gradient (‰), curvature (1/R) (1/m), cross level (mm), speed (km/h), age (years), rail type (kg/m), rail length (m) and sleeper type. From the segmentation operation 820 uniform AS were obtained with the average length of 220 m. For each AS track structure data, traffic characteristics data, track layout data, environmental factors, track geometry data, and maintenance and renewal data were collected.

The average deterioration rate of the geometrical parameters was determined for each AS and ANN model was developed to model the effect of the influencing variables on the deterioration rate. From the goodness of fit between the predicted and observed values the coefficients of determinations were found to be 0.727, 0.795, 0.765, 0.831 and 0.742, respectively for twist, gauge, alignment, cross level and leveling and this result shows ANN models produced significant relationships between the deterioration rate and the influencing variables. From the study, it was concluded that ANN with its ability to discover input – output relationships is one of the best ways of modeling track geometry deterioration associated with a high inherent complexity (45).

Different degradation models are discussed in the above literatures and Table 2-1 describes the strength and weakness of the above-discussed models. From the above-discussed approaches of models, the statistical modeling approach is used for this research. This approach is selected due to the availability of historically recorded data for the study area and the capability of the approach to address uncertainty of degradation more efficiently than the mechanistic approach. Lack of appropriate laboratory tests is also another barrier to adopt the mechanistic approach.

Table 2- 1: Comparison of degradation models

Type of model	advantage	disadvantage
Physical models	1.Based on Mechanical properties of track component 2. Uses primary data source. (laboratory experiments) 3.Small amount of data can be used in model development	1. Model is unable to deal with heterogeneity of track structure materials. 2. The model is applicable in limited number of track sections rather than the whole section. 3. Time consuming and Intensive
Statistical Model	1. Ability to work with large set of data and provide realistic results 2.model development can be easily understood 3. Can solve complex and sophisticated problems better than mechanistic models.	Requires large data set to develop an effective model
Mechanical-Empirical Models	1. Uses both laboratory experiments and historical data 2.Applicable in different track segments	Show higher rate of degradation on turnouts and bridges compared to other parts

Type of model	advantage	disadvantage
Artificial Intelligent Models	1. Ability to solve complex problems through learning data pattern 2. Can generate estimations with higher accuracy	1.Lack of transparency during decision making process 2.Without enough data training of the model is not well done

2.4. Fractal Analysis of Railway Track Geometry Data

Fractal analysis is the process used to determine the fractal dimension (fractional dimensions that reside between conventional whole-number dimensions) which was first developed by Mandelbrot to characterize those patterns within nature that cannot be effectively quantified using classical geometry of whole-number dimensions. Such kinds of patterns are generally described as irregular, chaotic, or fragmented (46). According to (46)**,** fractal analysis has been applied to railway track geometry data to examine its use in indicating track geometry condition, planning maintenance, and evaluating the cause of substructure-related problems. In the study, fractal analysis of geometry data was used for track geometry condition monitoring .Through experiments with geometry data; it was showed that the fractal analysis method is an effective way for trend analysis and track geometry degradation assessment.

(47) Reported the effectiveness and usefulness of fractal analysis for ballast condition assessment. When applying fractal analysis to railway track geometry data, the roughness of the geometry signal for different wavelength regions is considered. It was showed that three orders of roughness could be identified for wavelengths below 3 m, between 3 and 30 m, and above 30 m. Based on real-world validation with in-situ conditions, the second dimension (3 to 30m) is an indication of the ballast quality (47). In addition to the above (18) and (30) also support that fractal analysis of track longitudinal measurements is a qualified method to develop a quality index of ballast condition.

There are various techniques for determining the fractal dimension of rough patterns; from these methods, the two widely practiced methods found on most of the literatures are the divider and box counting methods.

The divider method

The divider method is implemented in such a way to "walk" the divider along the line and record the number of steps required to cover the line. By systematically increasing the width of the divider and repeating the stepping process, the relation between step size and line length over a range of resolutions is determined (48). It was used by Mandelbrot to quantify curves with fractal dimensions >1.0. This method expresses the length of a rough line by equation 3.1

$L(\lambda) = n\lambda^{1-D_R}$ .... (3. 1)

Where;

λ : length of unit measurement,

$L(\lambda)$: length of the rough line based on unit measurement length λ,

n : number of steps of length λ, and

D_R : fractal dimension of the rough line.

Taking logarithmic function on both sides of equation the expression become

$\operatorname{Log} L(\lambda) = (1 - D_R) \log\lambda + \log n$

This is in the form of the equation of a line, (i.e. $y = mx + b$) and Therefore, a plot of λ versus L (λ) on a log–log scale yields a linear relationship with the slope of the line defined as $m = 1 - D_R$ and the fractal dimension (D_R) equals $1 - m$.

The fractal dimension is therefore determined from the rate of change of measured length to corresponding step length as expressed in log-arithmetic form [slope of log L (λ) versus logλ plot].Most authors such as (49) and (50) have found the method to be reliable when tested on curves of known fractal dimension. But (51) concluded that the divider method was flawed and produced unreliable results.

Box counting Method

The box counting method is widely used to determine the fractal dimension of many different phenomena. Since it can be applied with equal effectiveness to point sets, linear features, areas, and volumes, the box counting method is a widely used means of determining fractal dimensions (48).The basic implementation in the determination of fractal dimension for linear features is as follows:

Different square grids will be used to cover the line. When the width of square grid ε changes, the number of grid N (ε) will also change. The relationship between N (ε) and ε will be as defined in equation 3.2

$N(\varepsilon) = k\,\varepsilon^{-D}$ …………………………………………………… (3. 2)

When the width of square mesh takes ε1, ε2, ε3… εk, the number of grid correspondingly will take N(ε1), N(ε2), N(ε3), …, N(εk). Using linear regression method to analyze the data (εi, N (εi)) on log-log scale, getting the regression coefficient α (slope of the line). Then the fractal dimension (D) can be calculated by the relationship D = -α. The method requires a large number of data points in order to produce correct dimension (52). The box counting method requires a significant amount of computer memory and computational time since a very large number of cells have to be stored.

Most of the literatures sowed that the box counting method gives reliable results than the divider method. (53) Showed that Box Counting Method has much higher accuracy than dividers Method and the error from the actual fractal dimension in divider method becomes 27.79% but the error can be dropped to 7.86% by box counting method.

Summary

In this chapter, previous works on degradation models and fractal analysis have been reviewed. A number of mechanical, statistical, mechanistic- empirical and artificial intelligence models have been developed and different researchers have utililized different methods in their work. The literature review presented here discussed the works most suitable for integration into the present study. However, most of the researches done were focused on the track geometry degradation models as a general. Developing degradation models for track components individually is useful to investigate the root cause of track degradation and enhances the track maintenance system, for this purpose component focused degradation models shall be developed. The research presented in this thesis aims to fill this gap by developing degradation model for the ballast component. By comparing and contrasting the advantage and dis advantage of fractal analysis methods, box counting method is applied to determine fractal dimension of the track longitudinal level. From the modeling approaches, Regression model is applied to model ballast degradation in AALRT.

CHAPTER THREE

3. RESEARCH METHODOLOGY

3.1. Introduction

The current practices in analyzing and modeling of track and ballast degradation are discussed and Reviewed in the literature review section. After critically evaluating literatures and understanding the necessity of ballast degradation models, the methodology described in the following sections is adopted to solve the research problem. The study approach used to create the models to forecast ballast degradation in light rail transit is presented in this chapter. It makes use of measurements of track geometry, operational statistics data, and asset management data gathered over the previous years. This chapter present overview of the location of the study area, the study design, the data collected and the procedure followed in the determination of ballast quality index by fractal analysis and the development of regression model.

3.2. Description of the Railway Route Selected for the Study

The Railway route selected for the study is found in Addis Ababa. The double-track, 34.25 km long metropolitan electric railway in Addis Ababa has two lines: an east-west line that is 17.35 km long and a north-south line that is 16.9 km long. Megenagna, Legehar, and Mexico Square are all on the East-West route, which runs from Ayat Village to TorHailoch and the North-South line runs through Menelik II Square, Merkato, Lideta, Legehar, Meskel Square, Gotera, and Kaliti. Both routes follow the same path from Meskel Square to Lideta until the East-West line route splits in to the west toward the Torhayloch through coca cola station. The alignment of AALRT and some AALRT stations are shown in Figure 3.1. The maximum slope of the track is typically 5%, and the nominal track gauge is 1435 mm. 60,000 passengers may be moved along the two lines in an hour and 80 km/h is the highest service speed on the line. Currently, the LRT serves 39 stations and employs roughly 40 LRVs (Light Rail Vehicles), each of which can accommodate 286 passengers. Ballast is not present on the elevated sections of the railway (slab track) and ballast track is found in the non-elevated sections of the track. The ballast track is made with precast mono block reinforced sleepers. This research covers the ballasted sections of the AALRT.

Figure 3-1: Map of AALRT

3.3. Study Design

The purpose of this study is to develop ballast degradation model for the AALRT. To achieve this goal, the following methodological procedures are followed. To begin with, degradation models used for degradation prediction found in various publications, reference materials, articles, international journals, and websites was reviewed. After that, the data set including inspection data, asset management data and operation data of the line is collected from AALRT office and used as input for this research. Next to this, fractal analysis is applied on the track longitudinal level data to calculate the fractal dimension for different sections of the line and these fractal dimensions represent the ballast quality index for the respective sections. After determining the ballast quality index a multiple linear regression model was developed by using SSPSS software to estimate the effects of different contributing variables on the Ballast quality degradation. Ballast degradation models are developed for straight and curved sections separately to consider the effect of curve radius on ballast quality. The outputs of this research are ballast degradation models for the straight and curved sections and finally the research is concluded, and the necessary recommendation and further research studies are specified. The research framework is presented in Figure 3.2

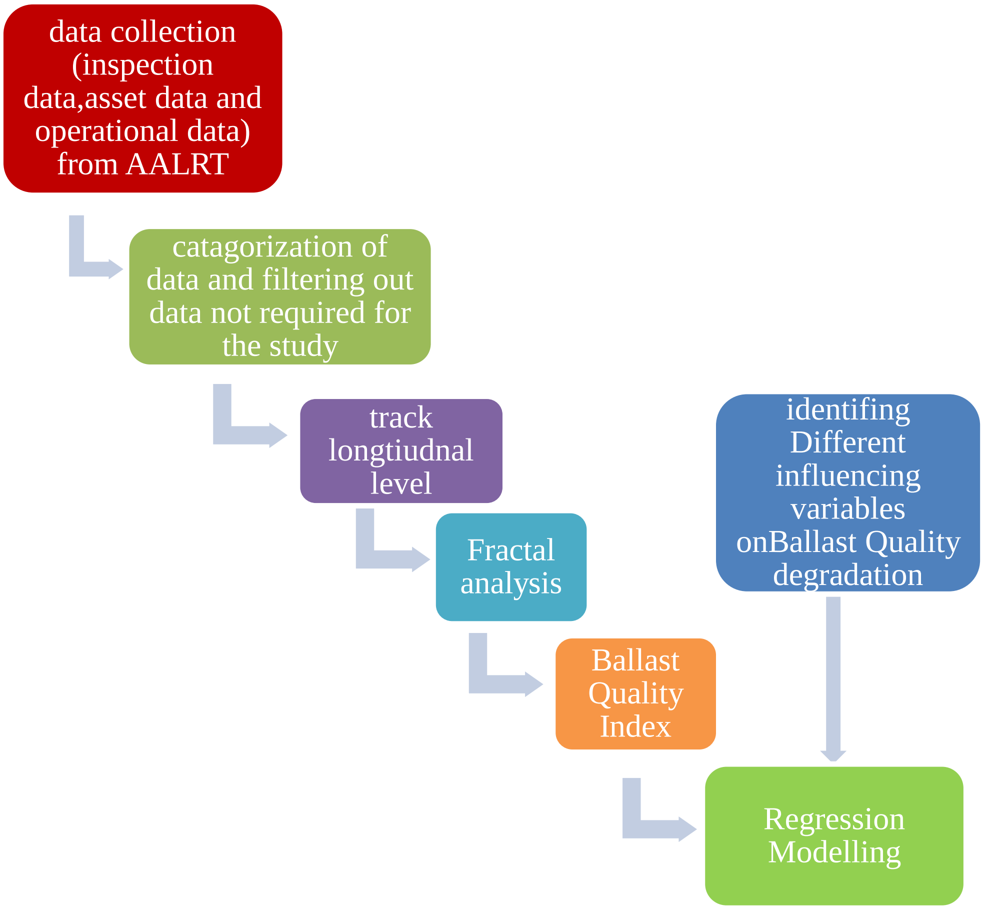


Figure 3-2: Research process framework

3.3.1. Dataset

Frequent inspections and maintenance works are conducted by the railway organization to guarantee the system's safe operation, and as a result, massive volumes of data are gathered in this way. The research's data set was gathered from the AALRT office as a secondary data and primary data were also gathered from the study route. Every day, AALRT inspects the track's portions on foot or by driving across the track at a speed that makes it possible to identify noncompliance with the standards. Qualified LRT staff carry out inspections. Visual inspection is utilized to find horizontal alignment errors, while portable equipment are used to capture geometric defects such as gauge, longitudinal level, cross level, and twist as presented in Figure 3-3. A section of track's components is inspected, and the condition of each is documented on an inspection form along with any deviations or defects and corrective action is undertaken for defects in accordance with the guidelines outlined in the Chinese standard. Copy of longitudinal level inspection documents showing the conditions of the track were acquired and are used as input data for this study.

Figure 3-3: Track inspection

Inspection Data

Addis Ababa Light Rail Transit inspects the track and the track's condition is recorded in the inspection form on a daily basis. This research used 11 months of inspection data from the years 2019 to 2020. The degradation modeling of the ballast utilizes this inspection data to evaluate the ballast degradation pattern of the line. The inspection data gathered includes measurements of the following track geometry parameters:

i. **Track gauge:** is the distance between rail tracks, measured at a right angle between the inner faces of the load-bearing rails. The typical rail gauge measures 1435 mm. The gauge imperfection for the line is measured using a gauge ruler.

ii. **Alignment:** This is the shift in the rail's curvature over a specific cord length. Visual inspections are done to measure the line's horizontal alignment.

iii. **Twist:** this is the difference in cross-levels of curve rail tracks over a certain cord length and it is expressed in mm/m. This is the height differential between two rail's top surfaces on a railroad track. When there is no height difference between the rail surfaces at the measurement point, the cross level is zero. When the outer rail of a curved track is shorter than the inner rail, the cross -level is negative (also known as reverse cross-level), while it is positive when it is not.

iv. **Longitudinal level (Profile):** is a projection of the track geometry along the longitudinal vertical plane. The longitudinal level (Profile defect) is the vertical deviation from this centerline. According to (46) and (18) it is usually the medium-wavelength of the longitudinal level measurement that shows the ballast condition . Hence, it is considered as the input parameter for this study.

On varied portions of the line, inspections are done every day, and each segment gets a second round inspection every month. The inspections are carried out along a line, 5m apart from one location to the next. The repair team is notified of the points that have defects that are above what is acceptable, and such maintenance is carried out at night. Table 3-1 shows the tolerance limit used by AALRT for maintenance of the track.

Table 3- 1: Tolerance limit values of geometric dimension of track

Items		Acceptance of operation (mm)		Regular Maintenance (mm)		Urgent repair (mm)	
		Main track and auxiliary track	Yard truck	Main track and auxiliary track	Yard truck	Main track and auxiliary track	Yard truck
Gauge		+6 -2	+6 -2	+7 -4	+9 -4	+9 -4	+10 -4
Level		4	5	6	8	10	11
High-Low		4	5	6	8	10	11
Direction		4	5	6	8	10	11
Twist of track (twisted)	Easement curve	4	5	5	7	7	8
	Straight line and circular curve	4	5	6	8	9	10

Asset Data

In this subsection, the geometrical characteristics of the line, including the track category (i.e. track sub grade condition, curve radius and Grade) and the construction elements, such as whether the track segment is straight or curved, were obtained.

Operational Data

Operation data are statistics about the use of the track. Both static and dynamic resources are used in Track Operations. Static resources include rail infrastructures such as tracks, bridges, stations, crossings, tunnels, etc., whereas dynamic resources include moving assets like as people and wagons. The operational data that are being used in this study include tonnage, number of trips made, speed, number of past tamping and weather conditions. Tonnage was calculated after obtaining information such as passenger flow for each month considered, weight of trains and average weight of passenger. Gross tons (MGT) are the result of total weight, which includes the weight of locomotives as well as the weight of the typical daily volume of traffic that passes the track. Number of trips represents the quantity of trains that cross each section of rail track. The speed used in passing through track section is used in this study to evaluate its contribution in ballast degradation. For this purpose, the author records the speed used in traveling through each section .To determine the average operation speed for each section; the speed for three trips with different drivers was recorded and averaged.

Filtering and Data Categorization

Due to the use of secondary data in this study, careful data analysis is necessary. This was done to ensure that the data was appropriate based on its dependability, sufficiency, and compliance with the problem's requirements. From the secondary data available in the inspection forms (i.e. longitudinal level, gauge, cross-level, and twist) the sole factor taken into consideration for this study is longitudinal level defect records. Thus, only the longitudinal level recorded on the inspection forms was used to sort the data. Data from copies of inspection forms was then entered into an excel file for simplicity of use. The descriptive statics of the longitudinal profile defect used for determining ballast quality index of sections is presented in Table 3-2

Table 3- 2: Summary of Data Used in fractal analysis (longitudinal level data)

Statistical Parameter	value
Minimum Value	-14
Maximum Value	11
Mean	0.373
Standard deviation	2.13
Total Sample data	16590

3.4. Fractal Analysis

As discussed in the literature review section, fractal analysis is a best method to characterize ballast quality from the recorded longitudinal profile defect data. From the two widely practiced methods determining fractal dimensions, the box counting method is employed in this study to determine the fractal dimension of different sections to characterize the ballast quality. For doing this, the available fractal-dimensioning software called Image J software with Fraclac as a plugin is employed in the study.

To do the fractal analysis the track is divided into several segments. The segments are divided by considering similarity of the influencing factors with in a segment. Based on this the ballasted section of the line is divided into 258 sections of length 80m. From the fractal analysis methods discussed under chapter two, box counting method is applied in this study to determine fractal dimension of each sections from the track longitudinal level. It is selected by considering its reliability and availability of software. The software used in calculating the fractal dimensions of track segments is FracLac plug in for ImageJ software. FracLac scans images using a shifting grid algorithm that can do multiple scans from different locations on each image. FracLac also scans images using an optional rotation function that analyzes an image from several angles and summarizes the results and it automatically generates different types of series of box sizes. FracLac works on binary images (black pixels on a white background, or white pixels on a black background) and on grayscale images to determine fractal dimensions by box counting algorithm. The program allows choosing the minimum and maximum size of boxes and this helps the user to determine the fractal dimensions with in different wavelength regions. The steps carried out to determine the fractal dimension of each sections is described in Appendix A and sample data used in fractal analysis is given in Appendix B

3.5. Selection of Influencing factors

By reviewing previous literatures and by considering the available data factors influencing ballast quality are identified. The author considers previously considered factors and others that are necessarily to be considered to develop a comprehensive ballast degradation model. The influencing factors on changes in the Ballast Quality Index (BQI) identified and taken into account in this study are tonnage, number of trips, speed, gradient, curve radius, number of past tamping, track subgrade condition, and weather condition (season).

3.6. Data Analysis Using SPSS

Using SPSS software, the influencing factors on changes in the Ballast Quality Index (BQI) are evaluated. The following variables were taken into account in this study: tonnage, number of trips, speed, gradient, curve radius, number of past tamping, track subgrade condition, and weather condition (season). The gathered information and the findings of the fractal analysis (BQI) are used to conduct analysis for just the straight and curved parts of the line. The analysis reveals the factors that have the greatest impact on ballast deterioration, which are then utilized to estimate the model of ballast degradation.

For continuous variables (such as the number of trips, tonnage, speed, grade, and curve radius), correlation analysis is utilized, whereas analysis of variance (ANOVA) is used for categorical variables (track sub grade condition and season).Correlation values, which range from -1 to +1, are used to assess these analyses. Values closer to 1 indicate a close association between the variables, whereas values 0 imply no relationship. The statistical significance test determines if the obtained correlation values are significant (p value should be less than 0.05). The variables that exhibit stronger correlation and have significant values are regarded as influencing factors and used in developing the ballast degradation model.

3.7. Regression Model

In this study, a regression model, a simplified predictive model known as the ballast degradation model is developed to estimate the future state of the ballast condition. The created model will help assess the ballast's condition, and these findings can be used to develop an ongoing maintenance program.

The most important factor that affects the change in the ballast quality index is discovered and with the use of the ballast quality index and the most relevant factors identified in Chapter 4, a regression model is developed using SPSS software. Results of the model, including estimated variables and model coefficients are produced using a sample of training data set and last, model performance is evaluated by plotting the observed values against the predicted values for both the straight and curved sections of the line and this allows one to evaluate the goodness of fit. This section discusses the procedures that were used to develop the regression model.

Multiple Regression Model development

A multiple regression model is used to assess ballast degradation of AALRT. A set of statistical techniques called regression analysis is used to calculate the relationships between variables. Regression analysis comprises in dependent variables, which are components required to arrive at the outcome and the dependent variable, which is the expected result at the end of the analysis. Regression analysis is useful for understanding the relationship between independent and dependent variables, estimating the relative strength of multiple independent factors' effects on a dependent variable, and forecasting future patterns of a series or activity. The format of the simplest regression model is defined in equation 3.1.

$$y = \alpha + \beta_1 x_1 + e \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots.. \quad (3.1)$$

Where;

y :the dependent variable which is variable to be estimated,

α: a constant or interception of regression model,

x_1: the dependent variable,

β_1 : the (regression) coefficient of the independent variable x.

: error or residual of the equation.

Multiple regression analysis is a set of technique for studying the straight-line relationship among two or more variables. The equation 3.2 defines a form of multiple regression model constructed through multiple regression analysis.

$$y_j = \alpha + \beta_1 x_{1j} + \beta_2 x_{2j} + \ldots + \beta_p x_{pj} + e_j \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (3.2)$$

Where:

X_s : independent variables,

y :the dependent variable,

j :represents the observation row number,

β_s : the unknown regression coefficients,

e_j: error (residual) of observation j.

In the present research, a multiple regression model is developed to predict trends and future estimates of the degradation of ballast quality.

Regression Model Build Up

In this study, a ballast degradation model is formulated using multiple regression analysis for both the straight and curved sections of the AALRT. The multiple regression model required a number of stages to be developed. These consist of collecting the information required for regression analysis, specifying and estimating the regression model, establishing and evaluating the assumptions of the regression model, interpreting the results, validating the regression model, and putting the model into practice. The steps used in developing the model are illustrated in Figure 3.1.

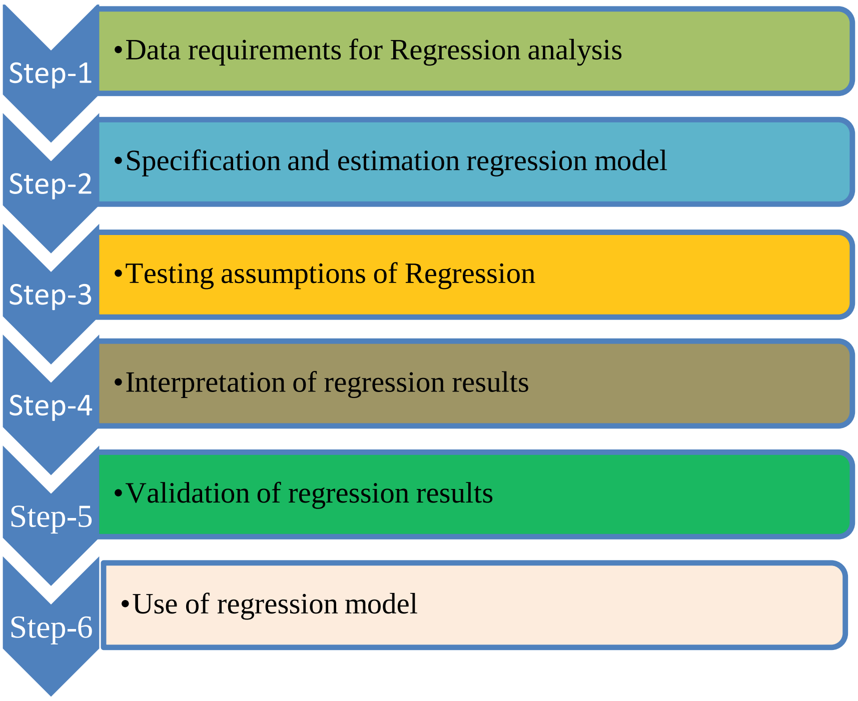


Figure 3- 4: Steps to conduct Regression Analysis

i. Data Requirement for Regression Analysis

Data requirement for Regression analysis include features such as sample size, variable variation, scale type, and collinearity of independent variables, among other things. A large enough sample size is required for regression analysis to be carried out. To achieve reliable results from regression analysis, the sample size that is allowed is essential. Green developed a technique called the "thumb rule" that is utilized to evaluate the effect of every individual Parameter. Green suggests a sample size of

140+k, where k is the number of independent variables (9). In this study, the regression modeling of the curved and straight sections of the AALRT used 1242 samples of data. Furthermore, 70% of the data is used for model training and 30% of the data is utilized for model testing. Training data is used in model development and testing data is used for model validation. Some of the Sample data used in regression modelling is given in Appendix C.

Variable variation is another feature of data requirement to do regression analysis. Regression analysis can only be carried out when the independent variables vary. Variations in the independent variable are one of the primary causes of variations in the dependent variable. Since the independent variables included in this study—tonnage, number of trips, curve radius, gradient, speed, number of past tamping, track subgrade condition, and season—have shown variation in their values, regression analysis can be used to forecast changes in the ballast quality index (BOQ). Another type of data requirement is scale type of variables and the scale types of variables, which favor regression analysis, are ratio and interval. Interval scale types of variables are used in this study, which supports regression analysis.

Another set of requirements for data is the diagnosis of collinearity. Collinearity is defined as the presence of high correlation between two independent variables and multi-co linearity is defined as the presence of more than two strongly correlated variables. The fundamental problem with co linearity across variables is that it often misrepresents important factors as unimportant; the VIF (Variance Inflation Factor), a reciprocal value of tolerance, is used to determine the presence of collinearity and multi-co linearity, when the VIF is less than 0.1 multi-collinearity exists and collinearity exists when the VIF is greater than 10. There are two techniques that can be used to eliminate co linearity in a data sample: factor analysis and re-specifying the regression model (54).

ii. Specification and Estimation of Regression Model

The process of specifying a regression model involves selecting the independent and dependent variables. The selection of independent and dependent variables is based on the objectives of the study as well as the research question that needs to be answered. The primary goal of this study is to predict the pattern of ballast degradation, and the dependent variable used in this study is the ballast quality index. Tonnage (MGT), number of trips, speed, gradient, curve radius, number of past tamping, subgrade condition, and season are the variables that have been selected as independent variables. These were the significant variables that were selected following the completion of the correlation analysis in chapter 4. The correlation analysis's findings indicate that these factors have the greatest impact on how the Ballast Quality Index changes along the line.

Based on the SPSS software used in this study, Model estimation involves choosing an estimation approach as well as an analysis method. Two methods for choosing an analytical method are Enter and Stepwise. The two subtypes of the stepwise method are backward and forward. Stepwise analysis allows the process to choose the best subset of variables, While enter approach enables the user or researcher to select the independent variables to be utilized in the analysis.

The techniques of step modeling use Forward and Backward types. Using the forward modeling approach, huge regression models are developed by starting with a constant next, using just one more independent variable from the other variables, it seeks to identify the optimal model. It then compares the outcomes of these two models. It continues by adding a second variable from the remaining variables if including an independent variable results in a model that is noticeably better. The resulting model is then contrasted with the prior model, which now contains the two independent variables with which includes one independent variable. This cycle is repeated until including a new variable has no noticeable impact on the model. Similar steps are taken by the backward technique, which first inserts all potential variables before deleting the one that contributes the least to the model's overall performance. The drawback of the stepwise method is that it only adds factors that are significant by coincidence, not just ones that happen to be actually interesting and valuable. Enter method is employed for this particular study since it allows the user to select the independent variables.

The choice of an estimation method entails the choice of estimation methodologies. There are numerous techniques, including weighted least squares (WLS), two-stage least squares (2SLS) and ordinary least squares (OLS). The SPSS Program analyzes data by default using OLS. By minimizing the squared differences between each observation and the regression line, OLS estimates a regression line. OLS prevents positive and negative departures from the regression line from canceling one another out by squaring distances. OLS also gives data that are far from the regression line more weight since it squares the distances between the observations.

iii. Testing Assumptions of Regression Analysis

For the regression analysis to yield trustworthy results, five presumptions must be met. These assumptions include linearity, zero expected mean error, homoscedasticity, lack of autocorrelation, and error distribution. If this requirement is not met by the regression analysis, it becomes difficult to determine the regression parametric significance even though the generated model will be reliable.

Linearity

To satisfy this assumption the relationship between the independent and dependent variables should be linear. A linear relationship enables the regression model equation, $y = \alpha + \beta 1 x 1$ + e, to be linear. Plotting independent variables against dependent variables can be used to determine whether the relationship between x and y variables is linear.

To assess whether there is a nonlinear pattern, scatter plots are used. Both straight and curved sections of the line were examined in this study to determine the relationship between the dependent and independent variables.

Homoscedasticity

This is a situation which errors variance are constant. Homoscedasticity means the relationship under investigation is the same for the entire range of the dependent variable. When homoscedasticity assumption is met, residuals will form a pattern less cloud of dots. The Homoscedasticity can be observed by creating graphs of errors against dependent variables(55). As dependent variable increases or decreases the errors variance should remain constant.

No autocorrelation of errors

This indicates the independence of the regression model errors. It denotes the absence of correlation between any two observations' errors. The Durbin-Watson test can be used to assess the autocorrelation of data in SPSS. By testing a null hypothesis of no autocorrelation against lower and higher bounds for negative autocorrelation and against lower and upper bounds for positive autocorrelation, the Durbin–Watson test determines if there is autocorrelation. To carry out this test, first sort the data on the variable that indicates the time dimension in your data.

Expected Mean Error of regression model is zero

The sum of expected errors must be zero in order to satisfy this assumption. The generated regression line will either overestimate or underestimate the true relationship between the variables if this assumption is not satisfied. the Ordinary Least Square (OLS) algorithm always chosen the ideal line, where the anticipated mean error is zero, So, as long as OLS is used as the estimation selection method, this assumption does not need to be tested(9).

Error Distribution

This assumption states that the errors in the regression model should approximately follow a normal distribution. Since the regression model still produces reliable estimates of the coefficients even in the case where the errors are not normally distributed, this assumption is regarded as optional(9). Plots and the execution of a formal test are two techniques used to determine whether the errors in a regression model are normally distributed. This test is not carried out in this particular study.

iv. Interpretation of Regression Results

Performance of regression model

The effectiveness of the constructed model is examined by determining the model's goodness of fit. R^2 and the mean squared error are used in this study to evaluate the model's performance. The R^2 (coefficient of determination) measures how well the model accounts for observed deviations from the mean in the dependent variable. Regression line closeness to observed values is shown by R^2. The R^2 is defined by equation 3.3

$$R^2 = \frac{(\hat{y}-\bar{y})^2}{(y-\bar{y})^2} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.\ 3)$$

Where:

$\hat{y}$: regression line

$\bar{y}$: mean absolute values

y: observed values.

In this research, y represents the calculated BQI from the observed longitudinal level.

The R^2 value explains the performance of the model in predicting the future ballast condition. The R^2 always lies between 0 and 1, where a higher R^2 indicates a better model accuracy. The closer R^2 is to 1 indicate the ability of the model to account the variation of depend variables by the independent variables.

Mean squared Error (MSE) is the average squared difference between the estimated values and the actual value. The MSE value is a non-negative value and values closer to zero are better and are used to measure the performance of the model .The formula governing the MSE value is given by equation 3.4.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \quad \text{...} \quad (3.\ 4)$$

Where:

MSE: Mean Squared Error,

n:Number of data points,

y_i: Observed values,

$\hat{y}_i$= Predicted Values.

v. Validation of Regression Model

It is essential to look at the regression model's stability. Stability describes the results as being consistent over time, not varying under various conditions and being heavily dependent on the model's specification. The regression results are validated by splitting sample data into two parts (called split-sample validation) and run the regression model again on each subset of data. From the total sample data 70% of the randomly chosen data are often used to estimate the regression model and the remaining 30% are used for validation purpose. The data can only split if the remaining 30% still meets the sample size rules of thumb discussed earlier. If the use of the two samples results in similar effects, we can conclude that the model is stable.

vi. Use of Regression Model

It is time to employ the regression model after a suitable regression model that conforms to regression analysis's presumptions has been found. Use of regression model in this study is prediction. Prediction entails calculating the values of the dependent variable based on predetermined values of the independent variables and their related but previously calculated unstandardized coefficients (β). From the results found in chapter four, the predicted BQI is a function of tonnage, number of past tamping speed, gradient subgrade condition and season for the straight sections of the line. For the curved sections, the predicted BQI is function of the above independent variables and curve radius. The functions can be defined by equation 3.5 and equation 3.6 for the straight and curved sections respectively.

BQI (t) = f (tonnage (t-1), number of past tamping, speed, gradient, subgrade condition, season)………………………………. (3. 5)

BQI (t) = f (tonnage (t-1), Curve Radius, number of past tamping, speed, gradient, subgrade condition, season)……………………………………. (3. 6)

CHAPTER FOUR

4. RESULTS AND DISCUSSION

4.1. Results

In this particular chapter, fractal analysis results (fractal dimensions of different sections of the track), factors influencing Ballast degradation followed with development of degradation models' results are discussed. Two models were developed to predict the Ballast degradation of AALRT. Such models are for straight sections and curved sections of the lines. Correlation analysis and ANOVA analysis results are illustrated in Tables 4-2 to Table 4-18. The results for Multiple regression models are illustrated in Tables 4-19 to 4 - 28.

4.1.1. Fractal dimensions

Fractal dimension results, which denotes the ballast quality index in this study, are within the range of one and two which shares the finding of (46) and (47)

4.1.2. Factors Influencing Ballast Degradation

Understanding the degradation mechanism of railway ballast is a crucial step in planning and optimizing the maintenance activities. Before development of degradation pattern of railway ballast there is a need of studying the factors causing such degradation. A number of researchers to contribute to Railway track ballast degradation has investigated various factors. These previously considered and other additional factors are evaluated in this study. Factors discussed in this study are tonnage, number of trips, curve radius, gradient, speed, number of past tamping, track sub grade condition and season.

Tonnage

Tonnage has been found by different researchers to be the main contributor on railway track ballast deterioration. Tonnage is a total load, which passes on a track including the weight of the trains and passengers. Monthly data of number of passengers transported in the line was collected from operation control center (OCC) of AALRT office and was used to generate the total tonnage passing through the line. The calculations of monthly tonnage passing through the line corresponding to the inspection data dates was done so as to perform correlation between the dependent and independent variable. Example of such calculation is presented in Table 4-1 for the month of September 2019. Eleven months of data were collected and applied in this study to analyze the effect on ballast degradation. Tonnage is taken as continuous variable in this study. Continuous variable is variables which can take an infinite number of values. The correlation analysis was conducted between the tonnage and change of Ballast quality index. The correlation measures how the variables relate. SPSS Software was used to perform correlation analysis between dependent and independent variables for straight and curved section of AALRT.

Table 4-1: Calculation done to obtain tonnage passing on the line

Item	Value
Passenger flow (a)	1655798
Average weight of passenger (b)	65Kg= 0.065tons
Passenger flow in tons $c = a * b$	107,626.87tons
Weight of trains (d) Consider empty vehicle	= 44 tons
Total weight on the line $g = c + d$	107,670.87 tons
Total weight in MGT	0.10767087 tons

The above calculation was done on all the available data to obtain the tonnage accommodated on the line for each month data from September 2019 to July 2020. The correlation analysis is made between the obtained tonnage (MGT) and change in BQI to test the significance of tonnage in affecting the ballast quality and the obtained results are presented in Table 4-2 and Table 4-3 for straight and curved sections respectively.

Table 4-2: Correlation analysis between changes in BQI value and tonnage for straight section of the line

		Ballast Quality Index	Tonnage
Ballast Quality Index	Pearson Correlation	1	-.865**
	Sig. (2-tailed)		.000
	N	573	573
Tonnage	Pearson Correlation	-.865**	1
	Sig. (2-tailed)	.000	
	N	573	573

**. Correlation is significant at the 0.01 level (2-tailed).

Table 4-2 shows that the correlation between the change in Ballast quality index and Tonnage for straight sections of the line is negative correlation with a correlation value of 0.865. This implies that as the passing tonnage increases on the line the ballast quality becomes degraded. Hence, the Significance (P-value) for the correlation analysis is less than 0.05 the variables tonnage significantly influences the ballast quality and ballast quality reduces with the gross tonnage on the line.

Table 4-3: Correlation analysis between changes in BQI value and tonnage for curved section of the line

		Ballast quality index	Tonnage
Ballast quality index	Pearson Correlation	1	-.870**
	Sig. (2-tailed)		.000
	N	347	347
Tonnage	Pearson Correlation	-.870**	1
	Sig. (2-tailed)	.000	
	N	347	347

**. Correlation is significant at the 0.01 level (2-tailed).

Table 4-3 shows that there is negative correlation between the Ballast quality index and tonnage with correlation value of 0.870. The determined Significance ((P-value) is less than 0.05, which implies that correlation between variables is significant and tonnage significantly affects the ballast quality for the curved section of the line as for the straight section.

In conclusion, the relationship between change in Ballast Quality Index and Tonnage are statically significant for both straight and curved sections of the line.

Number of trips

Number of trains passing through the track segment is termed as number of trips. AALRT records the number of trips for both lines every day. An average of 130 Number of trips are made by AALRT trains each day. Number of trips is also taken as continuous variable in this study. Correlation analysis is done to determine the effect of number of trip on the ballast quality and the results are illustrated in Table 4-4 and Table 4-5 for the straight and curved sections of the line.

Table 4-4: Correlation analysis between changes in BQI value and number of trip for straight section of the line

		Ballast Quality Index	number of trip
Ballast Quality Index	Pearson Correlation	1	-.813**
	Sig. (2-tailed)		.000
	N	573	573
number of trip	Pearson Correlation	-.813**	1
	Sig. (2-tailed)	.000	
	N	573	573

**. Correlation is significant at the 0.01 level (2-tailed).

From Table 4-4 it is shown that there is negative correlation between number of trips and Ballast quality index for Straight section of line with a correlation value of 0.813. The obtained correlation values imply that as number of trips increases the Ballast quality index decreases. The P- value for the correlation analysis becomes less than 0.05 implying that the relationship between number of trips and Ballast quality index is statically significant.

Table 4- 5: Correlation analysis between changes in BQI value and number of trip for curved section of the line

		Ballast quality index	number of trip
Ballast quality index	Pearson Correlation	1	-.809**
	Sig. (2-tailed)		.000
	N	347	347
number of trip	Pearson Correlation	-.809**	1
	Sig. (2-tailed)	.000	
	N	347	347

**. Correlation is significant at the 0.01 level (2-tailed).

As shown in Table 4-5 there exists negative correlation between number of trips and Ballast quality index for curved section of line as that of the straight section. The obtained correlation value imply that as number of trips increases the Ballast quality index decreases. The obtained P- value for the correlation analysis is less than 0.05 implying a significant relationship between number of trips and Ballast quality index.

As the number of trips increases on the line, the ballast quality degrades significantly.

Curve Radius

Addis Ababa Light Rail transit has a total of 87 curved section, 41 curves in East West East West (E-W) line which gives a total length of 5.55Km or 31.8% of total length of the line and 46 curves in North South (N-S) line with a total length of 6.59Km or 39.8% of total line. The maximum and minimum curve radius in N-S line is 2004m and 50m respectively and the maximum and minimum curve radius in E-W line is 3004m and 154m respectively. Presence of curves on a railway track are inevitable but are said to be one of the contributing factors on track geometry degradation. The presence of sharp curves on a line contributes greatly in track irregularity. To investigate the contribution of curve radius on change of Ballast Quality Index correlation analysis was conducted only on the curved section of AALRT where the effect can be observed. The results obtained in the correlation analysis are shown in Table 4-6.

Table 4- 6: Correlation analysis between changes in BQI value and curve radius for curved section of the line

		Ballast quality index	radius
Ballast quality index	Pearson Correlation	1	.188**
	Sig. (2-tailed)		.000
	N	347	347
Radius	Pearson Correlation	.188**	1
	Sig. (2-tailed)	.000	
	N	347	347

**. Correlation is significant at the 0.01 level (2-tailed).

The results illustrated in Table 4-6 shows that there is positive correlation between curve radius and change in Ballast quality index. This implies that the ballast quality degrades on sections with smaller curve radius than that of having larger curve radius. The obtained P- values are less than 0.05 which shows that curve radius significantly influences the ballast quality.

Speed

Based on different reviewed studies, which have included speed in degradation modeling, speed directly influence the deterioration of track geometry. Addis Ababa Light Rail transit has a maximum design speed of 80km/h and average speed of

18km/h. There are speed restrictions in sections of AALRT impacted due to different reasons such as construction control, track alignment, line route and rail joint scheme. In this particular study, Maximum speed used passing through track section is used and correlation with change in ballast quality index are calculated. Table 4-7 and Table 4-8 shows correlation analysis results between BQI and Speed for straight and curved sections of line respectively.

Table 4- 7: Correlation analysis between changes in BQI value and speed for straight section of the line

		Ballast Quality Index	speed
Ballast Quality Index	Pearson Correlation	1	-.154**
	Sig. (2-tailed)		.000
	N	573	573
Speed	Pearson Correlation	-.154**	1
	Sig. (2-tailed)	.000	
	N	573	573

**. Correlation is significant at the 0.01 level (2-tailed).

As shown from the Table 4-7 the correlation value becomes -0.154 implying that the speed and ballast quality index have inverse relationships such that as travelling speed increases the ballast quality becomes degraded. It is also shown that the relationship between speed and change in Ballast quality index for straight sections is significant (P-value less than 0.005).

Table 4- 8: Correlation analysis between changes in BQI value and speed for curved section of the line

		Ballast quality index	speed
Ballast quality index	Pearson Correlation	1	-.227**
	Sig. (2-tailed)		.000
	N	347	347
Speed	Pearson Correlation	-.227**	1
	Sig. (2-tailed)	.000	
	N	347	347

**. Correlation is significant at the 0.01 level (2-tailed).

The results illustrated in Table 4-8 also shows similar relationships with that of the straight section that there is negative relationship between speed and change in Ballast quality index for curved sections and the travelling speed has significant effect on the ballast quality of the line (P- value less than 0.05).

Gradient

Addis Ababa is one of highest capital city in the world located at 2.5355m above the sea level. Presence of other modes of transport such as roads forces rail to adopt to existing model to be accessible to users. Due to this, challenges such as strong vertical grades, vertical curves and horizontal circular curves with minimum radius arises. (56). AALRT is designed with maximum gradient of 50‰. The gradient of level grade in the section is not less than 3‰. The line profile drawing is used to correlate the grade of sections of line with the ballast quality index. Table 4-9 and Table 4-10 depict the correlation analysis results between ballast quality index and gradient for straight and curved sections respectively.

Table 4- 9: Correlation analysis between changes in BQI value and gradient for straight section of the line

		Ballast Quality Index	Grade
Ballast Quality Index	Pearson Correlation	1	-.130**
	Sig. (2-tailed)		.002
	N	573	573
Grade	Pearson Correlation	-.130**	1
	Sig. (2-tailed)	.002	
	N	573	573

**. Correlation is significant at the 0.01 level (2-tailed).

Table 4- 10: Correlation between changes in BQI value and gradient for curved section of the line

		ballast quality index	Grade
Ballast quality index	Pearson Correlation	1	-.190**
	Sig. (2-tailed)		.000
	N	347	347
Gradient	Pearson Correlation	-.190**	1
	Sig. (2-tailed)	.000	
	N	347	347

**. Correlation is significant at the 0.01 level (2-tailed).

The results illustrated in Table 4-9, 4-10 shows that gradient significantly affects the ballast quality of both straight and curved sections (P-value less than 0.005). There exists a negative relationship between the two variables in both straight and curved sections implying that for sections with highest gradient the ballast quality degrades higher than that of with lower gradient.

Number of past tamping

Tamping operations are carried out on railway track to correct track geometry of sections with geometry parameters that exceed the tolerable limit. The effect of number of past tamping on ballast quality index is evaluated by counting the number of tamping carried out in each section. Number of past tamping is taken as continuous variable and the correlation analysis results obtained between BQI and number of past tamping is illustrated in Table 4-11 and Table 4-12 for straight and curved sections of line respectively.

Table 4- 11: Correlation between changes in BQI value and number of past tamping for straight section of the line

		Ballast Quality Index	number of tamping
Ballast Quality Index	Pearson Correlation	1	-.157**
	Sig. (2-tailed)		.000
	N	573	573
number of tamping	Pearson Correlation	-.157**	1
	Sig. (2-tailed)	.000	
	N	573	573

**. Correlation is significant at the 0.01 level (2-tailed).

Table 4- 12: Correlation between changes in BQI value and number of past tamping for curved section of the line

		Ballast quality index	number of tamping
Ballast quality index	Pearson Correlation	1	-.140**
	Sig. (2-tailed)		.009
	N	347	347
number of tamping	Pearson Correlation	-.140**	1
	Sig. (2-tailed)	.009	
	N	347	347

**. Correlation is significant at the 0.01 level (2-tailed).

The results illustrated in Table 4-11 and Table 4-12 shows that there is negative correlation between number of past tamping and change in Ballast quality index in both straight and curved sections of the line. This implies that the ballast degrades rapidly with the repetition of tamping actions on the line. Hence, the obtained P- values are less than 0.05, the relationship between the two variables is statically significant in both straight and curved sections of the line.

Sub grade condition

Track sub grade condition includes whether track is constructed on cut, fill or level ground. The profile drawing of the line is used to identify the sub grade condition of the sections. Sub grade condition is considered as categorical variable and one-way ANOVA was used to correlate the two variables. In the effect analysis cut sections were assigned as 0, fill sections were assigned as 1 and level sections were assigned as 10. Table 4-13 and Table 4-14 shows the significance test analysis between BQI and sub grade condition for straight and curved sections of line respectively.

Table 4- 13: Correlation between changes in BQI value and sub grade condition for straight section of the line

Ballast Quality Index

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	.407	2	.203	9.561	.000
Within Groups	12.121	570	.021		
Total	12.528	572			

Table 4- 14: Correlation between changes in BQI value and sub grade condition for curved section of the line

Ballast quality index

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	.249	2	.125	14.878	.000
Within Groups	2.880	344	.008		
Total	3.129	346			

From the results illustrated in Table 4-13 and Table 4-14 it is shown that track sub grade condition significantly influences ballast quality for both straight and curved sections (P-value less than 0.005). To determine whether the ballast quality degrades on the fill or cut section, scatterplots of the variables is extracted and used to compare each of the sections. Figure 4-1 depicts the variation of the ballast quality over sections with different subgrade condition. As shown from Figure 4-1 most of the cut sections have lower ballast quality index than that of the fill sections.

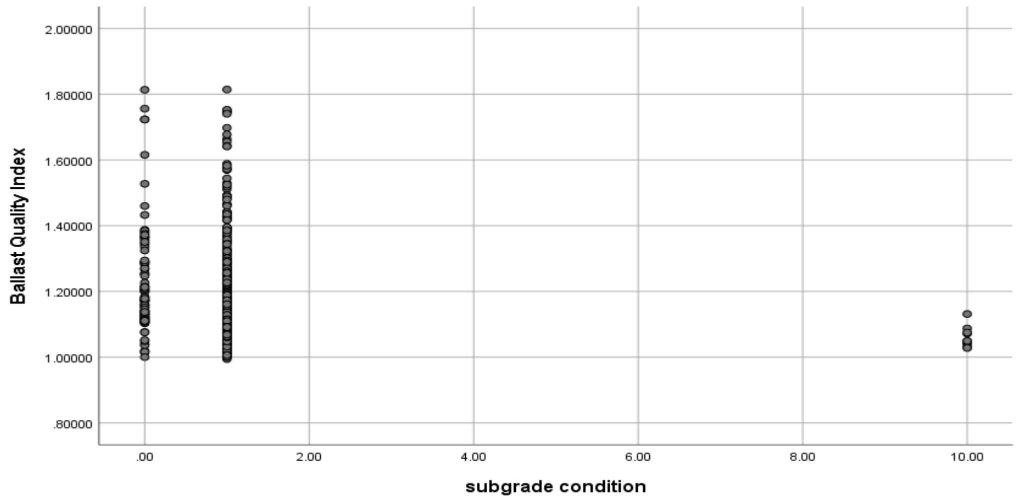


Figure 4- 1: comparison of ballast quality over subgrade condition

Whether condition

Ballast quality may change with the change in climatic condition. The effect of whether condition on the ballast quality index is taken in to account by considering the climatic condition of Addis Ababa city. The variable is considered as discrete variable and one way ANOVA is used to correlate with the change of ballast quality. According to different climatic studies done, the climatic condition of Addis Ababa City falls in to three climatic seasons, such as dry(February-May),semi dry(October-January)and Rainy(June-September). In the effect analysis 1 assigns Rainy season, 2 assigns semi dry season and 3 assigns dry season. Table 4-15 and Table 4-16 shows the relationships of BQI and weather condition for straight and curved sections of line.

Table 4- 15: Correlation between changes in BQI value and season for straight section of the line

Ballast Quality Index

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	.565	2	.283	13.469	.000
Within Groups	11.962	570	.021		
Total	12.528	572			

Table 4- 16: Correlation between changes in BQI value and season for curved section of the line

Ballast quality index

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	.226	2	.113	11.495	.000
Within Groups	3.375	344	.010		
Total	3.601	346			

From the results illustrated in Table 4-15 and Table 4-16, the P- values are less than 0.05, this implies that weather condition significantly influences the ballast quality in both straight and curved sections the line. To determine on which season the ballast quality degrades more scatterplots of the variables is extracted and used to compare each of the seasons. Figure 4-2 illustrates the variation of the ballast quality over the considered seasons. As shown from Figure 4-2 there exists higher ballast degradation in rainy season than dry season.

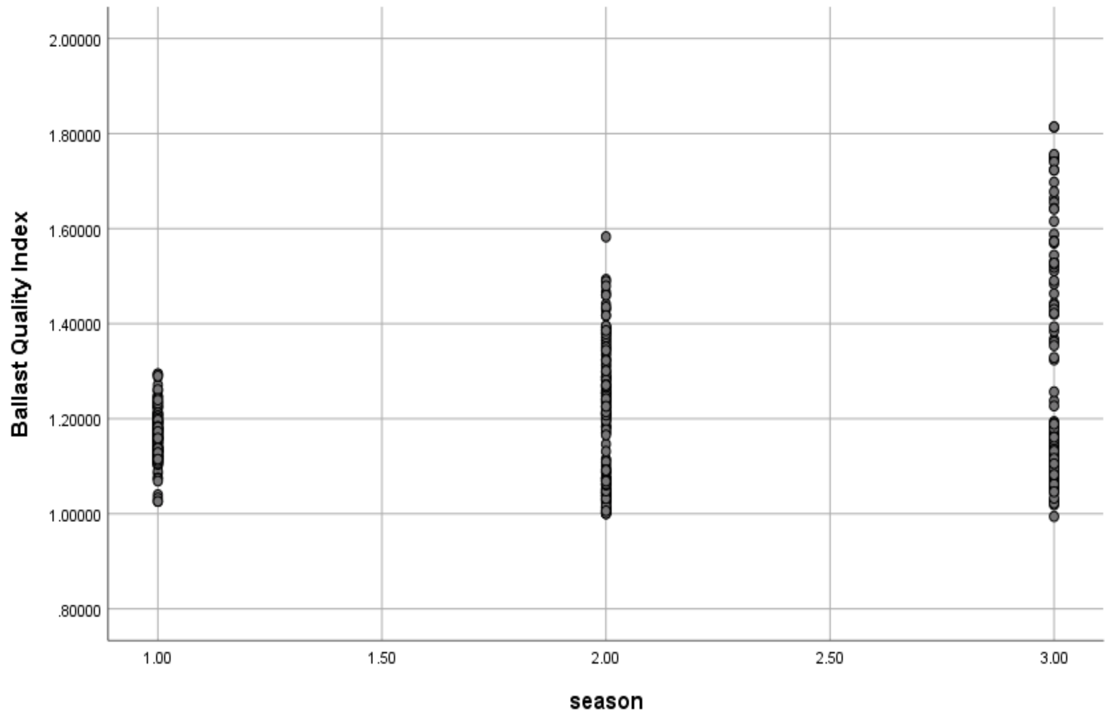


Figure 4- 2: comparison of ballast quality over different season

Tables 4-17 and 4-18 below provide the summary of the relationship between the dependent and independent variables. Table 4-17 presents the summary of correlations analysis results between changes in BOQ and continuous independent variables for straight and curved sections of AALRT and Table 4-18 gives the summary of ANOVA between changes of BOQ and Track sub grade condition and whether condition (season) for straight and curved sections of AALRT.

Table 4- 17: Correlations between changes in Ballast Quality Index and independent variables for straight and curved section

Track geometry	**Straight**		**Curved**	
Variables	Correlation Values	P-Values	Correlation Values	P-Values
Tonnage	-0.865	0.000	-0.87	0.000
Number trip	-0.813	0.000	-0.809	0.000
Curve radius			0.188	0.000
Speed	-0.154	0.000	-0.227	0.000
Gradient	-0.13	0.0.002	-0.19	0.000
Number of tamping	-0.157	0.000	-0.14	0.009

Table 4- 18: ANOVA between changes in Ballast Quality Index and Track Sub grade condition and season for straight and curved section

Track geometry	**Straight**		**Curved**	
Variables	F- Values	P-Values	F- Values	P-Values
Sub grade condition	9.561	0.000	14.878	0.000
Season	13.469	0.000	11.495	0.000

4.1.3. Multiple Regression Model Results

The procedures followed in the development of regression model are described in details in chapter 3. These procedures include data requirement for regression analysis, specification and estimation of regression model, testing assumptions of regression analysis, interpretation of regression results, validation of regression results and use of regression model. Data requirement for regression analysis include aspects such as sample size, variables variation, scale type of variables and collinearity of independent variable. Sample Size, variable variation and scale type of variables is clearly stated in chapter three.

Collinearity diagnosis

Collinearity of independent variables checks for correlation between the independent variables. For this study number of trips, tonnage curve radius, number of past tamping, subgrade condition and season are independent variables used. Table 4-19 and Table 4-20 illustrate collinearity diagnosis results between the independent variables for straight and curved sections.

Table 4- 19: Collinearity diagnosis between independent variables for the straight sections

Variables	Coefficients	Tolerance	VIF
Tonnage	-0.735	0.060	16.582
Number of Trips	3.061E-7	0.065	15.334
Gradient	-6.874E-5	0.978	1.022
Speed	0.000	0.971	1.030
Number of tamping	-0.004	0.956	1.046
Cut	0.014	0.882	1.134
Fill			
Dry	0.104	0.366	2.733
Rainy	0.110	0.620	1.612

Table 4- 20: Collinearity diagnosis between independent variables for the curved sections

Variables	Coefficients	Tolerance	VIF
Tonnage	-0.385	0.071	13.696
Number of Trips	4.89E-7	0.070	14.238
Gradient	0.000	0.812	1.232
Radius	4.349E-6	0.921	1.086
Speed	0.000	0.939	1.065
Number of tamping	-0.003	0.970	1.031
Cut	0.021	0.278	3.592
Fill	0.011	0.309	3.241
Dry	0.041	0.391	2.556
Rainy	0.055	0.789	1.267

As shown from Table 4-19 and Table 4-20 the tolerance value obtained for the first two variables (tonnage and number of trip) is less than 0.1 and VIF value is greater than

10, which implies that there exists collinearity between the two variables (57). However, there is neither collinearity nor multi-collinearity between the other independent variables.

Now severe collinearity exists in the two independent variables and to remove this collinearity, one of the independent variables that shows collinearity must be removed from the regression model. In this case, number of trip is excluded from the regression model due that it shows opposite relationship from the effect analysis. After removing number of trip from the regression model, collinearity between the independent variables is removed as shown in Table 4-21.

Table 4- 21: Collinearity diagnosis between independent variables

Variables	Coefficients	Tolerance	VIF
Tonnage	-0.347	0.764	1.309
Gradient	0.000	0.87	1.149
Radius	4.376E-6	0.919	1.088
Speed	0.000	0.943	1.061
Number of tamping	-0.003	0.972	1.028
Cut	0.016	0.302	3.314
Fill	0.007	0.326	3.071
Dry	0.035	0.716	1.397
Rainy	0.054	0.844	1.184

Specification and estimation of regression model

This section involves the selection of dependent and independent variables for the regression model and the model estimation methods used to develop the model. The dependent variable in the model is the BQI and the independent variables are Tonnage, curve radius (only for the curved sections), number of past tamping, gradient, speed, subgrade condition and season. Selection of estimation method involves selection of methods to be used in estimation and in this study; ordinary least square (OLS) method is used.

Testing Assumptions of Regression Analysis

For testing assumption of regression analysis, five assumptions have to be satisfied to get valid results. These assumptions are; linearity, expected mean error is zero, homoscedasticity, No autocorrelation and error distribution.

Linearity

In checking linearity, independent variables are plotted against dependent variable. In this study, the relationship between the dependent and independent variables was checked for both straight and curved sections of AALRT. These relationships were checked using scatter plots as illustrated in Figure 4-3 to Figure 4-14.

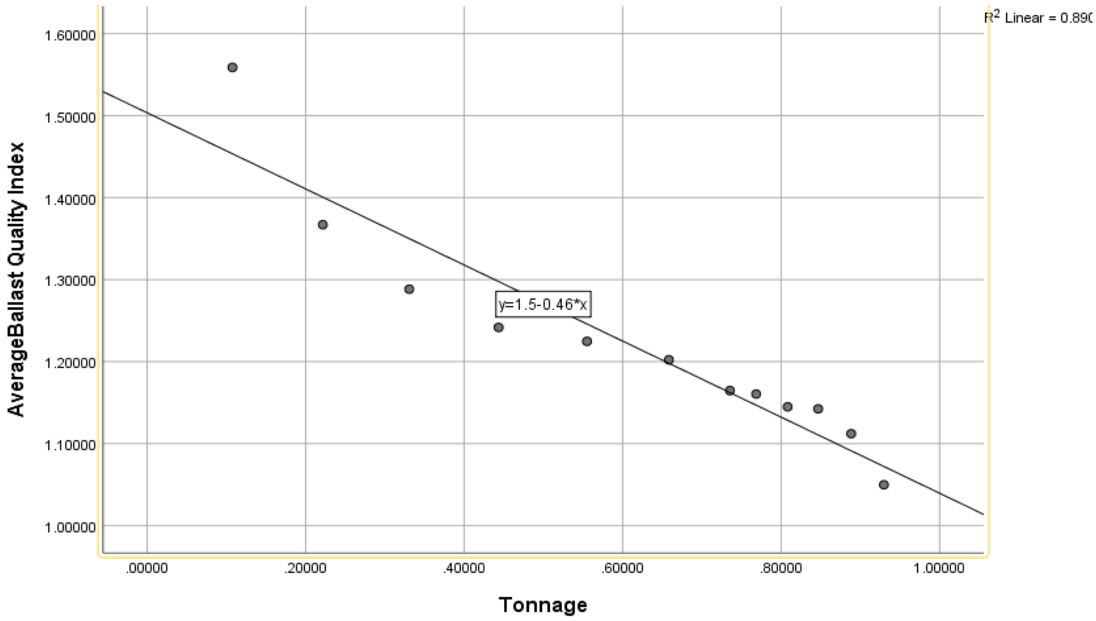


Figure 4- 3: Relationship between Ballast quality index and Tonnage for straight section

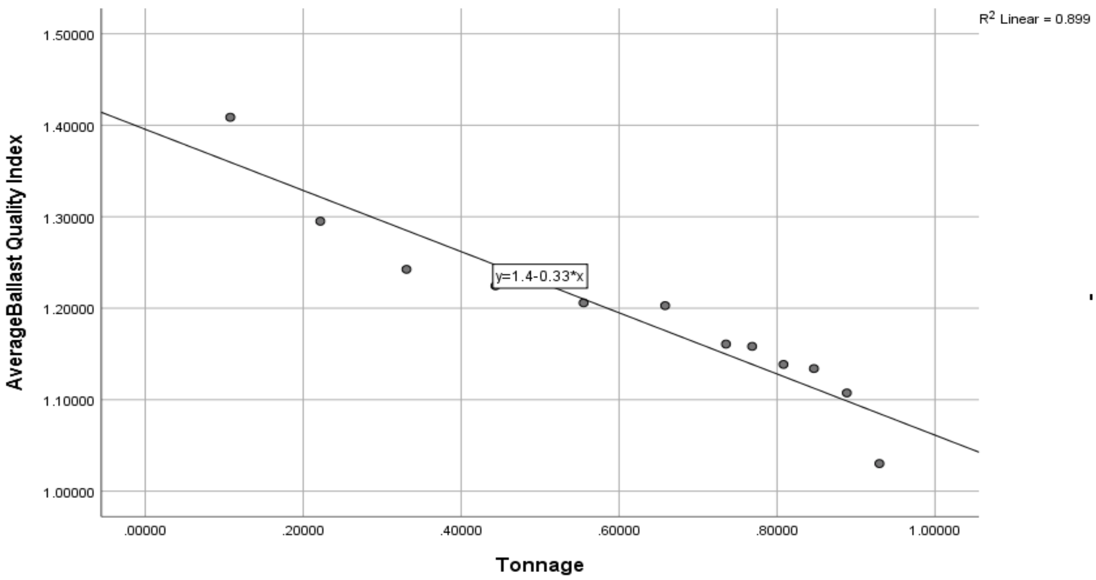


Figure 4- 4: Relationship between Ballast quality index and Tonnage for straight section

As shown from Figure 4-3 and Figure 4-4 there exists a linear relationship between ballast quality index and tonnage in both straight and curved sections of the line

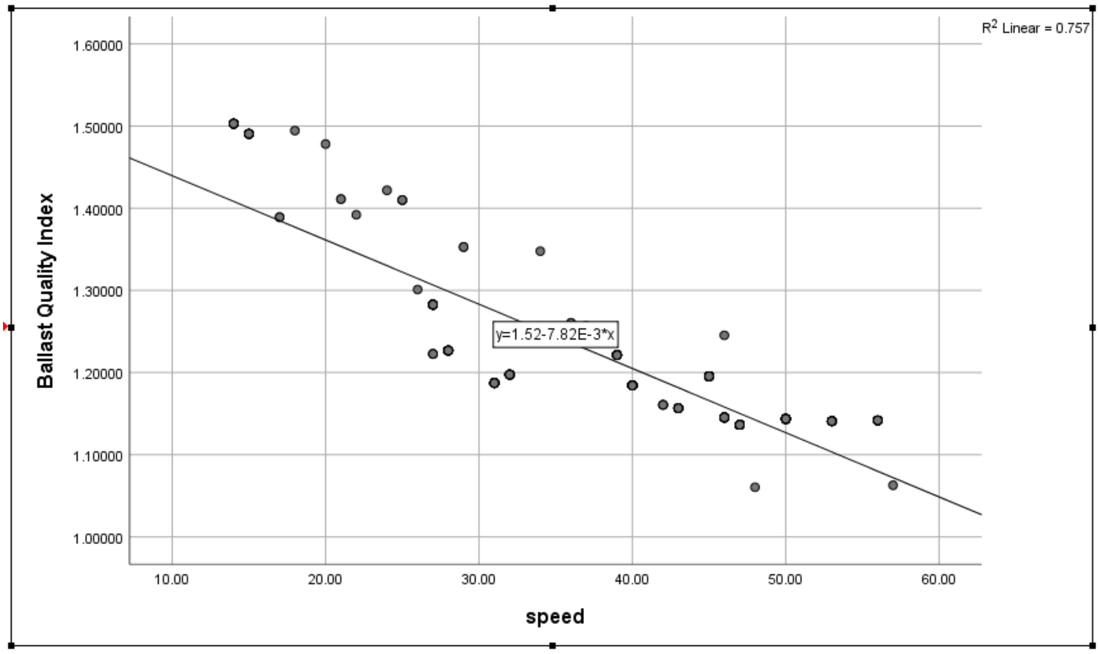


Figure 4- 5: Relationship between Ballast quality index and Speed for straight section

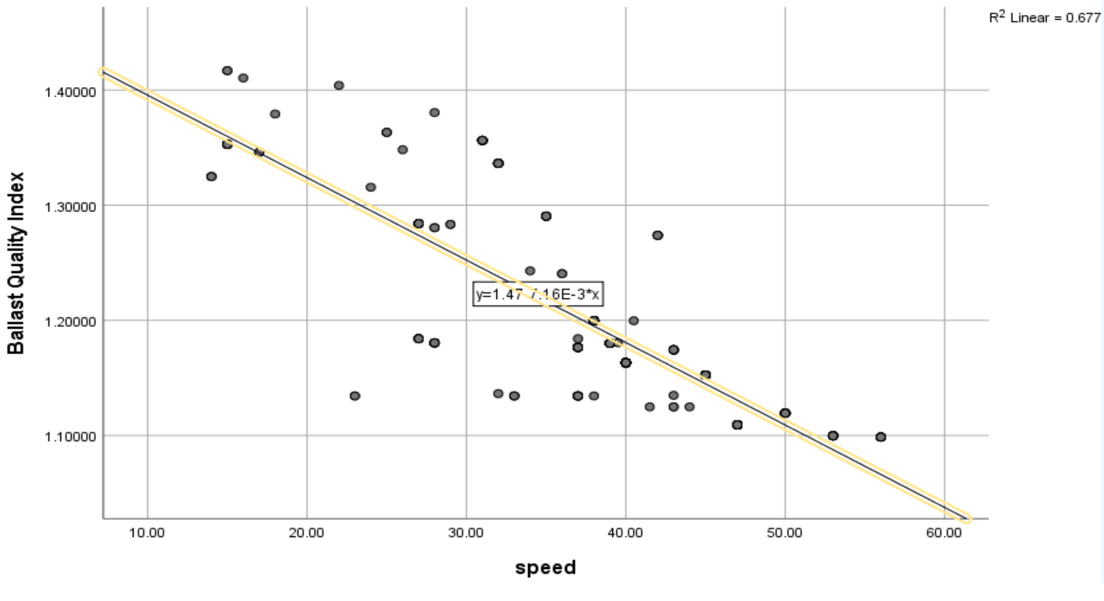


Figure 4- 6: Relationship between Ballast quality index and Speed for curved section

Figures 4-5 and 4-6 shows that a linear relationship exist between ballast quality index and speed in both straight and curved sections of the line

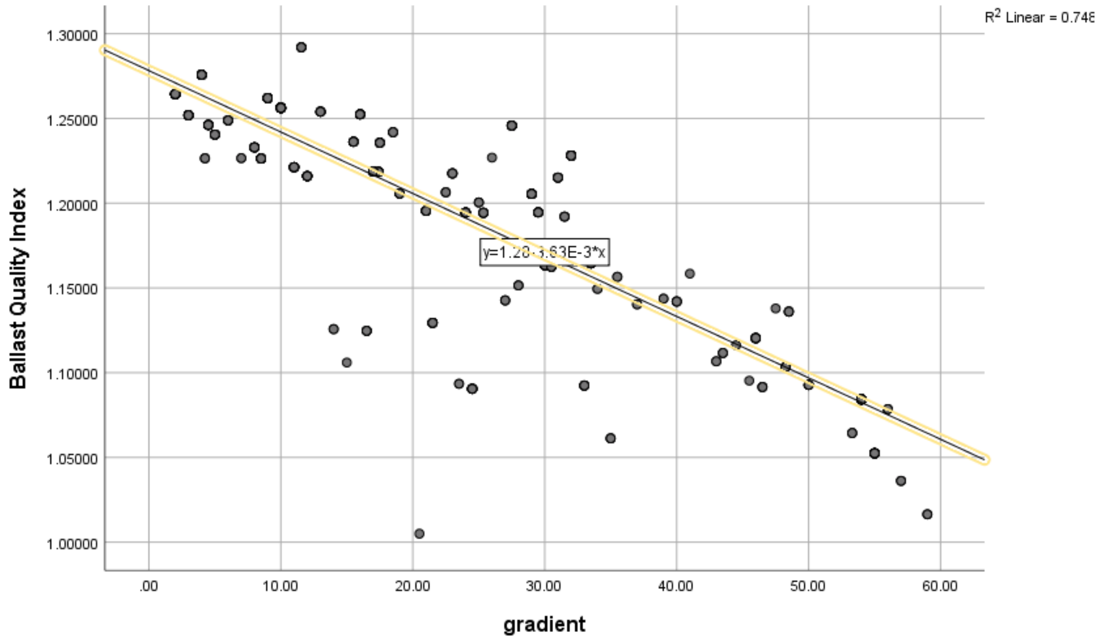


Figure 4- 7: Relationship between Ballast quality index and Gradient for straight section

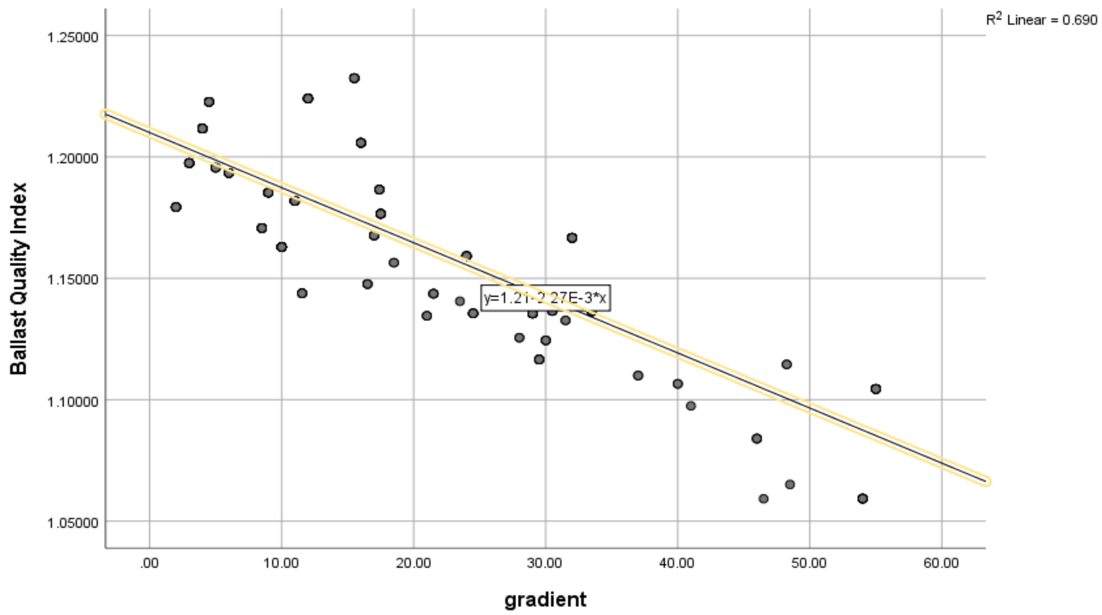


Figure 4- 8: Relationship between Ballast quality index and Gradient for curved section

As illustrated in Figure 4-7 and Figure 4-8 there exist a linear relationship between ballast quality index and gradient in both straight and curved sections of the line

Graph

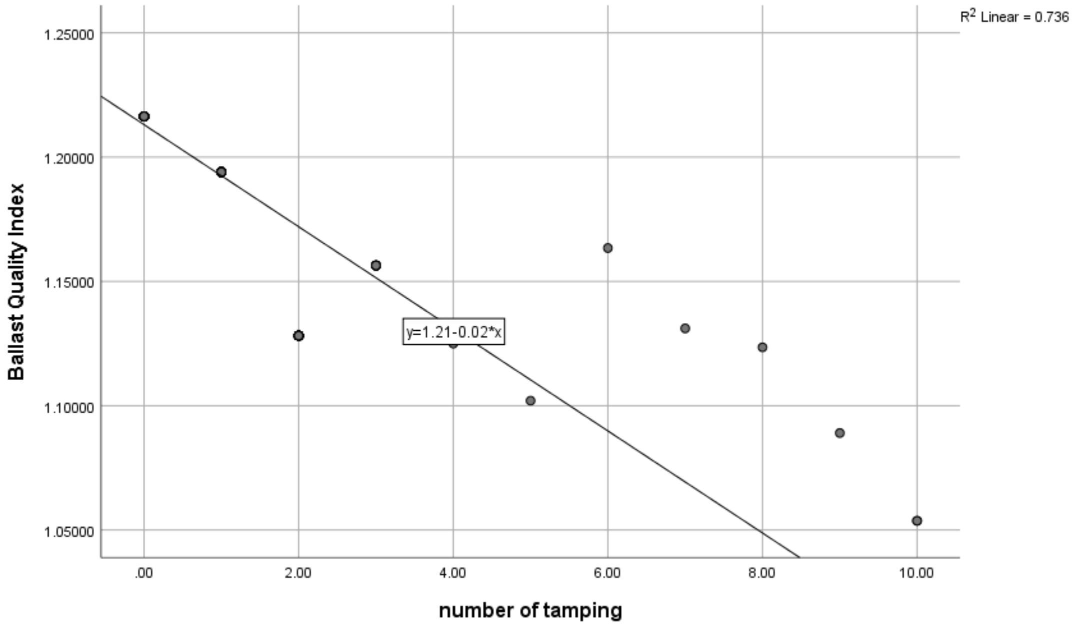


Figure 4- 9: Relationship between Ballast quality index and number of past tamping for straight section of the line

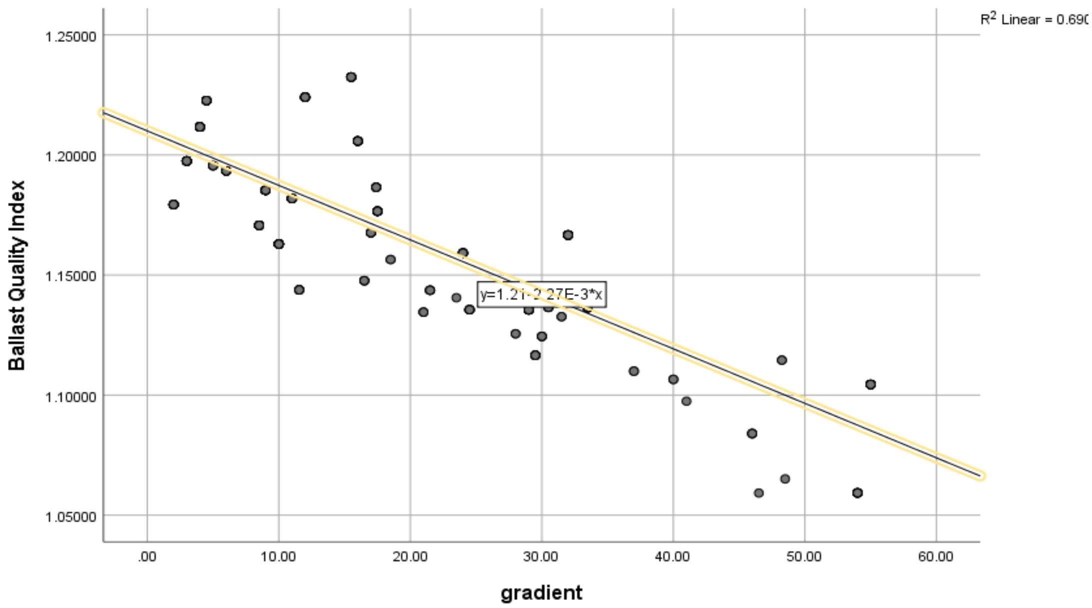


Figure 4- 10 : Relationship between Ballast quality index and number of past tamping for curved section of the line

The number of past tamping and ballast quality index exhibits a linear relationship in both straight and curved sections as shown in Figure 4- 9 and Figure 4- 10

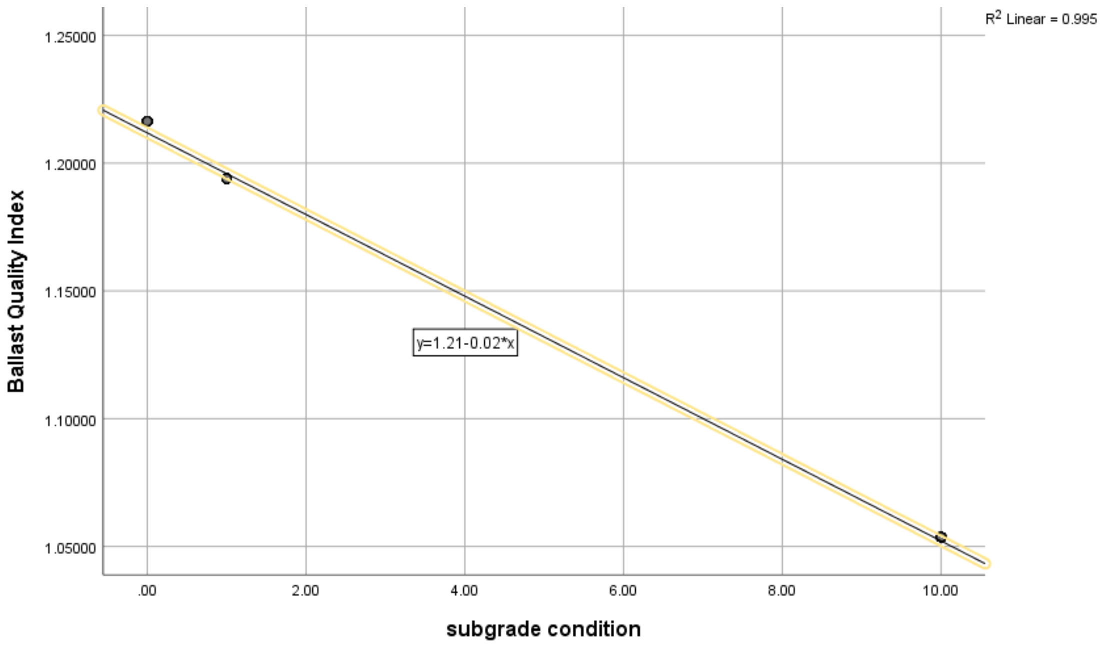


Figure 4- 11 : Relationship between Ballast quality index and subgrade condition for straight section of the line

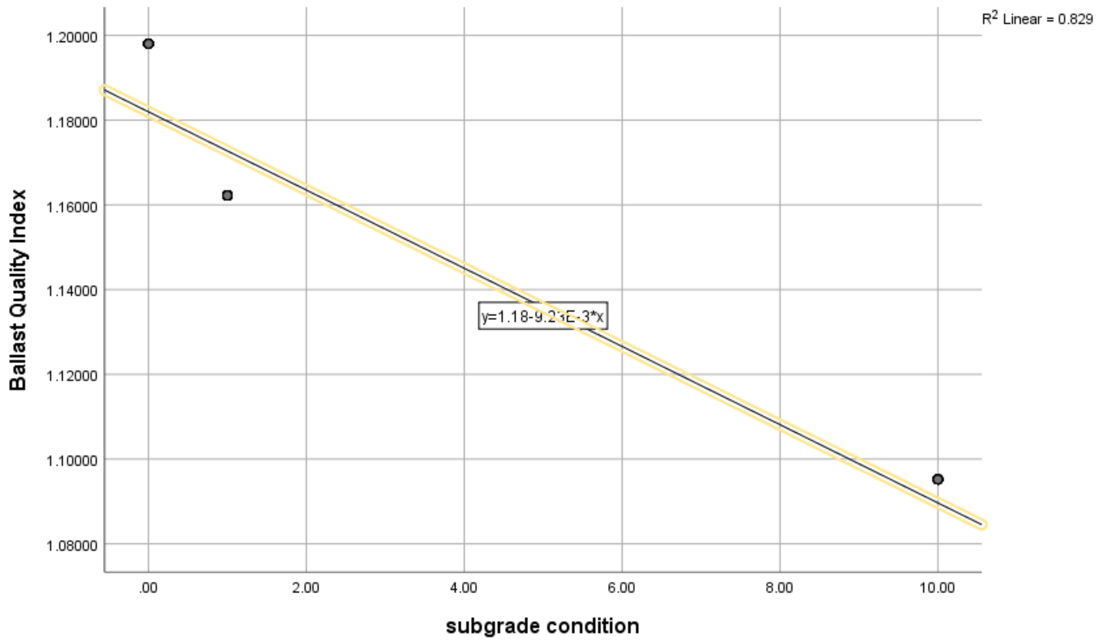


Figure 4- 12: Relationship between Ballast quality index and subgrade condition for curved section of the line

Figures 4- 11 and 4- 12 depicts that a linear relationship exists between ballast quality index and subgrade condition in both straight and curved sections of the line.

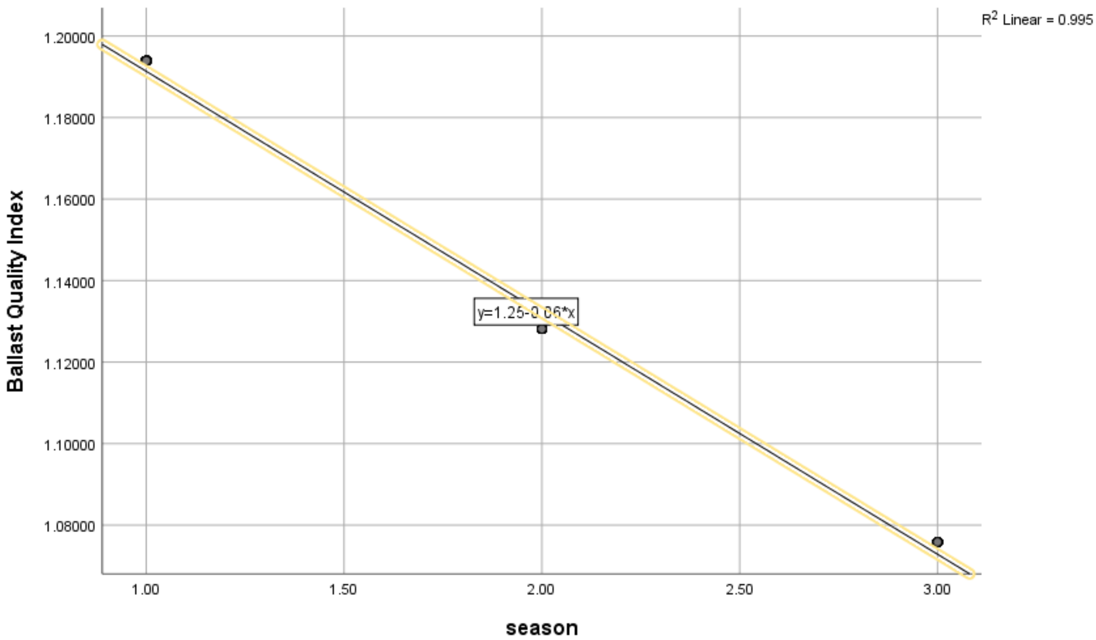


Figure 4- 13: Relationship between Ballast quality index and season for straight section of the line

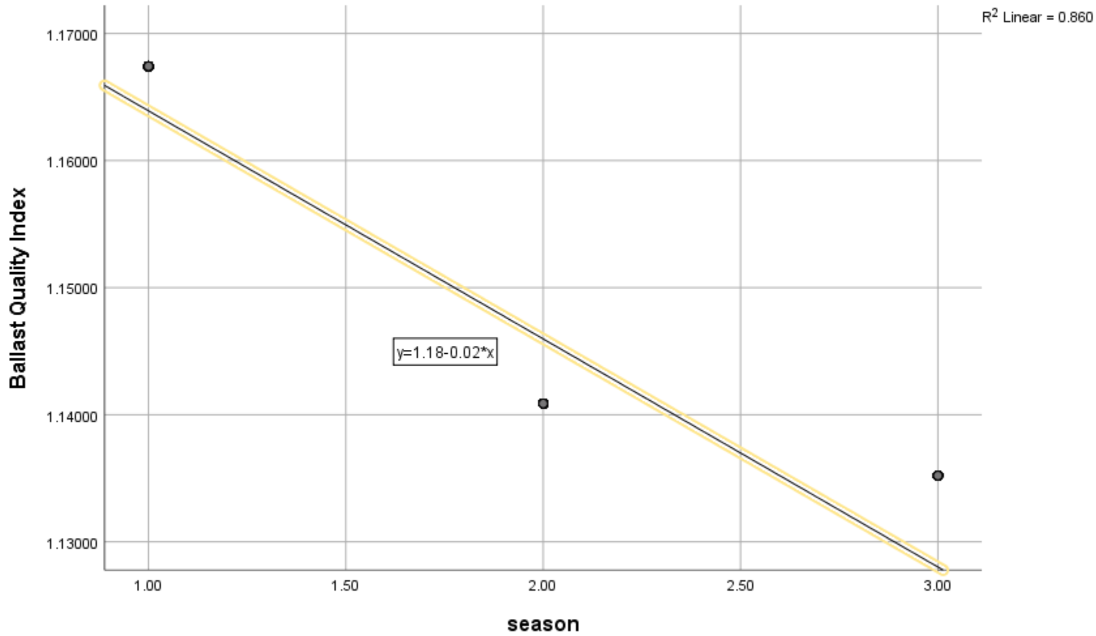


Figure 4- 14: Relationship between Ballast quality index and season for curved sections of the line

There also exists a linear relationship between ballast quality index and season in both straight and curved sections of the line as shown in Figures 4- 13 and 4- 14.

Homoscedasticity

To satisfy the assumption of homoscedasticity there should be a non-significant association between the predicted value and the residuals which implies that the multiple regression equation working equally well across the whole continuum of the independent variables(55). The results are as shown below

Table 4- 22: Pearson correlation between predicted ballast quality index and errors for straight sections

			Standardized Predicted Value	ABS_ZRE_1
Pearson	Standardized Predicted Value	Pearson Correlation	1	-.017
		Sig. (2-tailed)		.718
		N	479	479
	ABS_ZRE_1	Pearson Correlation	-.017	1
		Sig. (2-tailed)	.718	
		N	479	479

From the results shown in Table 4- 22 there exists a non-significant relationship between the predicted ballast quality index and the residual errors in straight sections of the line, this implies that homoscedasticity assumption is satisfied.

Table 4- 23: Pearson correlation between predicted ballast quality index and errors for curved sections

			Standardized Predicted Value	ABS_ZRE_1
Pearson	Standardized Predicted Value	Pearson Correlation	1	.026
		Sig. (2-tailed)		.667
		N	275	275
	ABS_ZRE_1	Pearson Correlation	.026	1
		Sig. (2-tailed)	.667	
		N	275	275

As in the straight sections, results illustrated in Table 4-23 shows a non-significant relationship between the predicted ballast quality index and the residual errors in curved sections that the homoscedasticity assumption is also satisfied in the curved sections of the line. Figure 4-15 also shows that homoscedasticity assumption is satisfied because residuals form a pattern of cloud of dots.

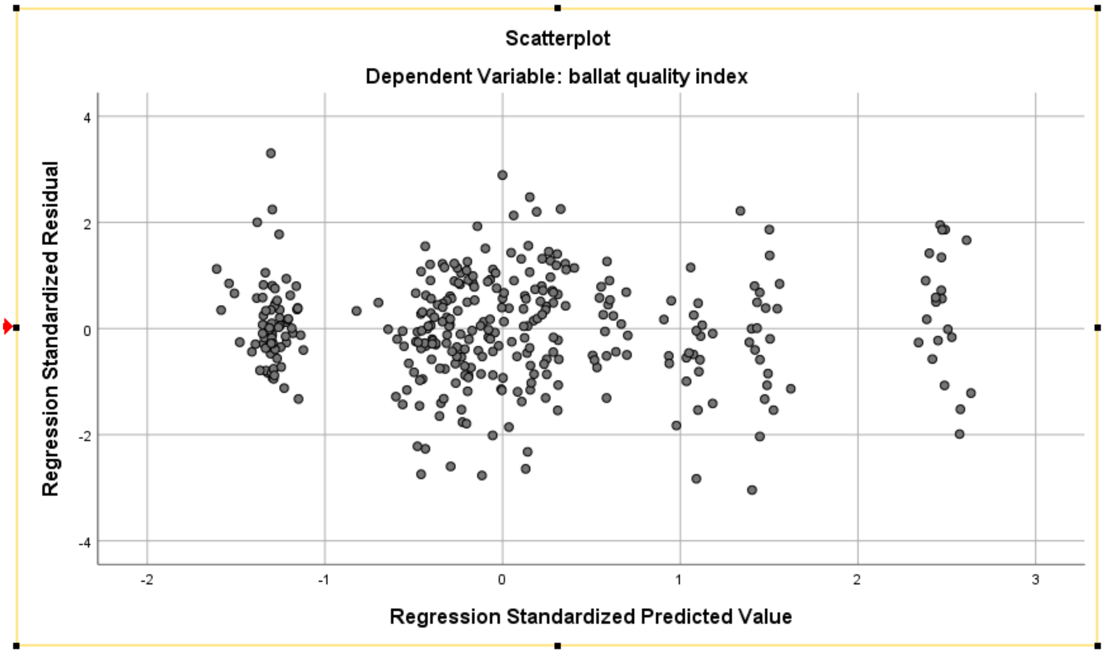


Figure 4- 15: homoscedasticity test by scatterplot

No autocorrelation of errors

The Durbin-Watson test can be used to assess the autocorrelation of data in SPSS. The Durbin-Watson coefficient should be between 1.5 and 2.5to satisfy this assumption(55). As illustrated in Table 4.25 and Table 4.27 the Durbin-Watson coefficient is 1.561 for straight sections and 1.748 for curved sections which lies between 1.5 and 2.5 showing there is no autocorrelation of errors.

Coefficients

Hence the assumptions for multiple regression are fulfilled the regression analysis is carried out by considering the independent variables that pass the testing session and the coefficients determined from the regression analysis for each of the independent variables are presented in Tables 4-24 and 4-26 for straight and curved sections of the line respectively. The model summary for each of the regression analysis models is also illustrated following the coefficients in Tables 4-25 and 4-27.

Table 4- 24: Estimated variables coefficients from regression model for straight section

Variable	coefficient	Values
Std. Error of the Estimate	E	0.0651
Constant	A	1.478
Tonnage	β1	-0.472
Gradient	β2	-7.865E-5
Speed	β3	-3.52E-10

Variable	coefficient	Values
Number of tamping	β4	-.006
Cut	β5	0.017
Fill	β6	0.013
Rainy	β7	0.085
Dry	β8	0.056

Table 4- 25: Model Summary of regression model for straight section

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.900[a]	.809	.807	.06509508	1.561

a. Predictors: (Constant), dry, grade, number of tamping, speed, fill, Tonnage, rainy, cut

b. Dependent Variable: Ballast Quality Index

Table 4- 26: Estimated variables coefficients from regression model for curved section

Variable	coefficient	Values
Std. Error of the Estimate	e	0.04454
Constant	a	1.379
Tonnage	β1	-0.342
Gradient	β2	-1.39E-10
Radius	β3	3.967E-6
Speed	β4	-3.89E-10
Number of tamping	β5	-0.003
Cut	β6	0.018
Fill	β7	0.008
Rainy	β8	0.034
Dry	β9	0.053

Table 4- 27: Model Summary of regression model for curved section

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.902[a]	.814	.809	.04453832	1.748

a. Predictors: (Constant), rainy, fill, number of tamping, speed, grade, radius, Tonnage, dry, cut

b. Dependent Variable: ballast quality index

Model Performance

The model performance is checked by the Goodness of fit of the model, which is checked by using R^2 and MSE to test the accuracy of the developed model. The values illustrated in Table 4-28 are obtained from the regression analysis results

Table 4- 28: model performance values for the developed models

	R^2	MSE
Straight section model	0.809	0.004
Curved section model	0.814	0.002

As shown from table 4-28 the R^2 value for both straight and curved sections of the line is closer to one indicating a better model accuracy. The MSE value obtained is smaller and closer to zero; this shows a better model accuracy.

Validation of Regression Model

The regression results are validated by running the regression model again on 30% of the sample data left for validation purposes and if the use of the two samples results in similar effects, we can conclude that the model is stable. While validating the developed model on 30% of the sample data in this study, similar effects were obtained with R^2 value of 0.80 and MSE of 0.005 in straight sections and R^2 value of 0.803 and MSE of 0.002 in curved sections, which guarantees the models stability. To show the validation the developed model the existed ballast quality index for the sections is plotted against the predicted ballast quality index as shown in Figure 4-16 and Figure 4-17.

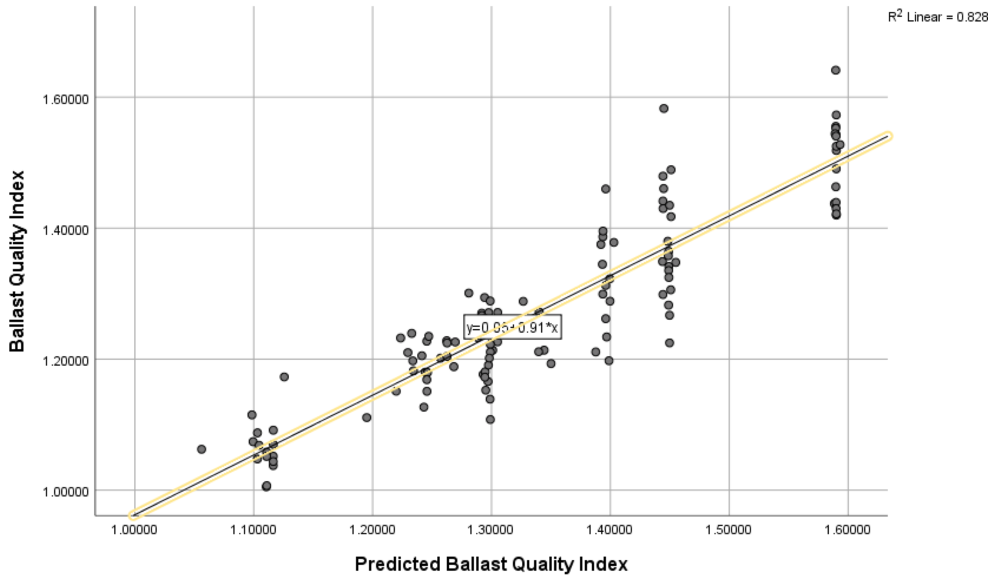


Figure 4- 16: validation plot for straight section model

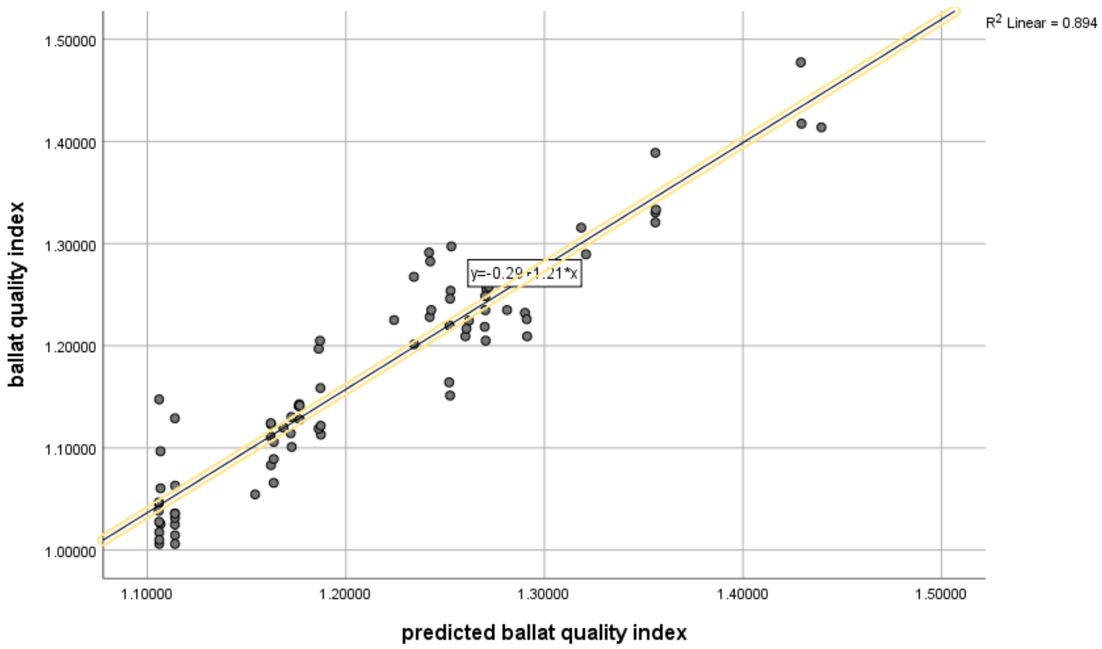


Figure 4- 17: validation plot for curved section model

Use of Regression Model

The use of the regression model in this study is prediction of ballast quality index from different variables. Prediction entails calculating the values of the dependent variable based on predetermined values of the independent variables and their calculated unstandardized coefficients (β). The estimation equations developed includes the dependent variable (Ballast quality index), independent variables (Tonnage, number of past tamping, speed, gradient, Curve radius, subgrade condition and season), regression coefficients and constants that are formulated as equation 4.1 and equation 4.2 for straight and curved sections respectively.

$$\text{BQI (t)} = 1.478 - 0.472\,Tonnage\,(t-1) - 7.865\text{E-}5 \text{ gradient} - 3.52\text{E-}4 \text{ speed} - 0.006 \text{ Number of tamping } (t-1) + 0.017\,(\text{Cut}) + 0.013\,(\text{Fill}) + 0.085\,(\text{Rainy}) + 0.056\,(\text{Dry}) + 0.0651 \quad \ldots\ldots (4.1)$$

$$\text{BQI (t)} = 1.379 - 0.342\,Tonnage\,(t-1) - 1.39\text{E-}10 \text{ gradient} + 3.967\text{E-}6 \text{ Radius} - 3.89\text{E-}10 \text{ speed} - 0.003 \text{ Number of tamping } (t-1) + 0.018\,(\text{Cut}) + 0.008\,(\text{Fill}) + 0.034\,(\text{Rainy}) + 0.053\,(\text{Dry}) + 0.04454 \quad \ldots\ldots (4.2)$$

Where: t is the time the ballast quality is calculated

t-1 one month before the time the ballast quality is calculated

4.2. Discussions

Two multiple regression models are developed in this study to predict the ballast quality degradation of AALRT. Two statistical measures namely R-squared and Mean Squared Error are applied to check the accuracy of developed models. The results of estimated coefficients from the models show that increasing MGT, number of tamping and being in Rainy season, all have increasing effect on ballast degradation and this finding shares the finding of (23). According (18) increasing the travelling speed minimizes ballast quality which is the finding of this study.

From (32) it was found that tamping increases ballast deterioration and for higher line speeds the track will require more maintenance which is also the finding of this study that for sections with higher travelling speed their exists higher ballast degradation and as tamping operations repeated the ballast becomes more degraded. It is found that tamping operations significantly degrades the ballast quality which shares the findings of (24)

This study addresses its objectives such that the ballast quality index is determined from the track longitudinal level by fractal analysis and the effect of various influencing variables on the ballast quality is evaluated by utilizing suitable procedures. The ballast degradation models for estimating the ballast quality from the influencing variables were developed for the straight and curved sections of AALRT and the models were validated by split sample validation in the study. The performance of the developed models were examined by R^2 and MSE and the results illustrated in table 4-30 entails the developed models are acceptable.

CHAPTER FIVE

5. CONCLUSIONS AND RECOMMENDATIONS

5.1. Conclusions

This study utilized a methodology for developing ballast degradation model to predict ballast quality from different parameters. Track geometry inspection data were used to evaluate ballast quality. The methodology employs fractal analysis of the track longitudinal level to estimate the degree of vertical profile irregularities related to the ballast. Box counting method which is the universal and widely accepted method of fractal analysis is used in the study to calculate the mid wavelength fractal dimensions from the inspected longitudinal level data. After determining the fractal dimensions that denotes the ballast quality index, the effect of different influencing variables including tonnage, number of trip, curve radius, number of past tamping, grade, speed, subgrade condition and weather condition on the ballast quality index are evaluated. To study the effect of these variables SPSS software is used. After identifying the significant variables multiple regression model is developed by SPSS software for straight and curved sections separately. The following conclusions were drawn from this study;

- ✓ Eight influencing variables were analyzed in this study namely; Number of trips, Tonnage, Speed, Grade, Curve Radius and Track subgrade condition and season. The analysis was done separately for straight and curved sections of AALRT. All analyzed factors were found to influence Ballast quality for both straight and curved sections of the line. All the variables except number of trips were employed to model the ballast degradation.
- ✓ Increasing tonnage and number of tamping reduces the ballast quality and ballast on cut sections degrades rapidly than the fill sections. It is also concluded that as the gradient of the line increases, the ballast quality degrades and the ballast degrades highly on rainy seasons than dry seasons. Furthermore, the ballast on sections with smaller radius degrades higher than that of with larger radius.

5.2. Recommendations

The degradation model was developed to predict the ballast quality of curved and straight section of AALRT. The following are recommendations based on the results of investigation in this research

- ✓ AALLRT need to compile daily track inspection data in a digital manner such as compiling all records in to excel sheet to made them ready for any research work.
- ✓ The inspection method also needed to be automated by employing use of track inspection car which provide wide range of data with high precision
- ✓ To allow further researches on the study area ballast renewal activities with their corresponding section need to be recorded

5.3. Future Work

- ✓ Developing ballast Maintenance model by including ballast renewal in the modelling process
- ✓ Developing overall track quality index and defining a relationship with ballast quality index
- ✓ Developing overall ballast Maintenance model by considering permeability coefficient of subgrade and track geometrical quality indicators

REFERENCES

1. Song M, Zhang G, Zeng W, Liu J, Fang K. Railway transportation and environmental efficiency in China. Transp Res Part D Transp Environ. 2016;48:488–98.
2. Sadeghi J, Askarinejad H. An investigation into the effects of track structural. 2009;223:415–25.
3. Jovanovic S, Evren G, Guler H. Modelling railway track geometry deterioration. Proc ICE - Transp. 2011 May;164:65–75.
4. Selig ET, Sluz A. Ballast and Subgrade Response To Train Loads. Transp Res Rec. 1978;(694):53–60.
5. Shenton MJ. Ballast Deformation and Track Deterioration. 1985;253–65.
6. Sato Y. Vehicle System Dynamics : International Journal of Vehicle Mechanics and Mobility Japanese Studies on Deterioration of Ballasted Track Japanese Studies on Deterioration of Ballasted Track. 2007;(September 2013):37–41.
7. Gebremariam SA, Alamnie MM. The Effect of Traffic on Ballast Settlement. 2021;669–87.
8. Fathali M, Chalabii J, Astaraki F, Esmaeili M. A new degradation model for life cycle assessment of railway ballast materials. Constr Build Mater. 2021;270(xxxx):121437.
9. Engineering R, Infrastructure C. African Railway Center of Excellence Analysis and Prediction of Track Geometry Degradation Case Study of Addis Ababa-Light Rail Transit. 2021;
10. Abadi T, Pen L Le, Zervos A, Powrie W. Improving the performance of railway tracks through ballast interventions. Proc Inst Mech Eng Part F J Rail Rapid Transit. 2018;232(2):337–55.
11. Abadi TC. Effect of sleeper and ballast interventions on rail track performance, PHD Thesis, Faculty of Engineering and the Environment, University of Southhamton, Southampton, England. 2015;(January):264.
12. Ababa A. Addis Ababa Institute of Technology MSc Railway Engineering (Civil Infrastructure) By : Advisor : Dr . Tensay Gebremedhin. 2021;(October).

13. Muinde MS. Railway track geometry inspection optimization. 2018;61.
14. Camacho D, Le TH, Rapp S, Martin U. Light rail ballasted track geometry quality evaluation using track recording car data. Comput Railw XV Railw Eng Des Oper. 2016;1(Cr):303–15.
15. de Melo ALO, Kaewunruen S, Papaelias M, Bernucci LLB, Motta R. Methods to Monitor and Evaluate the Deterioration of Track and Its Components in a Railway In-Service: A Systemic Review. Front Built Environ. 2020;6(September):1–15.
16. Soleimanmeigouni I, Ahmadi A, Nissen A, Xiao X. Prediction of railway track geometry defects: a case study. Struct Infrastruct Eng. 2020;16(7):987–1001.
17. Vale C, Simões ML. Prediction of Railway Track Condition for Preventive Maintenance by Using a Data-Driven Approach. Infrastructures. 2022;7(3).
18. Asadzadeh SM, Galeazzi R. An integrated methodology for the prognosis of ballast degradation in railway turnouts. Proc Inst Mech Eng Part F J Rail Rapid Transit. 2020;234(8):908–24.
19. Veit P. Prognosis of Switches Analogies and Differences to Open Track Life Cycle Management : A Step towards Sustainability Life Cycle Management. 2017;(August).
20. Soleimanmeigouni I, Ahmadi A, Kumar U. Track geometry degradation and maintenance modelling: A review. Proc Inst Mech Eng Part F J Rail Rapid Transit. 2018;232(1):73–102.
21. Chrismer S, Selig ET. Computer model for ballast maintenance planning. In: Proceedings of 5th international heavy haul railway conference, beijing, china. 1993. p. 223–7.
22. Ramūnas V, Vaitkus A, Laurinavičius A, Čygas D, Šiukščius A. Raudteeballasti tööea prognoos lähtudes selle mehhaanilistest omadustest. Balt J Road Bridg Eng. 2017;12(3):203–9.
23. Asadzadeh SM, Galeazzi R, Hovad E, Andersen JF, Thyregod C, Rodrigues A. Ballast Degradation Modeling for Turnouts based on Track Recording Car Data. Eur Conf PHM Soc. 2018;4(1).
24. Shi S, Gao L, Cai X, Yin H, Wang X. Effect of tamping operation on mechanical qualities of ballast bed based on DEM-MBD coupling method. Comput Geotech [Internet]. 2020;124(January):103574. Available from:

https://doi.org/10.1016/j.compgeo.2020.103574

25. Sadeghi J, Askarinejad H. Development of improved railway track degradation models. Struct Infrastruct Eng. 2010;6(6):675–88.
26. Yousefikia M, Moridpour S, Setunge S, Mazloumi E. Modeling Degradation of Tracks for Maintenance Planning on a Tram Line. J Traffic Logist Eng. 2014;2(2):86–91.
27. Jovanović S, Guler H, Čoko B. Track degradation analysis in the scope of railway infrastructure maintenance management systems. Gradjevinar. 2015;67(3):247–58.
28. Falamarzi A, Moridpour S, Nazem M, Hesami R. Rail degradation predication: Melbourne tram system case study. ATRF 2017 - Australas Transp Res Forum 2017, Proc. 2017;2018.
29. Chang H, Liu R, Li Q. A multi-stage linear prediction model for the irregularity of the longitudinal level over unit railway sections. WIT Trans Built Environ. 2010;114:641–50.
30. Galeazzi R. Springer Series in Reliability Engineering Intelligent Quality Assessment of Railway Switches and Crossings.
31. Wen M, Li R, Salling KB. PT US CR. Eur J Oper Res. 2016;
32. Audley M, Andrews JD. The effects of tamping on railway track geometry degradation. Proc Inst Mech Eng Part F J Rail Rapid Transit. 2013;227(4):376–91.
33. Elkhoury N. Analysis and Prediction of Tram Track Degradation. 2018;(April).
34. Khouzani AHE, Golroo A, Bagheri M. Railway maintenance management using a stochastic geometrical degradation model. J Transp Eng. 2016;143(1):1–9.
35. Scanlan KM, Hendry MT, Martin CD. Evaluating the impact of ballast undercutting on the roughness of track geometry over different subgrade conditions. 2017;0(0):1–11.
36. Design L cycle, Soleimanmeigouni I, Ahmadi A, Nissen A, Xiao X. Prediction of railway track geometry defects : a case study. Struct Infrastruct Eng. 2019;0(0):1–15.
37. Taylor P, Caetano LF, Teixeira PF. Structure and Infrastructure Engineering : Maintenance , Management , Life-Cycle Design and Performance Optimisation

model to schedule railway track renewal operations : a life-cycle cost approach. 2014;(May 2015):37–41.

38. Faiz RB, SIngh S. Information analysis of rail track for predictive maintenance. WSEAS Trans Comput. 2009;8(7):1123–33.

39. Lyngby N. Railway track degradation: Shape and influencing factors. Int J Performability Eng. 2009;5(2):177–86.

40. Sadeghi J, Ph D, Heydari H, Doloei EA. Improvement of Railway Maintenance Approach by Developing a New Railway Condition Index. 2017;143(8):1–10.

41. Sadeghi J, Askarinejad H. Application of neural networks in evaluation of railway track quality condition. J Mech Sci Technol. 2012;26(1):113–22.

42. Khajehei H, Ahmadi A, Soleimanmeigouni I, Haddadzade M, Nissen A, Latifi Jebelli MJ. Prediction of track geometry degradation using artificial neural network: a case study. Int J Rail Transp. 2022;10(1):24–43.

43. Hitihamillage L. Tram Track Degradation Prediction. Sch Eng Coll Sci Eng Heal RMIT Univ. 2018;(June).

44. Karimpour M, Hitihamillage L, Elkhoury N, Moridpour S, Hesami R. Fuzzy Approach in Rail Track Degradation Prediction. 2018;2018.

45. Taylor P, Guler H. Structure and Infrastructure Engineering : Maintenance , Management , Life-Cycle Design and Performance Prediction of railway track geometry deterioration using artificial neural networks : a case study for Turkish state railways. 2013;(March):37–41.

46. Hyslip JP. Fractal analysis of track geometry data. Transp Res Rec. 2002;(1785):50–7.

47. Landgraf M, Hansmann F. Fractal analysis as an innovative approach for evaluating the condition of railway tracks. 2018;0(0):1–10.

48. Although I, Columbia B. A Review o f M e t h o d s Used to D e t e r m i n e the Fractal D i m e n s i o n o f Linear Features I Brian Kiinkenberg 2. 1994;26(1):23–46.

49. Boatwright J. Fractal Analysis Applied to Characteristic Segments 0 Small Ruler Example •.../ Large Ruler Example. 1987;92:331–44.

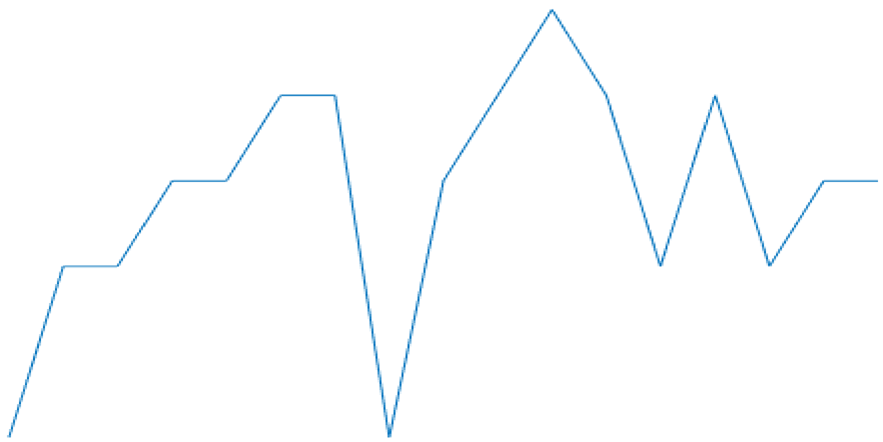
50. Maletsky E, Hösselbarth C, Peitgen HO, Perciante T, Jürgens H, Yunker L, et al. Fractals for the Classroom: Part One Introduction to Fractals and Chaos. Springer New York; 2012.

51. Gilbert LE. Are Topographic Data Sets Fractal ? 1989;131.
52. Microanalyse L De, Gray R De, Cedex B, Zucker SW. fractal of B. 1989;39(3).
53. Liang Z, Feng Z, Guangxiang XU. Procedia Engineering Comparison of Fractal Dimension Calculation Methods for Channel Bed Profiles. 2012;28(2011):252–7.
54. Tikkanen M, Heikkila R. The influence of clear felling on temperature and vegetation in an esker area at Lammi, southern Finland. Vol. 169, Fennia. 1991. 1–24 p.
55. Garson BGD. TESTING STATISTICAL ASSUMPTIONS. 2012;1–54.
56. Novales M, Orro A, Bugarín MR. Track geometry for light rail systems. Transp Res Rec. 2010;(2146):18–25.
57. Midi H, Sarkar SK, Rana S. Collinearity diagnostics of binary logistic regression model. J Interdiscip Math. 2010;13(3):253–67.

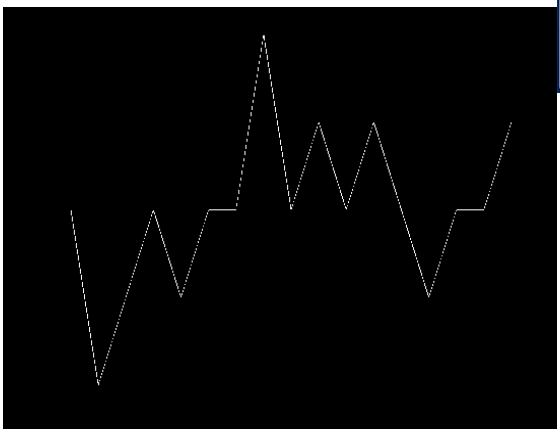
APPENDIX

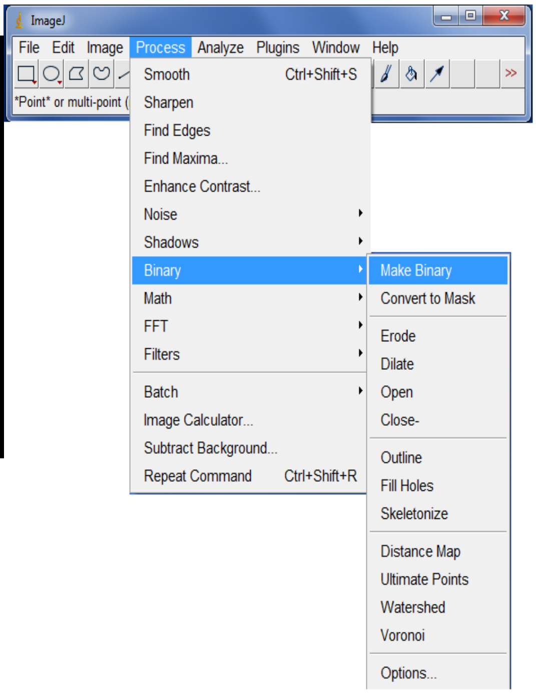
Appendix A: steps carried out to determine the fractal dimension of sections.

Step 1: The longitudinal level inspection data of each section are copied to excel and imported to mat lab software to plot the track longitudinal level. Hence the fraclac software works on black pixels on a white background, or white pixels on a black background to allow the software to only detect the profile graph the numeric ruler from the graphs should be in visible. The graph is made as shown below

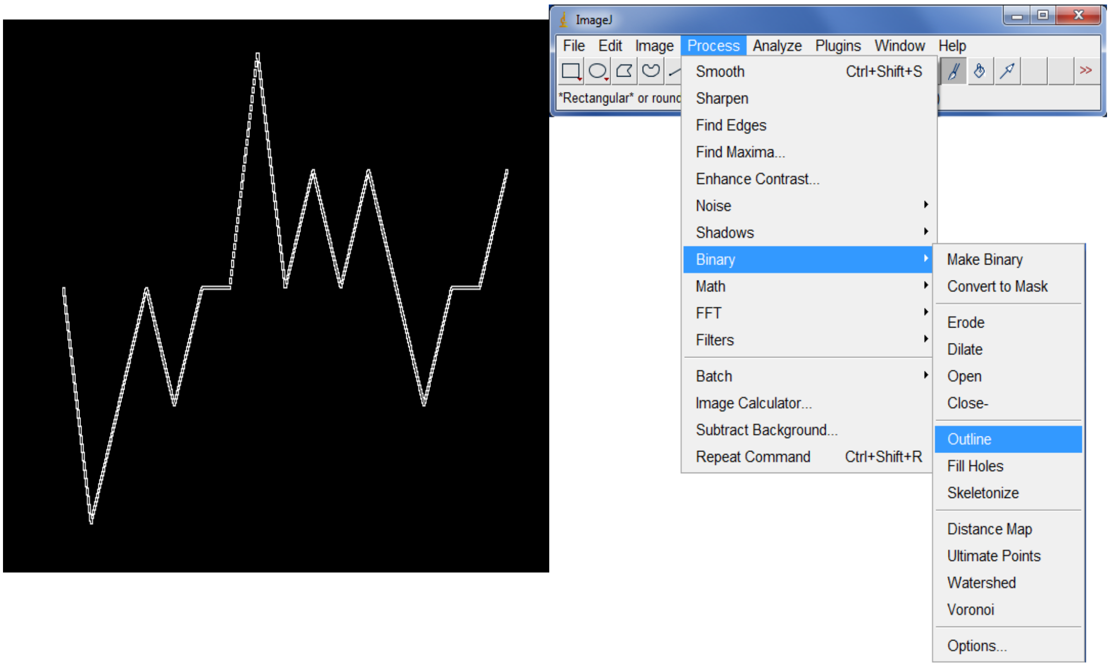


Step 2: In ImageJ click **File**, **import,** and import the image series of the sections
Step 3: On the ImageJ toolbar select process, go down to Binary, and select Analyze.

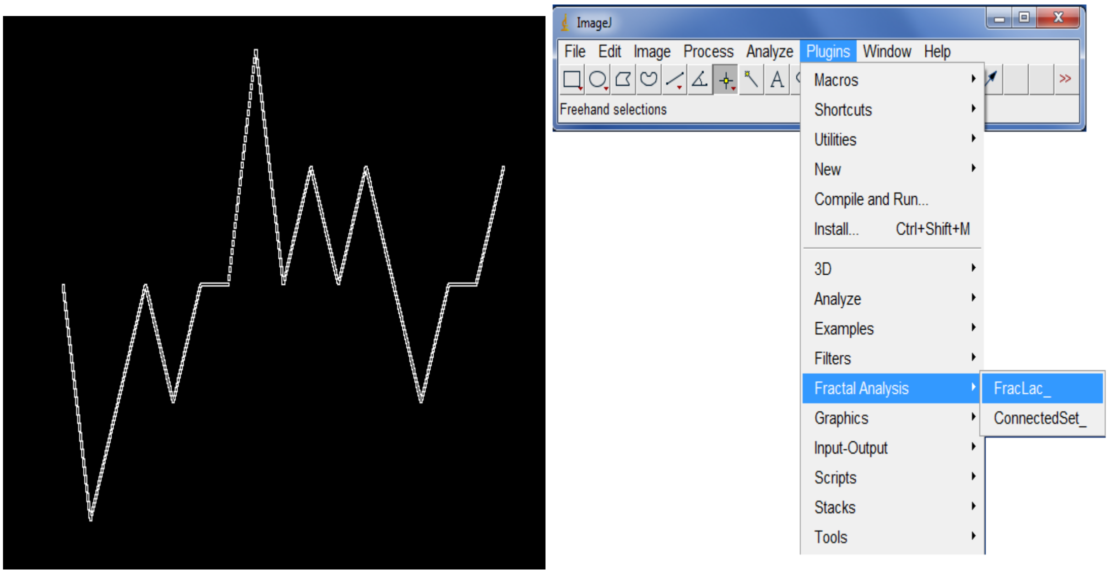




Step 4: Again, select Process on the toolbar, go down to Binary and now select outline to outline the image.

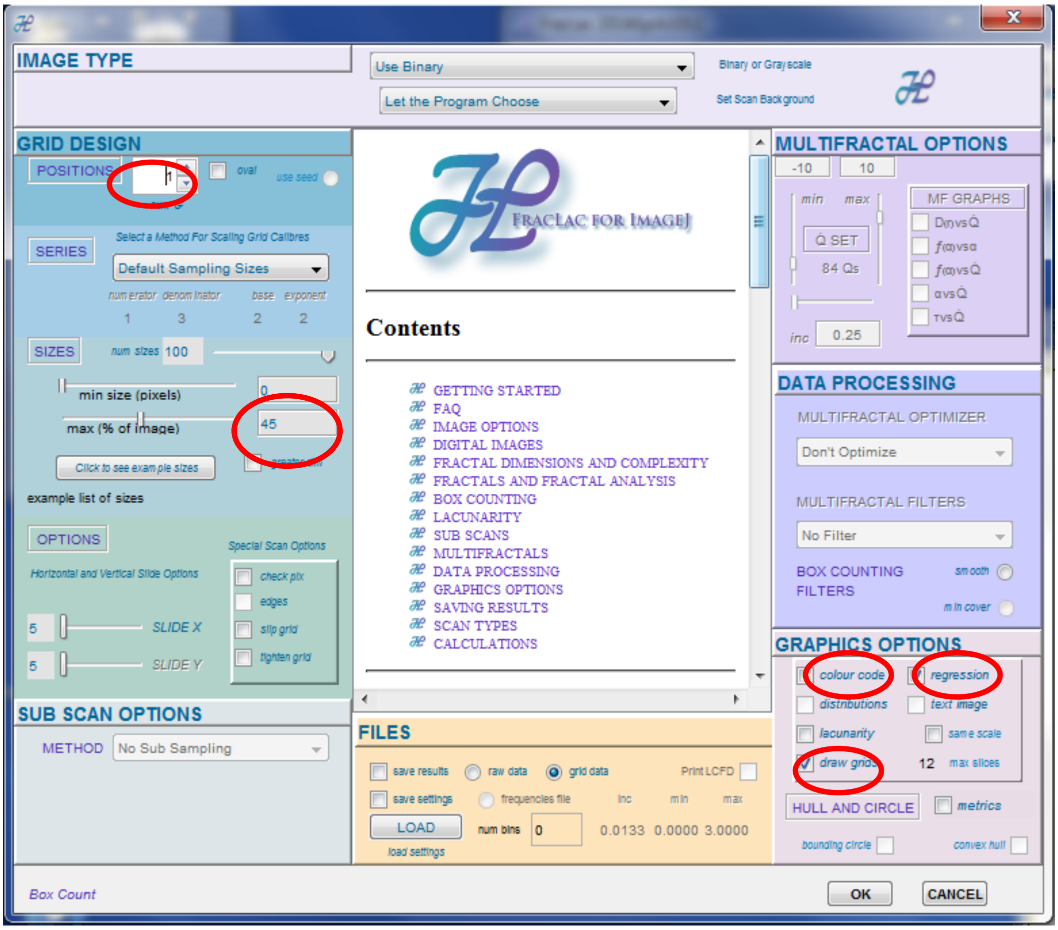


Step 5: The image is now ready to be analyzed. In the toolbar, select plugins, go down to fractal analysis, and open FracLac.

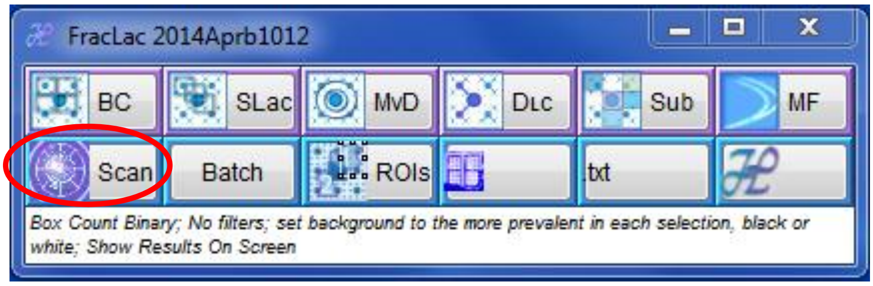


Step 6: Select “BC” to set up the box counting analysis.

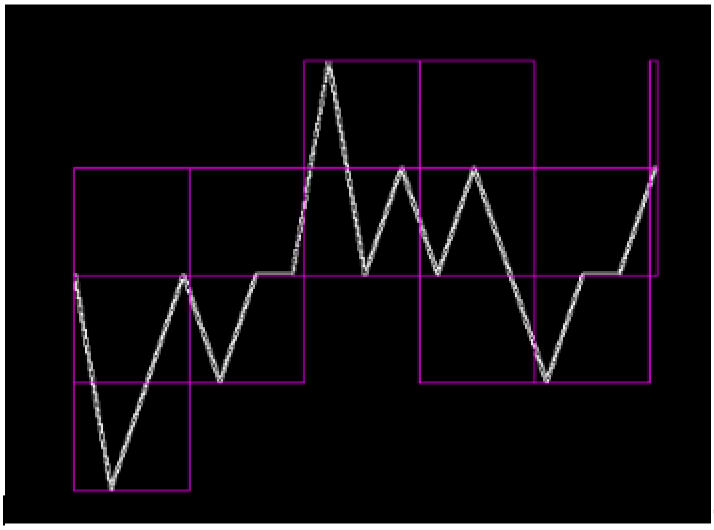
Step 7: In the FracLac analysis window shown below under GRID DESIGN set number of positions to 1 to count the number of boxes starting from one position. Under the series panel the minimum and maximum box size (ε) in pixels can be chosen by the user, in this study to consider the medium wave length range irregularity that characterizes the ballast quality box sizes of 3m to 25 m are chosen. When these lengths are measured on the image J software it becomes 16 to 130 pixel (25% of the image size). From GRAPHICS OPTIONS turn off "colour code" and turn on the "draw grids" and "regression" options. Then Click ok and now the program is set to scan.

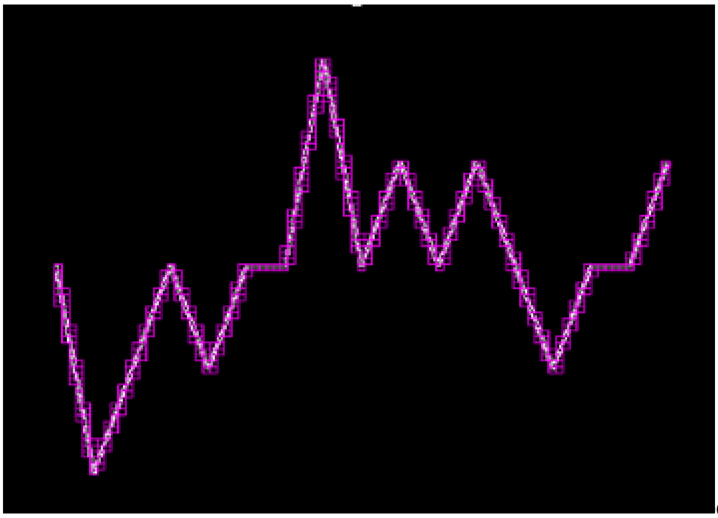


Step 8: Click the scan button on the FracLac toolbar.



Now the scan will open up several windows. The One contains the image with grid lines drawn. By using the scroll bar at the bottom of the window the program visualizes the different sample scales used in the analysis.





The other window that will appear contains a graph of size of sampling element, the samples of size containing the foreground pixels (profile graph) and the fractal dimension of the section (D_B) as shown below

FracLac 2015Sep090313a9330 Box Count Data Per Grid

File Edit Font

FracLac 2015Sep090313a9330: File Slice (ROI) start position	SCAN TYPE for DB	FRACTAL DIMENSION for DB
month1sec1png$1_ (0,0_560x420) 72,31--508,375	Box Count Binary; No filters; black; Show Results at G1	1.2609
Slopes for Data vs ε		
sampling size SIZE (G1)	ε (G1)	DB(counts) = slope ln(F) vs ln(ε): (G1)
1	1	-1.2609
DATA		
SIZE OF SAMPLING ELEMENT (G1)	ε = Sampling Element Size/Image Dimension (G1)	F = Samples of Sizeø Containing Foreground Pixels (F) Location (G1)
5	0.0115	473
6	0.0138	374
7	0.0161	316
8	0.0183	271
9	0.0206	242
10	0.0229	200
11	0.0252	183
12	0.0275	165
13	0.0298	153
14	0.0321	139
15	0.0344	131
16	0.0367	118
17	0.039	111
18	0.0413	103
19	0.0436	98
20	0.0459	86
21	0.0482	87
22	0.0505	80
23	0.0528	75
24	0.055	75
25	0.0573	72
26	0.0596	64
27	0.0619	69
28	0.0642	66
29	0.0665	59
30	0.0688	59

In this manner fractal dimension of each sections are calculated and this fractal dimension represents the ballast quality index.

Appendix B: Sample data for Fractal analysis

	19-Oct		19-Nov		19-Dec		20-Jan	
sec1	x	y	x	y	x	y	x	y
	0	1	0	2	0	0	0	0
	5	1	5	0	5	0	5	2
	10	0	10	1	10	0	10	0
	15	1	15	0	15	1	15	1
	20	0	20	1	20	1	20	0
	25	2	25	0	25	1	25	0
	30	1	30	2	30	0	30	1
	35	2	35	0	35	1	35	2
	40	2	40	2	40	0	40	2
	45	0	45	0	45	1	45	1
	50	1	50	2	50	1	50	1
	55	3	55	2	55	2	55	3
	60	2	60	0	60	-1	60	1
	65	0	65	0	65	0	65	0
	70	1	70	0	70	2	70	0
	75	2	75	0	75	0	75	3
	80	2	80	2	80	2	80	2
sec2	x	y	x	y	x	y	x	y
	0	0	0	1	0	-1	0	0
	5	2	5	-1	5	1	5	1
	10	1	10	0	10	1	10	1
	15	2	15	0	15	1	15	2
	20	2	20	-1	20	0	20	1
	25	3	25	0	25	1	25	2
	30	3	30	-1	30	0	30	2
	35	1	35	0	35	0	35	0
	40	2	40	0	40	-1	40	1
	45	2	45	0	45	0	45	1
	50	5	50	1	50	1	50	3
	55	1	55	1	55	1	55	0
	60	2	60	0	60	1	60	0
	65	2	65	0	65	2	65	1
	70	3	70	1	70	0	70	2
	75	1	75	-1	75	2	75	4
	80	5	80	3	80	2	80	4

Appendix C: Sample input data used for regression modelling

Sample input data for straight section modelling

BOQ	Tonnage	Gradient	Speed	Number of past Tamping	subgrade	season
1.5299	0.10767	4	17	0	1	3
1.3541	0.2217	17.5	25	1	1	2
1.1932	0.44346	17.4	40	0	0	2
1.28881	0.44346	10	53	1	1	2
1.2144	0.55509	17	56	2	1	2
1.196	0.65827	4	53	0	1	1
1.1208	0.73531	3	38	0	1	1
1.2348	0.76849	2	45	1	0	1
1.1156	0.80816	24.5	39	1	1	1
1.1887	0.88833	24	53	1	1	3
1.0871	0.84666	55	32	0	1	3
1.151	0.88833	56	45	0	1	3
1.0643	0.98695	57	40	1	1	3
1.1229	1.08559	59	31	1	1	1
1.1162	0.84666	55	56	0	1	3
1.0329	0.88833	56	53	1	1	3
0.9943	0.98695	57	53	2	1	3
1.0263	1.08559	59	37	3	1	1
1.0794	0.84666	55	40	0	1	3
1.0516	0.88833	56	43	1	1	3
1.0591	0.98695	57	47	2	1	3
1.033	1.08559	59	47	3	1	1
1.1198	0.84666	10	47	0	1	3
1.1276	0.88833	11	47	0	1	3
1.107	0.98695	12	14	0	1	3
1.1335	1.03024	13	15	0	1	1
1.1349	1.08559	14	38	0	1	1
1.1472	0.84666	12	40	0	1	3
1.15	0.88833	13	40	0	1	3
1.1165	0.98695	14	27	0	1	3
1.106	1.03024	15	23	0	1	1
1.111	1.08559	16	39	0	1	1
1.0932	0.84666	2	40	0	1	3
1.1249	0.88833	2	40	0	1	3
1.1184	0.98695	2	40	0	1	3
1.1729	1.03024	2	31	0	1	1
1.074	1.08559	2	56	0	1	1
1.0628	0.9296	23	57	1	1	2

Sample input data for curved section modelling

BOQ	Tonnage	Gradient	Radius	Speed	Number of past Tamping	subgrade	season
1.3151	0.10767	15.5	2500	43	0	0	3
1.3332	0.2217	4	470	38	0	1	2
1.243	0.33082	3	1000	45	0	0	2
1.235	0.44346	11	1304	37	0	1	2
1.1642	0.55509	32	350	27	1	0	2
1.1873	0.65827	55	190	40	1	1	1
1.2956	0.73531	3	1000	40	0	0	1
1.1872	0.76849	8.5	240	14	0	1	1
1.1614	0.80816	9	1200	50	0	0	1
1.1586	0.84666	9	1200	50	0	0	3
1.1392	0.88833	32	350	28	2	0	3
1.1503	0.55509	4.5	190	40	0	1	2
1.204	0.65827	4.5	190	43	0	1	1
1.1494	0.73531	4.5	190	47	0	1	1
1.1478	0.76849	4.5	190	47	0	1	1
1.1116	0.80816	4.5	190	47	1	1	1
1.1156	0.55509	17.5	190	47	0	1	2
1.1358	0.65827	17.5	190	14	0	1	2
1.0721	0.73531	17.5	190	15	0	1	1
1.1198	0.76849	17.5	190	38	3	1	1
1.0807	0.80816	17.5	190	40	1	1	1
1.151	0.55509	55	190	40	0	1	2
1.1705	0.65827	55	190	27	0	1	1
1.1083	0.73531	55	190	23	2	1	1
1.0764	0.65827	8.5	240	40	0	1	1
1.0514	0.73531	8.5	240	43	1	1	2
1.0775	0.76849	8.5	240	47	1	1	1
1.024	0.80816	8.5	240	47	1	1	1
1.131	0.55509	32	1400	47	0	0	2
1.1268	0.65827	32	1400	28	0	0	1
1.1518	0.73531	32	1400	50	0	0	1
1.1081	0.76849	32	1400	50	0	0	1
1.1198	0.80816	32	1400	15	2	0	1
1.235	0.55509	11	1304	38	3	1	2
1.2167	0.65827	11	1304	40	0	1	1
1.1578	0.76849	3	1000	40	0	0	1
1.1216	0.80816	3	1000	43	0	0	1
1.009	0.9296	48.25	1304	38	0	10	2