ADDIS ABABA UNIVERSITY

ADDIS ABABA INSTITUTE OF TECHNOLOGY

AFRICAN RAILWAY CENTER OF EXCELLENCE

ANALYSIS AND PREDICTION OF TRACK GEOMETRY DEGRADATION

CASE STUDY OF ADDIS ABABA-LIGHT RAIL TRANSIT

A Thesis in Railway Engineering (Civil Infrastructure)

By

Aisha Kabange

Advisor: Dr. Bikila Teklu Wodajo

Co Advisor: Mr. Daniel Haile Michael

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)

June, 2021

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

THE BOARD OF EXAMINERS

Dr. Bikila Teklu Wodajo	______________________	________________
Advisor	Signature	Date
Mr. Daniel Haile Michael	______________________	________________
Co Advisor	Signature	Date
Mr. Zewdie Moges	______________________	________________
Internal Examiner	Signature	Date
Dr. Getu Segni Tulu	______________________	________________
External Examiner	Signature	Date
Dr. Abraham Gebre	______________________	________________
Chairperson of Department	Signature	Date

DECLARATION

I hereby declare that the work which is being presented in this thesis entitled "**Analysis and Prediction of Track Geometry Degradation**" is original work of my own, 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.

____________________ _______________

Aisha Kabange Date

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

____________________ _______________

Dr. Bikila Teklu Wodajo (Advisor) Date

____________________ _______________

Mr. Daniel Haile Michael (Co Advisor) Date

ACKNOWLEDGEMENTS

Foremost, I would like to offer my sincere appreciation to my advisor Dr. Bikila Teklu Wodajo and co-advisor Mr. Daniel Haile Michael for their constructive comments and suggestions throughout this study. I would also like to express my gratitude to Mr. Zerihun Terefe and Mr. Tadewos Worku from Civil Infrastructure department of Addis Ababa Light Rail Transit for providing me with necessary data for my work. My deepest gratitude goes to my parents and my siblings for their continue encouragement throughout my studies. To my husband thank you for continued support and weathering love. Lastly, to my classmates their advice and on-time information regarding the thesis are appreciated and their friend ship over the last years.

Abstract

Development of good maintenance policy requires better understanding of the long-term behavior of railway track systems. Addis Ababa Light Rail Transit is new system which has been in operation for the last five years. Understanding the degradation pattern of such system will help in scheduling and controlling maintenance activities, which is a problem to many railway operating entities. Track safety and maintenance effectiveness are improved by clearly understanding future condition of the track.

This paper analyses and predicts the track geometry degradation of Addis Ababa Light Rail Transit. Track geometry degradation models developed by different researchers were critically reviewed to determine suitable model for AALRT. Factors which contribute to track geometry degradation were analyzed to determine the relationship between the factors and track degradation. Factors such as Number of Trips, Curve Radius, Speed, Grade, Tonnage and Track Surface were analyzed. All analyzed factors except Track Surface were found to have influence on rate of degradation for curved and straight sections of AALRT. The selected influencing variables and longitudinal leveling changes were used to develop track degradation model. Multiple regression model was developed in SPSS while Artificial Neural Network model (ANN) was developed on MATLAB software. The performances of two models were evaluated for straight and curved sections of North South line.

North South line is a double line, model development was performed per each direction (Uplink line and Downlink line) and whole line regardless of direction. Six multiple regression models were developed. For straight sections three models were developed in which Downlink line displayed tremendous performance (R^2=98%) compared to Uplink line (R^2=85%) and North South line (R^2=88%). For curved sections performance of 91% was observed on Uplink line and Downlink line while North South line had performance of 88%. Six ANN models were also developed as a complex model to predict future condition of track. Performance of 97% R-squared was obtained for both curved and straight sections of Uplink line while performance of 96% was observed in straight section and 93% performance was observed in curved sections of Downlink line. North South line display performance of more than 90% for both straight and curved sections of line. Comparison of two models showed that both can predict the longitudinal levelling changes with very low error. Although the ANN model has better performance than Regression model.

Keywords: Track Geometry Degradation, Regression model, ANN model, AALRT

TABLE OF CONTENTS

List of Tables

List of Figures

Glossary of Abbreviations

AALRT	Addis Ababa Light Rail Transit
MGT	Million Gross Tonnes
TQI	Track Quality Index
USA	United States of America
ERC	Ethiopia Railway Cooperation
EMGT	Equivalent Million Gross Tonnes
M&R	Maintenance and Renew
SDHA	Standard deviation of Horizontal Alignment
SDLL	Standard Deviation of Longitudinal Leveling
N-S Line	North South Line
E-W Line	East West Line
LRT	Light Rail Transit
ANOVA	One Way Analysis of Variance
HBM	Hierarchical Bayesian Models
ANN	Artificial Neural Network
C.R	Curve Radius

CHAPTER ONE

1.1. Introduction

Public transport aims at transportation of goods and people from one area to another in safe and reliable manner. Increase in population causes increase in transportation demand which required the service providers to expand existing networks to meet the demands. In order to alleviate the increasing traffic pressure, the Ethiopian Government (Transport authority) has introduced Light Rail Transit around 2014 to help in transporting passenger to their destination.

Light rail transit systems are becoming increasingly important as a solution to the problem of metropolitan public transport in medium- sized cities and as feeders for high-capacity modes in bigger cities. Light rail transit systems have great flexibility for running over different types of streets and layouts, because they can adapt to very strict conditions. Light rail systems are characterized by their type of right-of-way, which suppose that they are longitudinally separated from other traffic by curbs, barriers, grade separation, and other physical means but with at grade crossings for vehicles and pedestrians, including regular street intersections.(Novales, Orro and Bugarín, 1887)

Transportation infrastructures such as road pavements and rail tracks, although most of the time deterioration process is very slow, but it might lead to massive failures with an enormous financial lose. As a result it is very important to decide when and how to perform maintenance operations for such systems and how to allocate the resources (manpower, materials, machines and funds) to the parts of the system with highest need.(Yousefikia *et al.*, 2014)

Detection and rectification of track defects are major issues in the railway industry. The defects can be categorized into one of two groups: track structural defects and track geometry defects. Track structural defects include the condition of the rail, sleeper, fastening systems, subgrade and drainage systems generated from the structural conditions of the track. On the other hand, track geometry defects indicate severe ill-conditioned geometry parameters such as profile, alignment, gage etc. (Sadeghi and Askarinejad, 2010)

On a paper titled Modeling of track geometry deterioration presented by authors (Guler, Jovanovic and Evren) a curve showing the hypothetical deterioration of the track geometry was developed. On this curve the track goes through three phases namely youth phase, intermediate phase and old phase. The youth phase occurs during the early life of track which is after completion of construction or completion of major renew work. The duration of this stage is quite short and considerable deterioration is observed due to initial settlement of track (Figure 1.1, marked with (a)). This stage is highly unpredictable and differ from one track section to the other which makes it difficult to model.

The intermediate phase occurs during majority of track life-time and characterized by linear deterioration pattern. Linear deterioration pattern is depicted by the sufficient stabilized condition of track (Figure 1.1, marked with (b)). Quasi-exponential deterioration patter is observed in the third phase (Figure 1.1, marked with (c)). This phase is characterized by increasingly rapid deterioration due to old condition of track. Maintenance and Renew works are applied earlier before maintenance threshold (horizontal black line) to prevent this phase to occur. Safety of traffic and reliability of the track are affected on this phase if proper M&R activities are not employed. . (Guler, Jovanovic and Evren, 2011)

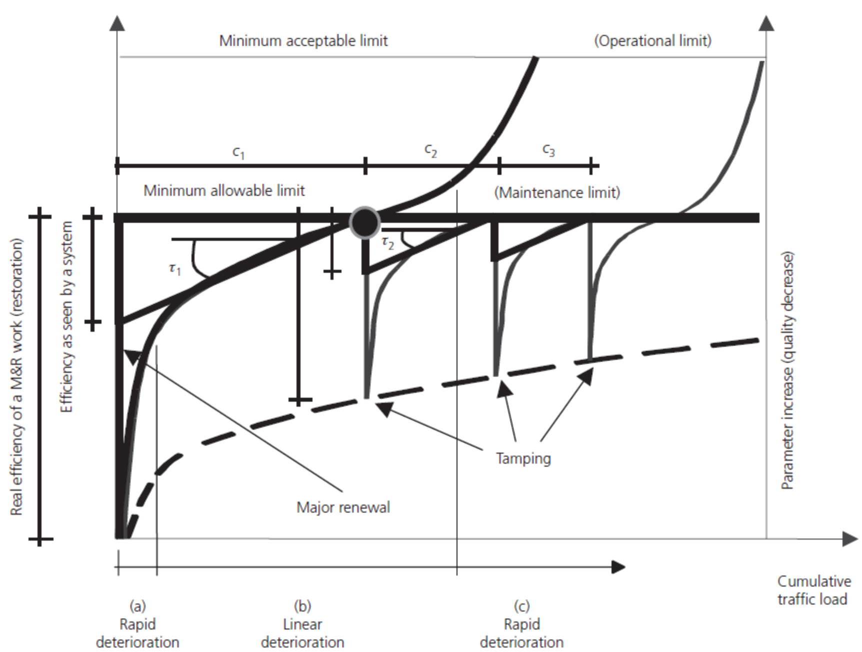

Figure 1.1: The hypothetical track geometry deterioration curve.(Guler, Jovanovic and Evren, 2011)

Hence a reliable maintenance plan, an accurate inspection program is needed. In this way, track geometry condition can be evaluated, and the failures and risks can be mitigated. An optimal inspection plan creates room for timely detection track geometry defects and for planning maintenance activities.

Therefore, there is a great need to formulate an effective inspection and maintenance plan for AALRT which will identify the defects in early stages and rectification measures can be performed accordingly and at a cost-effective means. This can be achieved through development of degradation model of track geometry in which the failure mechanism of track will be understood and appropriate and cost effect measures can be employed.

1.2. Problem Statement

Railroad track geometry, the three-dimensional layout and positioning of the track components, is an important indicator of safety of tracks. Geometric deviations, characterized by displacements in the range of millimeters, can significantly decrease the safety and reliability of the infrastructure. Because degradation can occur across multiple components, there is a need to understand how the different track components, including ballast and crossties, interact with one another. Effective maintenance can be employed to compensate for the shortcomings of railway track functionality and reliability. The International Union of Railways (UIC, 2006) states that, costs of permanent way and its maintenance and renewal (M&R) are substantial and form a large part of total infrastructure expenditure. Any reduction in these costs has a significant impact on the overall efficiency of the management of infrastructure. This study was done as a stepping stone towards optimization of maintenance of new railway system as Addis Ababa Light Rail Transit.

Ensuring the safety of railway traffic and the comfort of passengers are important management tasks for railway transportation enterprises, and they are greatly affected by track irregularities (or track geometry defects). Maintaining railway tracks in satisfactory regularity conditions requires a large amount of maintenance and replacement activities, which affect the operational efficiency of rail transport. This problem can be solved if managers accurately understand the changes in the track conditions over time, which allows them to develop scientific and rational maintenance plans.

In order to meet the required safety and comfort, high track quality is required. If track quality is poor, it may directly or indirectly lead to such issues like speed reduction, derailment, high maintenance costs as well as increased degradation rates. To avoid these issues, maintenance optimization is essential, which can only be achieved by clearly understanding the degradation behavior of the track.

Manual inspections are used in AALRT to identify defects, level and time period for maintenance of the damaged track. Inspection and maintenance activities require large number of people working in groups hence human errors in inspection and prediction of maintenance time frame are inevitable. Appropriate maintenance strategies will help to attain RAMS (reliability, accessibility, maintainability and safety) on a new infrastructure such as AALRT and in turn lowering the Life Cycle Cost. As one of the solutions, this thesis will base on the development of degradation model of AALRT. This will help the maintenance of degraded track to be identified within specific time period. Furthermore, there will be reduction in maintenance cost, save time and prevent unnecessary maintenance activities.

1.3. Objectives

The main objective of the research is carry out analysis and to develop track geometry degradation model for the straight and curved section of Light Rail Transit. And more specifically the study is intended to meet the following specific objectives:

- To study and evaluate the factors influencing the track geometry degradation of the straight and curved section of Light Rail Transit.
- To develop the track geometry degradation model for straight and curved section of LRT as a function of the influencing variables.

1.4. Research Framework

This research covers the following major tasks: Chapter 2 involves reviews of past studies on the degradation models, their categorization, and limitation of each category and the identification of the gap present. Chapter 3 involves the information on data collection techniques, research framework and procedures followed on development of degradation models. Chapter 4 shows results of analysis of factors influencing the track geometry degradation model specifically for the AALRT using SPSS further results of regression analysis and Artificial Neural Network (ANN) models with consideration of the variables affecting railroad tracks are displayed. Conclusions and Recommendations followed by future works for broadening of this research are explained in Chapter 5.

1.5. Scope

This research covers different aspects of railway track geometry degradation. In this research study, longitudinal level is used as track quality index for degradation modeling. Therefore, other track geometry parameters, i.e., alignment, cant, twist, and gauge are left out of the scope of this study. In addition, the effects of environmental condition and maintenance history on track geometry degradation are not considered in this study. The study will only focus on the straight and curved section of North-South line of Addis Ababa Light Rail Transit.

CHAPTER TWO

LITERATURE REVIEW

2.1. General

This chapter gives an overview of the railway structure and detailed review of degradation models developed by different researchers. The advantages and limitation of each model will be evaluated. The aim of chapter is to obtain the best methodology for this study by determining the gap and limitation of existing studies.

2.2 Railway Structure

The rail tracks are of two types namely ballasted track and concrete slab track. Ballast track can be classified in two Super structure (Rail, Sleeper, Rail pads, Fastening and Top Ballast) and Sub structure (Bottom Ballast, Sub ballast, Fill material and subgrade). Rail are mounted on sleeper which can be wooden, steel or concrete and main function is to transfer load to the sleepers. Sleepers are laid on ballast which provides lateral and longitudinal stability. Ballast also acts as shock absorber by reducing the vibrations. Under load geometry deviations of track in range of millimeters occurs so routine maintenance (tamping) is applied to restore the appropriate condition,. In addition to this ballasted track is a massive structure that makes it impossible to be used in tunnels, subway and especially in the urban roads as the tram track.(Yousefikia *et al.*, 2014)

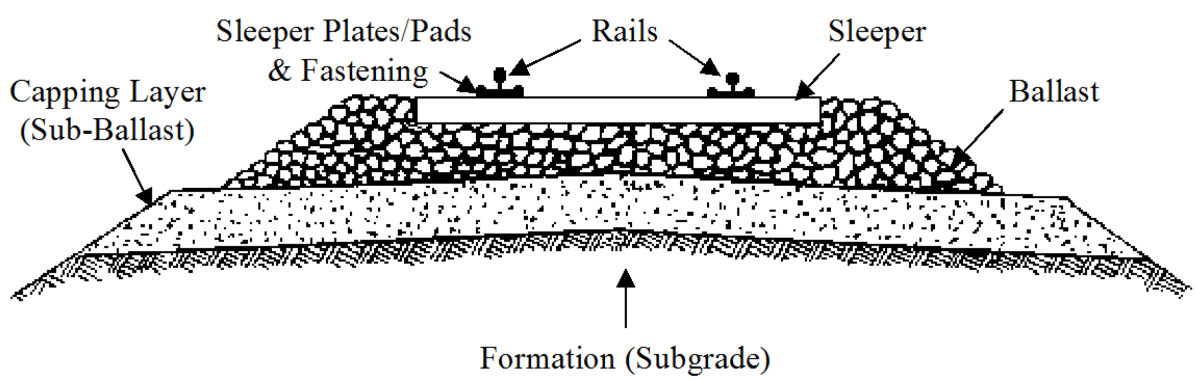

Figure 2.1: Cross-section of typical railway track.(Hawari, 2007)

2.3 Railway Track Geometry

The track geometry describes the position that each rail, or the track center line, occupies in space. During operation of railway track defects across track structure and track geometry are inevitable. Regular inspections are done either by qualified personal or track recording car to identify such defects. Sections with defects are rectified through scheduled maintenance.

Changes in the geometry are caused by wheel/rail interactions, which result in various types of track irregularities. Track irregularities are a major source of disturbance in wheel/rail systems and are the main cause of vibrations and wheel/rail dynamic forces. Importantly, these irregularities are an embodiment of the overall performance of the track structure. They significantly affect the operational safety and passenger comfort in trains.(Bai *et al.*, 2016)

Track geometry measures can be divided into five classes:

(1) Longitudinal level/Profile is the measure of non- uniformity of top surface of the rail measured in a short distance. The vertical deviation of this center line is termed as longitudinal level/ Profile defect.

(2) Alignment is the track geometry of the track center line projected onto the longitudinal horizontal plane. The horizontal deviation of this center line is termed as alignment defect.

(3) Gauge is the distance between the inner sides of the rail heads. Gauge defect is the change in distance between the inner rail heads.

(4) Cant or (cross-level) is the difference in elevation of top surfaces of the two rails measured at specific point. The extra difference in, compared to the designed value is termed as Cross Level defect.

(5) Twist is the difference between two cross-levels taken at a defined distance apart. The extra difference between two cross-level measurements a certain distance apart, compared to the designed value is termed as twist defect. (American Railway Engineering and Maintenance of Way Association, 2006)

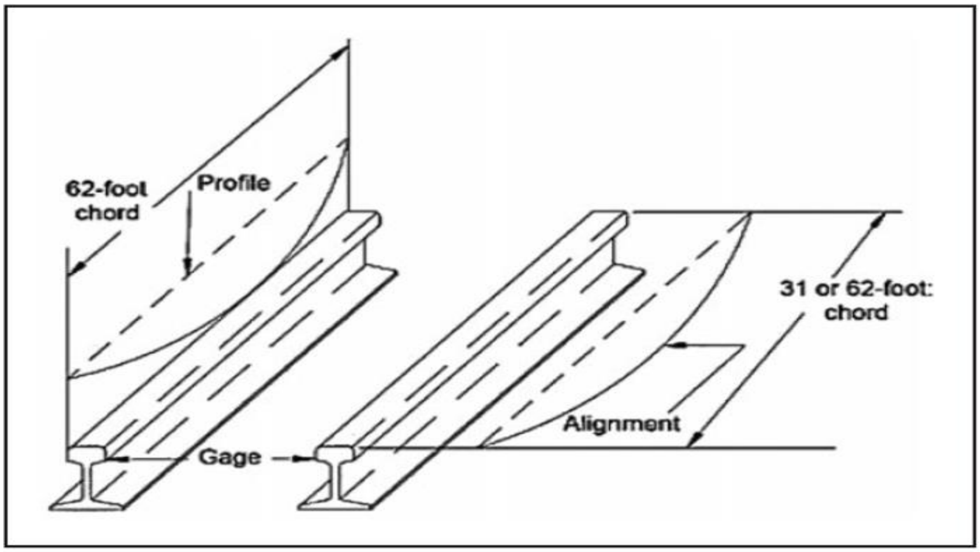

Figure 2.2: Schematic of Track Geometry and Type of defects.(Higgins and Liu, 2018)

2.4 Railway Track Degradation Modeling

Infrastructure deterioration models are mathematical relationships between a dependent variable, namely deterioration or change in condition, and a set of casual variables, including design attributes, traffic loading, environmental factors, age and maintenance history. Over the years a number of studies have been done on track geometry degradation. Railway industry is now focused on improving capacity of railway lines rather than construction of new lines. Such process requires understanding the degradation pattern of track and finally development of effective maintenance strategy. Degradation of rail track has many contributing factors including ill condition of track structures (*i.e*., sleepers, fastenings, drainage, subgrade and ballast), defects of track geometry parameters (*i.e*. rail profile, rail track elevation, rail track super-elevation, track curvature, alignment and cant) and operation of track including age of rails and axle load, speed, traffic density, traffic type, rail-wheel interaction, Million Gross Tonnes (MGTs) and rail track construction, and rail welding and rail lubrication.(Sadeghi and Askarinejad, 2010; Yousefikia *et al.*, 2014; He *et al.*, 2015)

Daily operation of track with passage of different vehicles with varying loading causes the deviation of track geometry from standard values . Environmental causes (i.e., snowfall, rainfall), and condition of wheel/ rail interaction also can cause the arise of defects on track geometry. Regular inspections will allow identification of defects early and schedule maintenance accordingly. (Guler, Jovanovic and Evren, 2011)

The railway track deteriorates very slowly and exhibits the bath curve failure characteristics. Figure 2.3 shows the stages any structure follows through its life time.

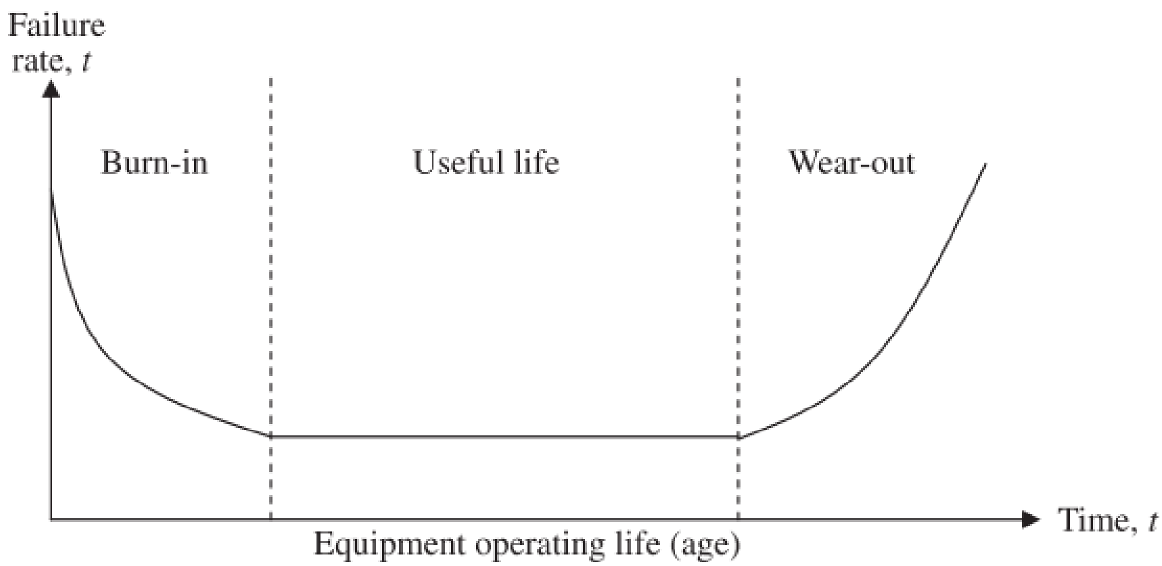

Figure 2.3: Bath tub Curve.(Ahmad and Kamaruddin, 2012)

The Bath tub Curve divides the infrastructure into three phases namely Burn-in, Useful life and Wear- out. During the Burn- in phase rapid failure occurs caused by number of reasons such as initial settlement of structure, poor material quality used, improper construction techniques, human error, etc. Useful life phase occurs during most of life of infrastructure at this stage linear or constant hazard rate is observed also failures occurrence can be easily predicted at this stage. Wear- out phase is the last stage of infrastructure life. At this stage high failure is also observed caused by number of reasons such as aging associated with failures such as (cracking, corrosion and creep) and poor maintenance practice. It is at this stage renewal of particular part of infrastructure is required .(Yousefikia *et al.*, 2014)

2.4.1 Classification of Track Degradation Models

The degradation of the track geometry varies along the track line, due to many structural and environmental factors that exert a different influence on the railway track. Researchers have developed different degradation models based on influencing factors such as vehicle characteristics, operation characteristics (i.e., speed, tonnage etc.), track characteristics such as curve radius, track surface, rail type and size and Environmental factors (i.e., mountains terrain, snowfall etc.). Accordingly, rail degradation models can be classified into four general approaches, as shown below.(Soleimanmeigouni *et al.*, 2018)

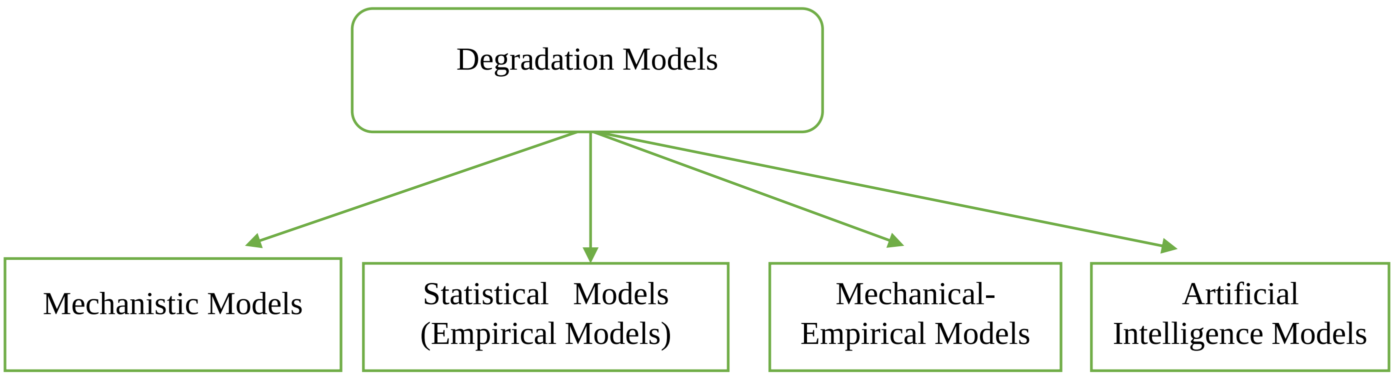

Figure 2.4: Classification of Track degradation Models(Elkhoury, 2018)

2.4.1.1 Mechanistic Models

Mechanistic approach uses fundamental theories of infrastructure behavior for modeling. This model tends to focus on interaction between wheel and rail. it relies on mechanical properties of track parameters which are difficult to quantify and very different from place to place.

Various researchers have used this model to describe and predict the track degradation for better understanding of behavior of track in turn to optimize the maintenance activities.

Japanese study on deterioration of ballasted track conducted by Sato (1995) who conducted a study of ballast settlement on a heavy haul narrow gauge track, developed the following formula for estimation of track settlement;

$$y = \gamma(1 - e^{-\alpha x}) + \beta x \tag{2.1}$$

Where:

x represents Tonnage passing through the track.

γ is initial rapid settlement. It is cause by rearranging of ballast particles.

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

β is lateral movement of ballast particles under sleeper.

Sato investigated the growth of track irregularity in 100 days by evaluating the 25m wavelength longitudinal level defects recorded by geometry car. He found that there is relationship between average growth of track irregularity and influencing variables such as Tonnage, Speed, Joint or Continuously welded rail, Structure factor and quality of subgrade as shown in formula 2.2.

$$S = 2.09 \times 10^{-3} \cdot T^{0.31} V^{0.98} M^{1.10} L^{0.21} P^{0.26} \tag{2.2}$$

Where:

S is the average growth of track irregularity in a track section (mm/100 days),

T is the passing tonnage (million tons/year),

V is the average running speed (km/h),

M is the structure factor,

L is the influence factor for jointed rail or continuous welded rail (CWR) and

P is a factor related to the subgrade conditions.

The Structure factor M has to be calculated and it depends on sleeper pressure, ballast acceleration due to wheel impact and impact coefficient. The conclusions draw from this study are when joint rail is used irregularity growth will be 1.62 times to that on CWR and a bad subgrade impacts the irregularity growth by 1.82 times to that of good subgrade. (Sato, 1995)

The track rate of settlement due to vehicle passing through a dipped joint was calculated from experiments conducted in the controlled condition laboratory by the University of Munich. The ballast pressure is multiplied by the log of the number of axles passes as follows:

$$S = a \times p \times \ln \Delta N + b \times p^{1.21} \times \ln N \quad (2.3)$$

Where:

S represents the rate of settlement of track,

p represents fast settlement just after a maintenance action. It is the ballast pressure and is calculated using Zimmermann method,

N in the second part should express the total number of passing axles,

ΔN expresses a pre-loading period comprising the first passing axles. ΔN should be ≤ 10000,

The parameters a and b are constants suggested to be in the value range; 1.57–2.33 (a) and 3.04–15.2 (b).(Iwnicki S, 2000)

Primary investigations to understand the fundamentals of deterioration mechanism of railroad track were carried out by the Office for Research and Experiments (ORE) of the International Union of Railways (UIC) in the 1980s.The ORE examined . The ORE examined the inspection data from different entities and a mechanical model was developed. This model discovered there is relationship between deterioration with initial deterioration after tamping, traffic volume, dynamic axle load and speed. The model has been analyzed on the data obtained from American and Indian Railways respectively. The relationship reads:

$$e = e_0 + hT^{\alpha}(2Q)^{\beta}v^{\gamma} \quad (2.4)$$

Where:

e_0 is deterioration directly after tamping,

T is traffic volume,

2Q is dynamic axle load,

v is speed,

h is a constant and

α, β, and γ are the variables estimated from laboratory experiments.

2.4.1.2 Statistical (Empirical) Models

The statistical approach involves the analysis of many observations of track measurements collected for a long period and corresponding casual parameters including traffic, track components, environmental and maintenance variables. When actual observations are used for modeling, more realistic prediction of track condition is expected to be given. Aim of statistical based degradation models is to find a general pattern for the statistical distribution of the track geometry using inspected data of track condition. These models generates sample data from large population using the statistical assumptions. The statistical models can further be divided into three categories namely; Deterministic models, probabilistic models and stochastic model.

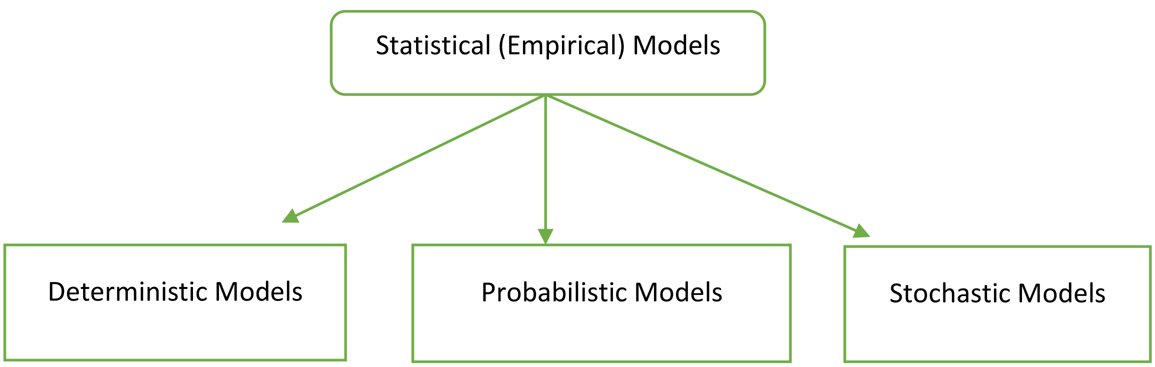

Figure 2.5: Classification of Statistical degradation Models(Elkhoury, 2018)

i. Deterministic Models

Deterministic models are statistical models which the same conditions or state develops same results meaning randomness is not involved in degradation modeling predictions. Large number of data is required in model construction which provide definite relationship between track geometry (dependent variable) and other casual parameters (track characteristics, vehicle characteristics and operation). Linear and exponential forms of deterministic models were the first attempt in rail degradation modeling, due to their simplicity of mathematical expression and their ability to show a direct relationship between the input and output variables.

M Audley and JD Andrews analyzed the effect of tamping on track geometry degradation for UK rail network. Short wave length (35m) measured Vertical alignment was considered as track quality indicator for this study. The track degradation is estimated with consideration of the effect of line speed and the maintenance history. The track geometry inspection data, line speed and maintenance history of UK railway for six years was used to generate the degradation model for this study. A linear function was used as it was observed to provide good fit for the model. The Weibull distribution was selected by the authors due to widely applicability, ability to deal with small sample size and its good performance. The authors concluded that as the number of cycles of tamping increase the rate of track deterioration also increases. The more maintenance will be required on a high-speed line than slow line.(Audley and Andrews, 2013)

The Multivariate regression analysis was used by Guler (2011) to develop the track degradation model for Turkish railway. The track geometry defects every 6 months, for 7 years period was collected from the Turkish State Railway. Data involving track characteristics, environmental conditions, maintenance and renew activities were collected and applied in development of track degradation model. The track line of 180km was divided into 820 nearly equal analytical segments in which information on track structure, traffic characteristics, maintenance & renewal history, track measurement records, environmental factors and track layout were collected. The model found environmental effect such as land slide and snow have no correlation with degradation rate while flooding and falling rock increases the deterioration rate. Through analysis of scatter plots and correlation matrices it was observed that curvature and gradient have positive correlation with deterioration rate.

Although cant showed effect on deterioration rate was excluded from modeling due to multi collinearity with curvature. Sleeper type showed mixed results in which there is positive correlation with vertical geometry parameters (twist, cant and level) while negative correlation with horizontal geometry parameters (gauge and alignment) was observed. Rail type and Rail length also had effect on the deterioration rate. Use of CWR's tend to reduce the rate of deterioration . Model accuracy was questioned when cumulative load and Speed showed negative correlation with deterioration rate. The following linear regression model that predicted the deterioration rate ($\bar{\tau}_i$) was expressed by the following:

$$\begin{aligned}\bar{\tau}_i = {} & \alpha_{i1}x_1 + \alpha_{i2}x_2 + \alpha_{i3}x_3 + \alpha_{i4}x_4 + \alpha_{i5}x_5 \\ & +\alpha_{i6}x_6 + \alpha_{i7}x_7 + \alpha_{i8}x_8 + \alpha_{i9}x_9 \\ & +\alpha_{i10}x_{10} + \alpha_{i11}x_{11} + \alpha_{i12}x_{12} + \alpha_{i0} \end{aligned} \tag{2.5}$$

Where:

x_1 =traffic loads, x_2=velocity, x_3=curvature (1/R),

x_4= gradient, x_5= cant, x_6= sleeper type,

x_7= rail type, x_8= rail length, x_9=falling rock,

x_{10}= land slide, x_{11}= snow and x_{12}= flood.(Guler, Jovanovic and Evren, 2011)

ii. Probabilistic Models

The probabilistic Models can be further divided into three models namely; continuous probability distributions (state-based), Bayesian models and Markov models (time-based). Although there are available models, it is not an easy task to estimate the probability of track degradation considering the influencing factors. Using available research works the three models will be explained.

Continuous Probability Distributions (State-Based)

A continuous probability distribution track degradation model was proposed by Zio et al. (2007), which relates railway system modelled as a stochastically degrading multi-state component with its performance. Monte Carlo simulation is used to simulate the complex stochastic dynamic nature of the railway system. The performance of each section is analyzed through speed and traffic loads. The Norwegian Rail network is divided into sections which are categorized based on the track condition (defects) as illustrated in Figure 2.6. The degradation model presents a state diagram of the defects. There are six sections in this diagram; the track condition h_j in each section j (j = 1, 2, …, n) is discretized in 6 levels. Each section implies the condition of section; where level 5 corresponds to a section with zero defects. The degradation levels from 4 to 1 corresponds to gradually change of track condition from good to worse. The section of level hj = 0 corresponds to rail breakage (i.e., complete failure). A study using this probability distribution model showed that the growth of defect relates to condition of track and can directs affect the performance of the line. The speed restriction on rail sections causes delays which reduce the planned performance. The authors recommend relaxation of speed restriction on low-risk sections to compact the delays observed in high-risk sections. (Zio, Marella and Ã, 2007)

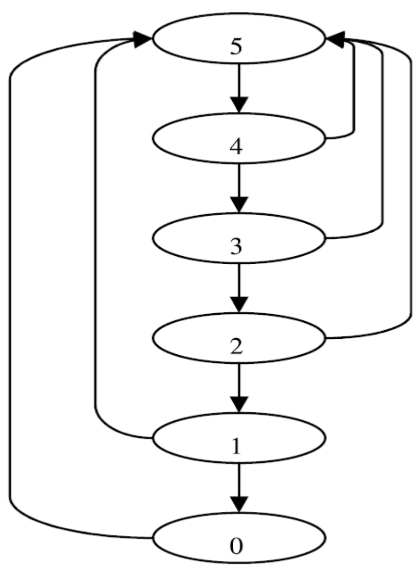

Figure 2.6: Model of defect growth(Zio, Marella and Ã, 2007)

Hierarchical Bayesian Models

Hierarchical Bayesian Models (HBM) are flexible statistical models that allows combination of prior information or from expert judgement and current information. HBM are developed for both quality indicators, conducting an extensive comparison between candidate models and a sensitivity analysis on prior distributions.

Andrade and Teixeira applied Hierarchical Bayesian Model to investigate the Portuguese Railway line (Lisbon-Porto) degradation behaviour. EM 120 inspection vehicle was used to collect the line inspection data in which the standard deviation of longitudinal leveling and standard deviation of horizontal alignment were considered for the study. Total number of 36 inspections data from February 2001 to October 2009 were considered. The accumulated tonnage and maintenance and renew data from operation and maintenance reports respectively where also considered.

The Authors considered the Bayes' rule in prediction of track degradation. Bayesian models are different than classical statistical models in the fact that they assume parameters as random variables. For this reason, prior distribution has to be defined initially. Before any statistical analysis is done prior distribution provide all necessary information to the researcher. This prior distribution p(θ) is then combined with the traditional likelihood p(y|θ) to obtain the posterior distribution of the parameters of interest. The posterior distribution p(θ|y) of the parameters θ given the observed data y can be computed according to Bayes' rule as:

$$p(\theta|y) = \frac{p(y|\theta)\cdot p(\theta)}{\int p\left(y\middle|\acute{\theta}\right)\cdot p(\hat{\theta})d\hat{\theta}}\,\alpha\, p(y|\theta)\cdot p(\theta) \qquad (2.6)$$

Where:

θ is a random variable

y is a random variable the value or probability distribution of which is known,

p (θ | y) is posterior distribution of θ given y which relates to θ via a model,

p (y | θ) is the likelihood to observe y given unknown θ or the sampling distribution of D given known θ, p (θ) is prior probability of θ.

Calculation of prior distribution which include prior mean and prior variance are important steps in Bayans model. Prior mean provides prior estimate of parameter of interest while prior variance provides concerns on certainty of the estimate. Markov Chain Mote Carlo (MCMC) methods are mostly applied in this approach which is used in integration with joint posterior distribution p(θ|y).

Overall, Authors found that Standard deviation of longitudinal level defect (SD_{LL}) provide better fit for the HBM model compared to the Standard deviation of horizontal alignment defect (SD_{HA}). They also suggested future extension of the models such as development of transportation demand model which uses predicted Tonnage as a variable and also use of HBM models to improve life cycle cost of system and safety impact.(Andrade and Teixeira, 2015)

- **Markov Models**

Markov Models are statistical Model which uses the transition probability Matrix to predict the future track condition. This model is based on the theory that the future condition of the track depends on the current state and not on the past system performance.

Jabbar-Ali Zakeri and Shahrbanoo Shahriari used the semi -Markov model to predict the rail wear failure of Lorestan district, Iran Railway line. The selected line consisted of many curves and mostly freight trains. In this research, rail wear in curved track with 250 meters radius was measured by a field scan on a six-month period and rail life was predicted using deterioration probabilistic model. The hazard rate function of the line was calculated and the transition probability of the line was predicted for the next 6 periods of 6-month interval. Passing traffic was considered as the most influencing variable for the rail wear degradation. 7 years passing traffic data was used in the study. They concluded that the rail will require replacement after 3 years based on prediction.(Zakeri and Shahriari, 2012)

By using Markov models a wide range of dependencies can be taken into account. Shafahi developed a Markov model where the track quality index is calculated in a range of 0-100 based on the track unevenness, twist, alignment and gauge measurements. The 100-unit range was then mapped onto 5 states in the Markov model. Transition probabilities for the transition matrix were then established from changes in the track quality index over time.

An alternative 50-state Markov model is proposed by Lyngby to represent the variation of twist over time. In this treatment each of the states represents the twist on a section of track up to 50mm. Degradation model and Maintenance and Renewal model are developed by the authors and impact of optimization of both on the operation cost is also computed. Information on the inspections and failures recorded for 13 years (1989-2002) were collected from Norwegian Railway line (Trondheim to Oslo part) and analyzed. Failures recorded are divided into two groups namely degraded failures (minor cracks and major cracks) and critical failures (failure due to degradation and shock failures).A continuous Markov chain is used to model development of degraded and critical failures within specific inspection interval. A discrete Markov chain is used to model the change of state of section at the end of inspection time. Optimization of track geometry inspections on the Norwegian railway network is done. Alternative deterioration rates were given for the model depending on whether the track section was straight, curved or a transition section. F_1 represents failures due to degradation, can be discovered in inspection. This failure can be avoided by preventive repair. F_2 represents shock failures which are very hard to be discovered in inspection and are caused by large external forces exerted on the rail . D_1 denotes minor degraded failure such as cracks. If not monitored closely the minor cracks can develop to critical failure due to degradation. D_2 denotes larger cracks the rectification of this kind of failure is conducted immediately. The Figure 2.7 shows that for the rail line to reach critical failure F_1 the rail has to grow through degraded state D_1 and D_2. There is a constant rate λ for a railway line to reach shock failure F_2. The railway line is divided into 1m segments so as the states OK,D1,D2,F1 OR F2 can fall in one segment.(Lyngby, Hokstad and Vatn, 2008)

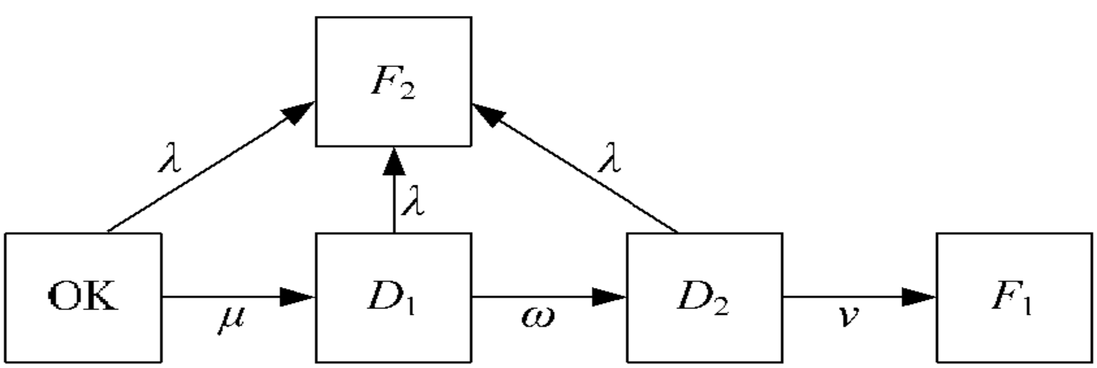

Figure 2.7: General Markov Failure Mode.(Lyngby, Hokstad and Vatn, 2008)

Lyngby also developed a model for the maintenance of the degradation failures. This model tries to predict weather the degraded state will be detected or not. A subscript *u* on the degraded states indicates that a degraded failure is undetected. Likewise, a subscript *d* indicates a degraded failure that is detected. The authors found that Mean time to failure (MTTF) for critical failures for 1km of the line can be reduced by half if the interval of inspections is doubled. Also, if time interval for inspection is extended from 1 year to 2 years the mean number of failures will be doubled. In terms of cost inspection interval of six to eight months gives lowest cost of maintenance per year.

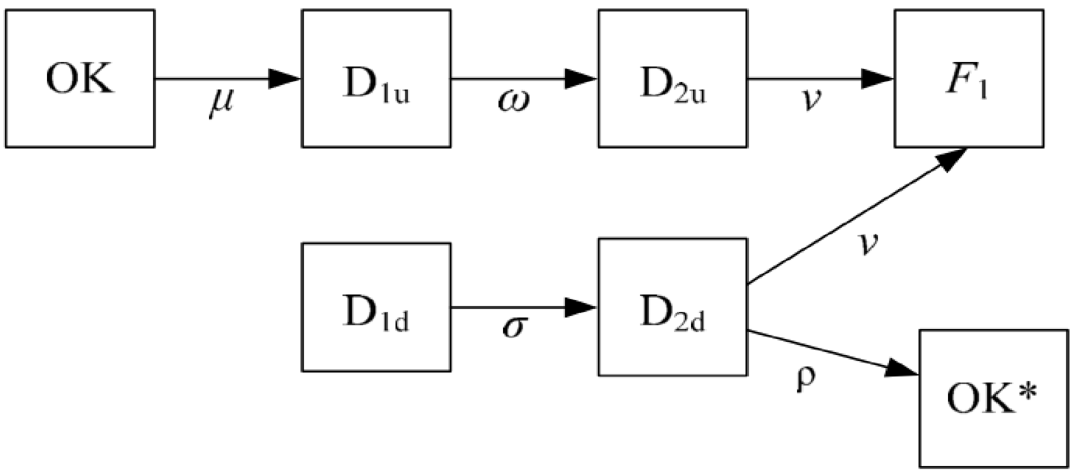

Figure 2.8: Markov model adapted to the railway case. (Lyngby, Hokstad and Vatn, 2008)

iii. Stochastic Models

Stochastic models are based on statistical theory for modeling phenomena where random noise in components exists. Although these models do not deliver based on the underlying physics insights, they use probability to achieve the desired predictions. To predict track geometry degradation, several statistical models have developed. Gamma process, Inverse Gaussian process, and Wiener process are some of the statistical methods that have been used in degradation modeling. It has been noted that most of the researchers have used Gamma process among the mentioned methods to model track geometry degradation.(Andrews, Prescott and De Rozières, 2014a; Hasan, 2015; Muinde, 2018)

Quiroga and Schneider developed an hybrid model (degradation-restoration model) for optimization of tamping schedule. Monte Carlo simulation was used to simulate track geometry degradation using tamping interventions. 20 years of track measurement data obtained from French railway operator SNFC were used. The authors do not consider measurement data of three month after the tamping activities performed on the lines. The reasoning is the track undergoes bedding process. Therefore, their model used data sets with at least 1-year period as this was perceived as a good way to improve precision on the model. Two assumptions are considered in model development, as follows:

• The first assumption is that the initial degradation values after the nth maintenance (tamping) action (an) follows log normal distribution with a mean and variance:

$$a_n = LN(\mu_a(n), \sigma_a(n)) \tag{2.7}$$

Where, μ_a is the mean value, and σ_a is the variance.

• The second assumption is that evolution of degradation values from one tamping to the next is exponentially distributed.

$$a_n e^{b_n(t-t_n)} + \varepsilon(t) \tag{2.8}$$

Where t is the time, t_n is the time for the last tamping activity, b_n is the rate of degradation and is log normally distributed stochastic variable with mean and standard deviation

$b_n \sim LN\ (\mu_b(n), \sigma_b(n))$, with an error (t), which is normally distributed variable with zero mean

$(t) \sim N(0, \sigma^2\varepsilon)$

The authors concluded that the hybrid model offers better performance compared to the normal model (Exponential smoothing method) although they have failed to justify the exponential nature of the model.

Andrews, Prescott, and De Rozieres (2014) analyzed the effect of the maintenance limit and the maintenance response time on the overall quality of the track geometry. Petri net method was used to model the track geometry degradation. Their study concluded that setting lower limits significantly reduced the time that the track spent in a poor condition. They found that an increase in the mean response time would decrease the total number of maintenance actions and would have no effect on speed reduction or line closure. The Peri Net Modeling is done by connecting different stages of deterioration including deterioration stage, inspection stage, Intervention stage and renew stage. Monte Carlo simulations were used to simulate track condition for different inspection interval of 7, 14 and 56 days. Track response to Change of renew period from 30 years to 20 or 40 years. Routine time to repair from 100days to 50, 150, 200 days). (Andrews, Prescott and De Rozières, 2014b)

Vale and Lurdes (2013) proposed a probabilistic model to predict the degradation pattern of the track. Portuguese Railway Northern Line was used as the case study for the development of their model. The degradation rate of standard deviation of longitudinal level is considered to be random. Statistical and probabilistic analyses were conducted on different vehicle speed groups. The dagum probabilistic distribution was selected as it better fit the degradation pattern of track over time. The authors concluded that the rails exhibit same pattern of degradation of longitudinal level.(Vale and M. Lurdes, 2013) The skewness coefficient was calculated from the formula below:

$$\gamma = \frac{\mu_3}{\sigma^3} \tag{2.9}$$

Where:

γ is the skewness of a random variable X,

μ_3 is the third moment about the mean and

σ is the standard deviation of the variable.

The Dagum model was selected to describe the geometrical track degradation process over time. It was selected because a good fit is achieved for representing the degradation of the standard deviation of the longitudinal level for all the time intervals and speed groups for Portuguese Railway Northern Line. The choice of the Dagum model is also supported by the fact that it provides a good fit to both extremes of the observed degradation rate. The Dagum distribution represents the model in the analysis of three variables of function F(x), which is defined by;

$$F(x) = \left[1 + \left(\frac{x}{\beta}\right)^{-\alpha}\right]^{-k}, x > 0 \qquad (2.10)$$

where:
α, β and k are positive variables.

Parameter β is a scale parameter, while α and k are shape variables.

A time series is an ordered sequence of values of a variable at equally spaced time intervals.. It is mathematically defined as a set of vectors $x(t)$, $t = 0,1, 2, \ldots$.. where t represents the time elapsed.

The main aim of time series modeling is to carefully collect and rigorously study the past observations of a time series to develop an appropriate model which describes the inherent structure of the series. This model is then used to generate future values for the series, i.e. to make forecasts. Time series forecasting thus can be termed as the act of predicting the future by understanding the past.

In general models for time series data can have many forms and represent different stochastic processes. There are two widely used linear time series models in literature, viz. Autoregressive (AR) and Moving Average (MA) models. Combining these two, the Autoregressive Moving Average (ARMA) and Autoregressive Integrated Moving Average (ARIMA) models have been proposed in literature. The Autoregressive Fractionally Integrated Moving Average (ARFIMA) model generalizes ARMA and ARIMA models. For seasonal time series forecasting, a variation of ARIMA, viz. the Seasonal Autoregressive Integrated Moving Average (SARIMA) model is used.(Agrawal and Ratnadip, 2013)

Elkhoury (2018) used a time series model named ARIMA to predict the degradation pattern of the Melbourne tram network. The track inspection data, Asset's data (track location, track categories and construction material) and operation data (Million gross tones and number of trips) were used to find the most influencing variable on tram track. The Gauge defect was used as indicating parameter and Million gross tones was found as the most influencing variable. Straight and curved sections of the track were considered in prediction. Using the Most influencing variable, the degradation of tram track was model. ARIMA model and linear regression were used in track degradation prediction and their results were compared. The coefficient of determination was used to validate both models. Both models had the coefficient of determination greater than 0.8 (Elkhoury, 2018)

2.4.1.3. Mechanical-Empirical Models

Mechanical-empirical models are available to improve the performance of the both mechanical and empirical model by using the strong features of each. By using the understanding the mechanical properties of the system coupled with the actual measurements ad operation features mechanical-empirical models produce better degradation model. These models are applied to combat the shortcomings of mechanical and empirical models and improve degradation predictions.

Akihito et al used Mechanistic- Empirical models to build two degradation models to estimate alignment irregularity growth on ballasted track. One was a degradation model to estimate the mean time between maintenance (MTBM), and another degradation model for the preparation of daily/monthly maintenance schedules. Japanese Ballast track technical standard was used in development of degradation model. The vertical load and maximum lateral force of vehicle were estimated first followed by the lateral spring constant of ballast taking into account the vertical load. The lateral deformation was estimated with consideration of traffic volume then the growth of standard deviation of alignment irregularity was obtained through conversion model. The comparison of the actual and estimated alignment irregularity for the different curve radius were done. The MTBM was tested by checking the Maintenance practice change by following the improvement plan (change from jointed rail to CWR) and estimation of cost associated with these changes. The exponential smoothing method was used to build the maintenance optimization model. Before maintenance optimization a degradation model was developed to estimate the surface irregularity growth. The model is expressed as follows;

$$\hat{\sigma}_z(t) = s \cdot \sigma_z(t) + (1-s) \cdot \hat{\sigma}_z(t-1)]$$

$$T_t = s \cdot \{\hat{\sigma}_z(t) - \hat{\sigma}_z(t-1) + (1-s)T_{t-1}\}$$

$$\bar{\sigma}_z(t) = \hat{\sigma}_z(t) + (1-s)/s \cdot T$$

$$\bar{\sigma}_z(t+L) = \bar{\sigma}_z(t) + L \cdot T_t \qquad (2.11)$$

Where:

$\hat{\sigma}_z(t)$: Expected standard deviation of alignment irregularity at t (mm)

$\bar{\sigma}_z(t)$: Estimated standard deviation of alignment irregularity at t (mm)

$\bar{\sigma}_z(t+L)$: Estimated standard deviation of alignment irregularity after L terms from t (mm)

$\sigma_z(t)$: Actual standard deviation of alignment irregularity at t (mm)

T_t : Tendency of increase

s : Smoothing constant (0＜ s ≦1)

To determine an effective track maintenance schedule by tamping machine, they developed a mathematical programming model by applying OR technology. The objective of the programming model was to minimize the mean value of the weighted standard deviation of surface and alignment irregularities within the period of the schedule. The formula used is expressed below;

$$r_{yz} = \alpha \cdot s_y/s_{ytg} \cdot e_y + (1-\alpha) \cdot s_z/s_{ztg} \cdot e_z \qquad (2.12)$$

Where:

r_{yz} : Quantity of track maintenance improvement with tamping machine

s_y : Standard deviation of surface irregularity

s_{ytg} : Criteria for surface irregularity

e_y : Improvement rate from maintenance for surface irregularities

s_z : Standard deviation of alignment irregularity

s_{ztg} : Criteria for alignment irregularity

e_z : Improvement rate from maintenance for alignment irregularities

α : Weight coefficient (0≦α≦1)

They concluded that the lateral deformation can be found as three times the standard deviation of the lateral irregularity. Further alignment irregularity growth tends to be prominent around normal joints in sections with EH-type rails, where ballast has suffered severe damage or where the condition of welded joints has deteriorated. (Kawaguchi, Miwa and Terada, 2005)

Mechanical and Empirical models were used by Sadeghi and Askarinejad to develop a track degradation model which incorporate the track structural condition. Mechanical model was used to investigate the mechanical properties of track structure and railroad vehicle and their contribution on track degradation. Empirical model are 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 with inspection data collected for a period of 2 years applied in analysis. Track quality, Traffic and Maintenance history were the main parameters considered in this study affecting rate of track deterioration. Two quality indices are considered in this study 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, subgrade and drainage system. Most effective traffic parameters considered are Total million gross tons (EMGT) and average running speed (V). For maintenance parameter time since any major maintenance operation was considered.

Mathematical equations are developed to correlate main effective parameters (loading, track maintenance and track quality) and track degradation coefficients (DC) . As a result of observation of track behavior for a period of 1 year an equation representing change of degradation coefficient in response to time is also developed.

The model is expressed in two forms: one based on the relationship between the track geometry conditions and time (Equation 2.13), and the second based on structural visual inspections of the mechanistic conditions of the track components over time (Equation 2.14).

$$\frac{TGI_2}{TGI_1} = \alpha_4 \exp\left(\beta_1 V + \beta_2 EMGT + \beta_3 TGI_1\right) \times \left(\lambda_1 T^4 + \lambda_2 T^3 + \lambda_3 T^2 + \lambda_4 T + 1\right) \qquad (2.13)$$

$$\frac{TSI_2}{TSI_1} = \acute{\alpha}_4 \exp\left(\acute{\beta}_1 V + \acute{\beta}_2 EMGT + \acute{\beta}_3 TSI_1\right) \times \left(\acute{\lambda}_1 T^4 + \acute{\lambda}_2 T^3 + \acute{\lambda}_3 T^2 + \acute{\lambda}_4 T + 1\right) \qquad (2.14)$$

Where:

TGI_2 is the future track geometry index,

TGI_1 is the present track geometry index,

TSI_2 is the future track structure index,

TSI_1 is the present structure index,

T is the time (in seconds).

Furthermore, a model was formulated providing the correlation between TGI_2 and TSI_2 to limit the study to fewer inspections, as follows

$$TSI_2 = \eta_1 \eta_2 \eta_3 \eta_4 TGI_2 \qquad (2.15)$$

Where:

η_1 represents train speed, η_2 represents equivalent million gross tons (EMGT), η_3 represents initial (TGI_1), and η_4 represents time (T)

Obtaining linear correlations between the ratio of TGI_2/TSI_2 and the influencing variables, the following relations are obtained for η_1 to η_4

$$\eta_1 = k_1 V + k_2 \qquad (2.16)$$

$$\eta_2 = k_3 EMGT + k_4 \qquad (2.17)$$

$$\eta_3 = k_5 TGI_1 + k_6 \qquad (2.18)$$

$$\eta_4 = k_7 T + k_8 \tag{2.19}$$

Where: k_1 to k_8 are constant coefficients.

Conclusion drawn from this paper are track geometry experience more degradation compared to track structure, also more attention should be given to the turnouts, transition parts and bridges as their deterioration rate are high. Hence, frequent inspections and better maintenance schedule should be involved. The Authors recommend evaluation of applicability of developed degradation models

models further emphasize on the use of the analytical methods to develop degradation models .(Sadeghi and Askarinejad, 2010)

2.4.1.4. Artificial Intelligence Models

Artificial Intelligence models are type of machine learning models used in the prediction of the rail track degradation. The artificial intelligence models can be divided into two categories Artificial Neural Network Model (ANN) and Neuro-Fuzzy Models.

i. Artificial Neural Network (ANN)

ANNs are biologically inspired computer programs designed to simulate the way in which the human brain processes information. ANNs gather their knowledge by detecting the patterns and relationships in data and learn (or are trained) through experience, not from programming. Recently, the neural network approach to data analysis and modeling has received much attention from researchers and practitioners. Over the past decade, there have been significant activities and important breakthroughs in the development of various theoretical and computational models for connectionist computing.

Sadeghi and Askarinejad applied Artificial Neural Network to evaluate railway track quality condition. The aim of their research was to develop a methodology for the establishment of correlations between the track structural conditions and the data obtained from automated inspections. Data from Iran railway network was used in this study. The track geometry data including gauge, profile, alignment and twist were obtained from the track recording car (EM120). The track structural defects including rail, sleeper, fastening and ballast defects were recorded and quantified using visual inspection carried out on tracks for 2 years. In this study, multilayered feed-forward networks with one hidden layer were utilized. The networks were trained by the back-propagation learning algorithm using the real-valued sample data. Figure 2.9 illustrate the three-layered neural network.

The mean squared error technique was used to assess the networks' performances in the training and validation procedures. The author concluded that deteriorations or improvements in the quality conditions of the track have influences on the accuracy of the network prediction. The proposed neural network models have a better efficiency and accuracy in tracks with medium and low-quality conditions. (Sadeghi and Askarinejad, 2012)

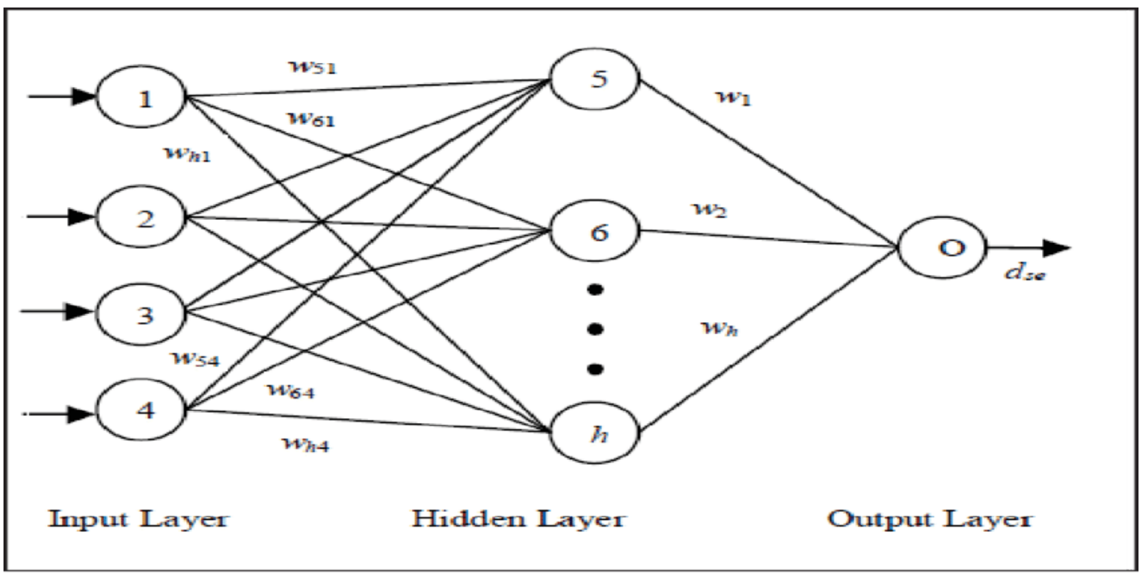

Figure 2.9: Typical architecture of a three-layered neural network.(Wu, Dandy and Maier, 2014)

ii. Neuro-Fuzzy Models

The combination of Artificial Neural Network (ANNs) with Fuzzy Interface System (FIS) is called neural fuzzy or neuro-fuzzy systems. Neuro-Fuzzy combines the learning and connectionist property of Artificial Neural Network and reasoning property of fuzzy neural network. Neuro-Fuzzy are universal approximators which are able to interpret IF-THEN rule.

Shafahi et al used Markov Chain, Artificial Neural and Neuro-Fuzzy Network methods to develop track degradation model. The developed models were compared to conventional ORE model of the International Union of Railway. The data was collected from the Iranian Railways network and incorporated four types of survey: topography, annual traffic and axle load, date of construction or reconstruction, and the track condition of the block in each year. The Jang's ANFIS model was used to build the neuro- fuzzy model. The ANFIS structure membership function parameters, if-then rule exertion and output parameters are calculated by train data set. The training algorithm is usually hybrid or back propagation. The model was able to predict the degradation pattern of the track with the performance of 73%.ANFIS model offers better performance compared to ANN model by 6%.(Shafahi, Masoudi and Hakhamaneshi, no date)

- **Classification of track degradation based on Methods applied in analysis**

The track degradation can also be classified by methods used in development of degradation models. Different methods are applied by researchers in developing degradation models. It has been observed from the literatures that some methods are most preferred by researchers than others, methods such as regression analysis were the first applied and are preferred due to their simplicity while the gamma process were mostly applied in stochastic modeling based on its ability to accommodate large data set. In the last two decades' researchers have been using probabilistic models' methods such as Weibull distribution and log- normal distribution to predict track degradation. New methods such as Artificial intelligence methods and data mining methods are currently applied in practice. Figure below show classification of track degradation based on method applied in the analysis

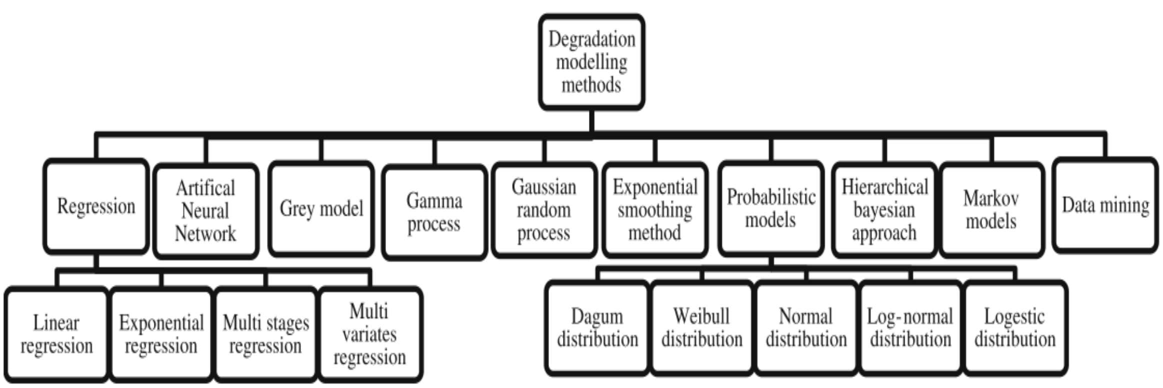

Figure 2.10: Track degradation modeling methods.*(Kumar* et al.*, 2016)*

2.5 Comparison of different track degradation models

Four degradation models (Mechanistic models, Stochastic models, Mechanic-Empirical models and Artificial intelligent models) extensively described above are compared. The strength and weakness of each model are described and finally appropriate model for AALRT track geometry degradation will be selected. Table 2-1 shows the comparison of different track degradation models.

Table 2-1: Comparison of different track degradation models

Approach	Strength	Weakness
Mechanistic Models	• Based on Mechanical properties of track component • Uses primary data source (laboratory experiments) • Small amount of data can be used in model development	• Model is unable to deal with heterogeneity of track structure materials. • The model is applicable in limited number of track sections rather than the whole section. • Time consuming and intensive.
Statistical Model (Deterministic model, Probabilistic Model and Stochastic Models)	• Ability to work with large set of data and provide realistic results • Transparent model (model development can be easily understood) • Can solve complex and sophisticated problems better than mechanistic models.	• Unable to capture the effect of uncertainty in model prediction modelling (deterministic models) • Memoryless property of process makes the function of model limited (Probabilistic models) • No Validation of exponential degradation pattern (Stochastic Models)
Mechanical-Empirical Models	• Uses both laboratory experiments and historical data	• Show higher rate of degradation on turnouts and

	• Applicable in different track segments (e.g., bridges, turnouts)	bridges compared to other parts
Artificial Intelligent Models	• Ability to solve complex problems through learning data pattern • Can generate estimations with higher accuracy	• Lack of transparency during decision making process • With no enough data training of the model is not well done.

2.6.Summary

In this chapter, previous works on track deterioration prediction have been reviewed. It is clear that a vast amount of work is available on this topic. A number of stochastic, deterministic and AI-type models have been proposed and a number of methods have been adopted by different researchers. The literature review presented here has revealed the components most suitable for integration into the present study. However, a research gap exists in understanding the effect of influencing factors on the deterioration models of railway track. Furthermore, there is scarcity of research on the light rail deterioration prediction. The research presented in this thesis aims to address this gap in knowledge. Regression model and Artificial Neural Network models will be applied to predict track geometry degradation of AALRT.

CHAPTER THREE

RESEARCH METHODOLOGY

3.1. Introduction

Review of literature has provided an insight on the current practice in prediction of track geometry deterioration. After critical evaluation of literature and understanding the peculiarity of light rail transit this methodology is adopted to solve the research problem. This chapter presents research methodology employed to develop models to predict degradation of light rail using the measurements of track geometry, operation statistics data and asset management data collected over the past years. This chapter includes overview of the location of the study, study design, the data collected and the procedure followed in development of Regression model and Artificial Neural Network.

3.2 Location of study

The metropolitan electric railway in Addis Ababa is a double track and has 34.25 km with a North-South line of 16.9 km and East- West line 17.35 km. East-West line covers the areas such as Ayat Village to Tor Hailoch passing through Megenagna, Legehar and Mexico Square and North-South line passes through Menelik II Square, Merkato, Lideta, Legehar, Meskel Square, Gotera and Kaliti (Jemere, 2012). The North-South corridor is a heavily-used route; it is utilized to access Africa’s largest open market at Merkato. From Lideta to Meskel Square both services have a common route until the North-South route bifurcates to the south heading to the new Gotera interchange, Saris and Kaliti. In its final phase the route will cross the ring road at Kaliti interchange and extend to a new railway hub at Akaki. The northern end of the route will be extended from Giorgis to Shiromeda. The initial phase of this route is 16.97km length and expected to extend up to 36km in the long-term plan. The nominal track gauge is 1435 mm with maximum grade typical 5‰. The two lines have a capacity to transport 60,000 passengers per hour. The maximum service speed is 80 km per hour. The LRT currently covers 39 stations with about 30 LRVs operating daily, each with a carrying capacity of 286 passengers. The elevated sections of the track are ballast less (slab track), while the ballast track is found in the non-elevated sections of the track. The ballast track is composed of mono block precast reinforce sleepers.

This research covers the North- South line which consist of double track starting from Meskel Square to Kalit. Figure 3.1 and Figure 3.2 shows the LRT alignment and LRT stations respectively.

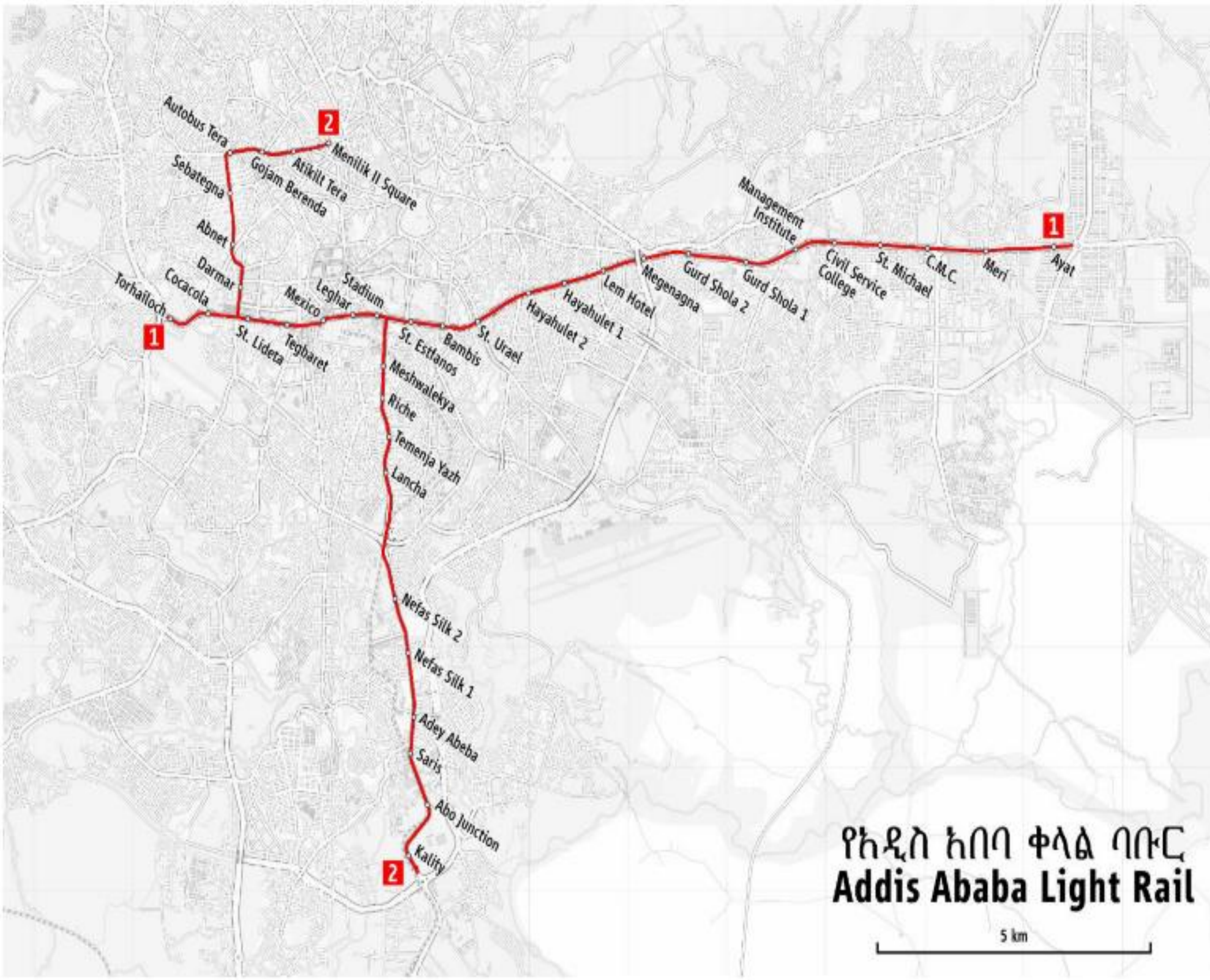

Figure 3.1: Existing LRT Stations on North South line and East West Line.(Fekadu, 2017)

3.3 Study Design

This study aims at predicting rail track degradation for Addis Ababa Light rail transit. To archive this goal, the following procedures are followed; Different literatures, reference books, internet web-sites, articles and international journals those focus on rail track degradation prediction were reviewed. The data set including inspection data, asset management data and operation data of North South line were collected from AALRT and used as input in this research. SPSS Software was used in analysis of data to determine track geometry degradation most influencing variables. Accordingly, regression degradation model and Artificial Neural Network Model (ANN) are developed for curve and straight sections so as to understand the degradation pattern of Addis Ababa Light rail Transit. The track geometry degradation models for straight and curved sections are the outputs of this research. Finally, the paper work is concluded and the necessary recommendation and future work are mentioned.

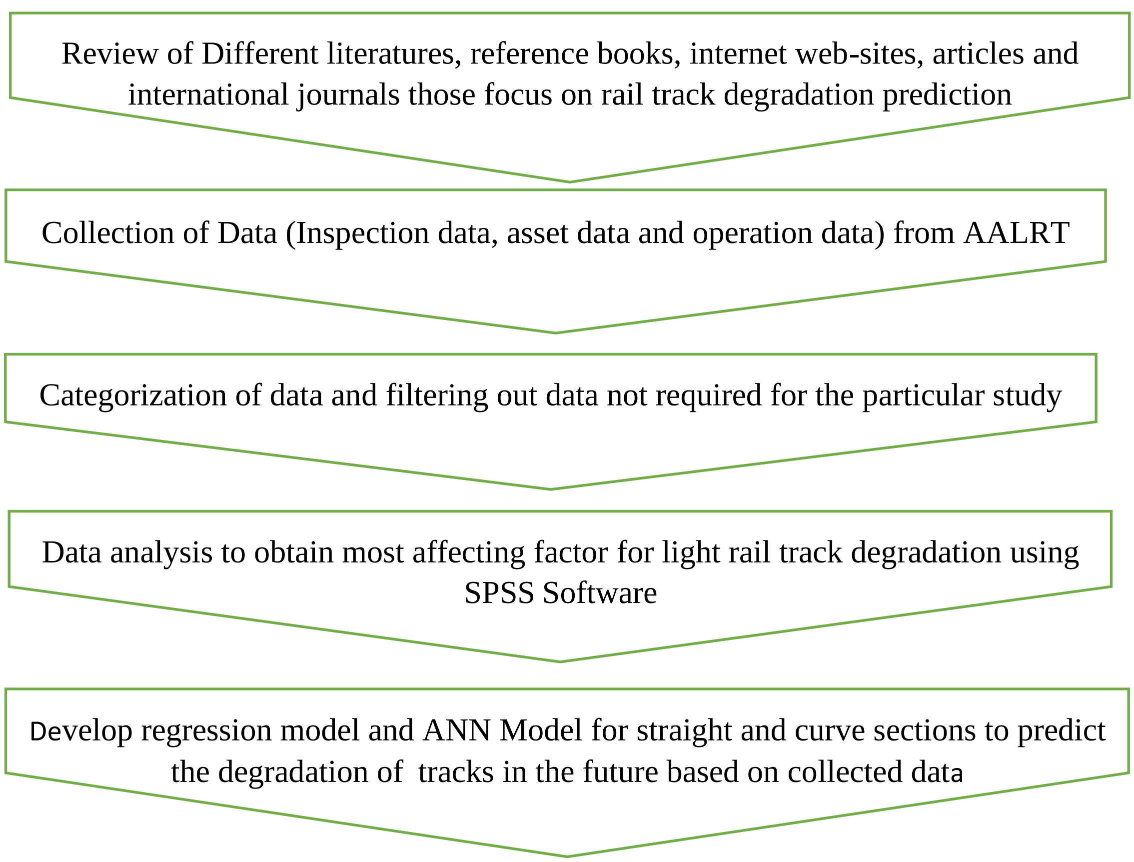

Figure 3.2: Research process framework

3.3.1 Dataset

Railway Organization perform inspections and maintenance activities frequently so as to ensure safe operation of system in so doing large amounts of data are collected. Data set for this research were collected from AALRT. AALRT conducts inspection of the sections of the track daily on foot, or by riding over the track in a vehicle at a speed that allows detection of noncompliance with the standards. Inspections are performed by qualified personnel from LRT. Portable devices are used to record the geometry defects such as gauge, cross level and twist while visual inspection is used to detect defect of horizontal alignment. Components of a section of track are inspected and their condition recorded on an inspection form, with all deviations or deficiencies recorded on the form. Remedial action for defects is taken in accordance with the parameters set forth in the Chinese standard. Copies of inspection forms depicting the condition of track were obtained and are used as input in this research.

Figure 3.3: Gauge ruler for measurement of track geometry(Francis, 2018)

3.3.1.1. Inspection Data

Addis Ababa Light Rail Transit conduct inspection of track daily and track condition are recorded in the inspection form. 10 months inspection data for year 2018 and 2019 were collected and used in this research. The inspection data for double track (Uplink and Downlink) of North South line of AALRT were applied in track degradation modeling. The inspection data collected included track geometry measurements of the following;

- Track gauge: This is the distance between rail tracks and is measured at a right angle between the inner faces of the load-bearing rails. The standard railway gauge is 1435 mm. Gauge ruler is used to measure the gauge defect for the line.
- Alignment (also defined as horizontal alignment): This is the change in the curvature of rail over a certain cord length. Visual inspections are used for measurement of horizontal alignment of the line.
- Twist: This is the change in cross-levels of curve rail tracks over a specified cord length (expressed in mm/m). This is the difference in height between the top surfaces of two rails of the railroad track. The cross-level is zero when there is no difference in height between the surfaces of rails at the measured point. The cross-level is negative (also known as reverse cross-level) when the outer rail of curve track has a lower height than the inner rail and is positive when otherwise.
- Longitudinal level/Profile is the track geometry of the track center line projected onto the longitudinal vertical plane. The vertical deviation of this center line is termed as longitudinal level/ Profile defect. According to the International Union of Railways, it is usually the short-wavelength of the longitudinal level measurement that drives the need for track geometry maintenance.(American Railway Engineering and Maintenance of Way Association, 2006)

  Hence it is considered as the main influencing parameter for this study.

The inspections are conducted on both curved and straight sections of the line daily. The inspections are done on a line at a distance of 5m from one point to the other. The points which show defects beyond tolerable limit are sent to the repair team and maintenance of such is done during the night. The table below shows the tolerance limit used by AALRT for maintenance of the track.

Table 3-1: The tolerance management values of static geometric dimension of track (China, 2010)

Items		Acceptance of operation (mm)		Regular maintenance (mm)		Urgent repair (mm)	
		Main track and auxiliary track	Yard track	Main track and auxiliary track	Yard track	Main track and auxiliary track	Yard track
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

3.3.1.2. Asset Data

The geometrical properties information of the line was obtained in this part such as whether the track segment is straight or curved (track category), track location (i.e., Uplink or Downlink) and construction features (i.e. track surface material, curve radius and Grade).

3.3.1.3. Operational Data

Operation data are data related to the use of track. Track Operations includes static and dynamic resources. Rail infrastructures such as lines, bridges, stations, crossing, tunnels, etc. are part of static resources while moving assets such as passengers and wagons are part of dynamic resource. Tonnage, number of trips and Speed are the operation data which will be used in this research. Tonnage was calculated after obtaining information such as Passenger flow for each month considered, weight of trains and average weight of passenger. Gross tons are the product of total weight including the weight of locomotives as well as the weight of the average annual daily traffic volume passing the track. Number of trips is the number of trains passing over each rail track segment. Maximum speed allowable for trains to pass through track section is used in this study to evaluate its contribution in track geometry degradation.

3.3.2 Filtering and Data Categorization

This research is based on secondary data collection; thus, data scrutiny was required. This was for the purposes of confirmation of suitability of the available data on basis of reliability, adequacy and if the source data meets the demands of the problem. The inspection forms included data for gauge, cross level and twist defects. Longitudinal level defects are the only consideration for this research. Therefore, the data was sorted to work with only longitudinal level recorded on the inspection forms. Data was then inserted into excel sheet for easy usage of data from copies of inspection forms.

Table 3-2: Summary of Data Used in degradation model development

Statistical Parameter	Straight Sections	Curved Sections
Minimum Value	5	5
Maximum Value	20	11
Mean	5.545	5.695
Standard deviation	5.198	4.962
Total Sample data	760	850

3.3.3. Data Analysis Using SPSS

Evaluation of influencing factors on changes of longitudinal leveling is done using SPSS Software. The factors considered in this study are Tonnage, Number of trips, Speed, Grade, Curve Radius ad Track Surface. Analysis is conducted for the straight and curved sections of the line using the collected data. From the analysis the most influencing variables on track degradation are discovered and used in prediction of track geometry degradation model. Maintenance works are not considered in this research. Correlation analysis is used for continuous variables (Number of trips, Tonnage, Speed, Grade and Curve Radius) while Analysis of Variance (ANOVA) is adopted for category variable (Track surface). These analyses are evaluated by observation of correlation values which range from -1 to +1 where values closer to 1 indicate close relationship between variables and 0 indicate no relationship. The significance checks whether the obtained correlation values are statically important (p value should be less than 0.05). The variables which shows stronger correlation and the values are significant are considered to be influencing factors and are applied in track geometry degradation modeling

3.3.4. Regression Model

Regression model is a simple predictive model applied in this study to determine the status of the track geometry in the future. The developed model will help in determining the status of track and maintenance plan can be developed from these observations.

The relationship between the change in longitudinal level and most influencing variable will be identified. The most influencing variables determined in Chapter 4 will be used together with longitudinal leveling parameter to develop regression model over time using SPSS Software. Training data sample is used to provide results of the model including estimated variables and coefficients of the model. Finally Model performance is checked by using the validation data in which goodness of fit of model is observed. The goodness of fit is checked by plotting the observed values versus the predicted values for both straight and curved section of the line.

The steps followed in the regression model development are described in details in this part.

3.3.4.1 Regression Model Development

i. Multiple Regression Model

The Multiple regression model is used to evaluate the track geometry degradation of N-S line of AALRT. Regression analysis is a set of statistical processes for estimating the relationships among variables. Regression analysis includes the dependent variable which is the outcome expected at the end of analysis and dependent variables which are instruments used to reach the outcome. Benefits of using regression analysis includes understanding the relationship between independent and dependent variables, obtaining the relative strength of different independent variables' effects on a dependent variable and to make prediction of future pattern of a series or activity. The simplest form of regression models has the format show below;

$$y = \alpha + \beta_1 x_1 + e \quad \text{.....}(3.1)$$

Where;

y is the dependent variable which is variable to be estimated, α is a constant or interception of regression model, $x_1$is the dependent variable, β_1 is the (regression) coefficient of the independent variable x. This coefficient represents the gradient of the line and is also referred to as the slope. And e represents 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 Multiple regression analysis applies the following equation in model development.

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

Where:

X's are the independent variables, Y is the dependent variable. The subscription j represents the observation row number. The β's are the unknown regression coefficients. The ε_j represents the 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 track data over time.

3.3.4.2 Regression Model Build Up

Multiple regression analysis is used in this study to produce degradation model for the straight and curved section of N-S line of AALRT. Several steps were employed to build the multiple regression model. Such steps include data required to conduct regression analysis, to specify and estimate regression model, to set and test assumptions of regression model, interpretation of results, validation of regression model and application of model. The figure 5.1 below illustrate the steps followed in the development of model.

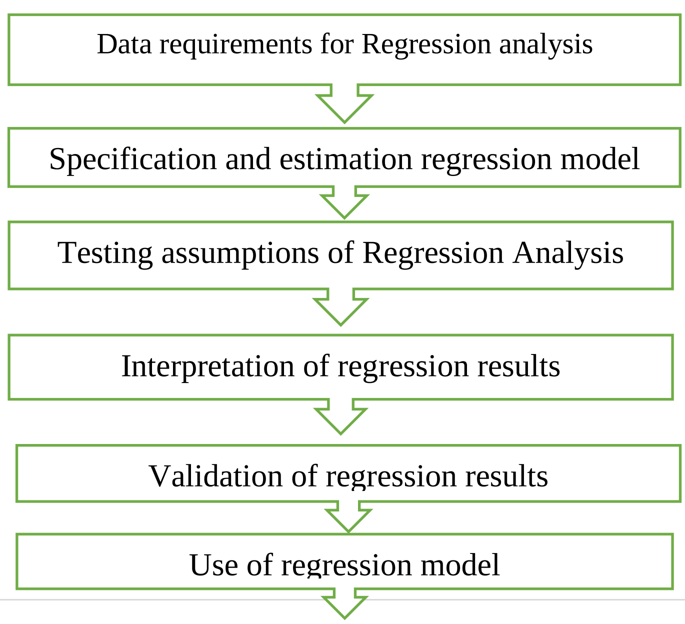

Figure 3.4: Steps to conduct Regression Analysis

i. Data Requirement for Regression Analysis

Data requirement for regression analysis include aspects such as sample size, variables variation, scale type of variables and collinearity of independent variables. In order to conduct regression analysis a sufficient large sample size is required. The acceptable size of sample data is a key point in obtaining regression analysis results. There is a thumb rule formula proposed By Green which is used to test the individual parameter effects. The acceptable sample size proposed by the author is 140+k where k is the number of independent variables. Another approach is using power analysis which is difficult and detailed parameter specifications such as expected effect size and maximum type of error allowed. In this study a total of 600 samples of data are used in the regression modeling of curved and straight section of N-S line. Further the 70% of data is used for model training and 30% is used for model testing. Training data is used in model development and testing data is used for model validation.

Variable variation is another data requirement for conducting regression analysis. Regression analysis can only be conducted when there is variation in independent variables. The variation in independent variable is a key explanation for variation of dependent variable. Independent variables (Number of trips and tonnage) used in this study have shown variation in its value which indicate regression analysis can be done to predict change in longitudinal leveling.

Ratio and Interval are scale types of variables which favors regression analysis. If the data used includes only two values (zero and one) also termed as binary number a binary regression analysis is recommended. The scale type of variables applied in this study are interval which favors regression analysis.

Collinearity occurs when two independent variables are highly correlated. Multi-collinearity occurs when more than two variables are highly collated. The main problem of having collinearity between variables is that it tends to disguise significant parameters as insignificant. VIF (Variance Inflation Factor) which is a reciprocal value of tolerance is used to detect collinearity between variables. VIF below 0.1 indicates presence of multi-collinearity and VIF above 10 indicates presence of collinearity. To remove collinearity in data sample two methods can be applied; namely factor analysis and re-specifying of regression model. (Mooi, 2016)

ii. Specification and Estimation of Regression Model

Regression model specification involves selection of dependent and independent variables. The selection of independent and dependent variables depends with the need of the study and the research question required to be answered. Prediction of track geometry degradation pattern is the basic aim of this study and longitudinal leveling is influencing parameter employed and used as dependent variable. Most studies and railway agency around the world use longitudinal leveling as indicating parameter for track quality. This is a main reason behind to selection of longitudinal leveling as influencing parameter. Other reasons are availability of referencing materials and availability of sufficient data from AALRT to conduct modeling. Number of trips and Tonnage (MGT) are selected as independent variables. These variables are selected after the correlation analysis conducted in chapter 4. The correlation analysis results showed that these two variables are most influencing variables on change of longitudinal leveling of N-S line of AART.

Model Estimation is explained based on SPSS Software used in this study. Model Estimation involves selection of analysis methods and selection of estimation method. Selection of analysis method include two methods such as Enter and Stepwise. Stepwise Method can be further divide into Backward and Forward. Enter method is the method which allows the user/researcher to select the independent variables to be used in the analysis while stepwise analysis allows the process to select the best subset of variables available. Forward and Backward method used the techniques of step modeling.

Forward method modeling begins with a constant which conducts development of large regression models Then it tries to find the best model by adding just one independent variable from the remaining variables. Subsequently it compares the results between these two models. If adding an independent variable produces a significantly better model, it proceeds by adding a second variable from the variables that remain. The resulting model (which includes the two independent variables) is then compared to the previous model (which includes one independent variable). This process is repeated until adding another variable does not improve the model significantly. The backward method does something similar but initially enters all variables that it may use and removes the least contributing independent variable until removing another makes the model significantly worse. The disadvantage of using Stepwise Method is it only adds variables which are only significant by chance and not truly interesting and useful. For this particular study Enter method is used because it gives mandate to the user to specify the kind of regression analysis to be conducted.

Selection of Estimation method involves selection of methods to be used in estimation. There are several methods such as Ordinary Least Square (OLS), weighted least squares (WLS) and two-staged least squares (2SLS). By default, the SPSS Software uses OLS in analysis. OLS estimates a regression line so that it minimizes the squared differences between each observation and the regression line. By squaring distances, OLS avoids negative and positive deviations from the regression line cancelling each other out. Moreover, by squaring the distances, OLS also puts greater weight on observations that are far away from the regression line.

iii. Testing Assumptions of Regression Analysis

Five assumptions have to be satisfied by the regression analysis to get valid results. These assumptions are linearity; expected mean error is zero, Homoskedasticity, No autocorrelation and Error distribution. If the regression analysis fails to meet this assumption the regression parametric significance determination becomes difficult although the developed model will be accurate.

- Linearity: There should be a linear relationship between the independent and dependent variables. Linear relationship allows the regression model equation to be linear in the form of $y = \alpha + \beta_1 x_1 + e$. To check the linearity between x and y variables can be done by plotting independent variables against dependent variables.

  Scatter plots can be used to analyze if there is nonlinear pattern. In this study the relationship between the dependent and independent variables was checked for both straight and curved sections of N-S line.

- Expected Mean Error of regression model is zero: This assumption requires the sum of expected errors to be zero. If such assumption is not fulfilled the obtained regression line will either over or under estimates the true relationship of variables. The Ordinary Least Square (OLS) always selects the best line in which the expected mean error is zero. Hence this assumption does not require testing as long as OLS is used as estimation selection method.

- Homoskedasticity: This is a situation which errors variance are constant. The presence of homoskedasticity in regression model causes the obtained regression coefficients not to be significant although the prediction model will be correct. The Homoskedasticity can be observed by creating graphs of errors against dependent variables. As dependent variable increases or decreases the errors variance should remain constant. Presence of Homoskedasticity can be observed on the graph in which the point becomes more or less spread across the graph and forms a tunnel shaped. Weight Least Square (WLS) is the method applied in SPSS to deal with homoskedasticity. Variable which causes the variance not to be constant is applied here and the obtained results are weighted by this variable.
- No autocorrelation of errors: This means the regression model errors are independent. It means errors from any two observations are not correlated. The autocorrelation of data in SPSS can be tested using Durbin–Watson test. The Durbin– Watson test assesses whether there is autocorrelation by testing a null hypothesis of no autocorrelation, which is tested against a lower and upper bound for negative autocorrelation and against a lower and upper bound for positive autocorrelation. To carry out this test, first sort the data on the variable that indicates the time dimension in your data, if you have this included in your data. Otherwise, the test should not be carried out. In this particular study this assumption is not taken into consideration.
- Error Distribution: This assumption indicates the regression model errors are approximately normally distributed. This assumption is considered as optional because even if the errors are not normally distributed the regression model still provides good estimation of coefficients. Use plots or performance of a formal test are methods used to check if regression model errors are normally distributed.

iv. Interpretation of Regression Results

On interpretation of regression results, how the developed model perform is checked. The performance of regression model is checked by evaluating the goodness of fit of the model. In this study the model performance is checked using Coefficient of determination or R^2 and Mean Squared Error.

R^2 or coefficient of determination indicates the degree to which the model explains the observed variation in the dependent variable, relative to the mean. R^2 shows how regression line is close to observed values. The R^2 adopts the following equation;

$$R^2 = \frac{(\hat{y}-\bar{y})^2}{(y-\bar{y})^2} \qquad (3.3)$$

Where:

$\hat{y}$ is regression line,

$\bar{y}$ is the mean absolute values,

y is the observed values.

In this research, y represents the observed values of longitudinal leveling defects. The R^2 value explains the performance of the model in predicting the future track 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. The adjusted R^2 is also taken into account in the regression analysis. The adjusted R^2 is a model accuracy measure that tends to estimate the fit of the linear regression. The adjusted R-squared consider number of independent variables used in the prediction of dependent variables. It checks whether increase of independent variables improve model fit. The value of the adjusted R^2 is always less than R^2. Adjusted R^2 ranges same as R^2 between 0 to 1 with value closer to 1 meaning good model fit.

Mean squared Error (MSE) is the average squared difference between the estimated values and the actual value. Almost always strictly positive (and not zero) is because of randomness and small values indicate better model. The formula governing the MSE value is shown below.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2 \qquad (3.4)$$

Where:

MSE= Mean Squared Error, n= Number of data points, Y_i= Observed values and $\hat{Y}_i$= Predicted Values.

v. Validation of Regression Model

The stability of the regression model has to be checked. Stability means that the results are stable over time, do not vary across different situations, and do not depend heavily 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. 70% of the randomly chosen data are often used to estimate the regression model and the remaining 30% are used for comparison purposes. 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

After a useful regression model that satisfies the assumptions of regression analysis is obtained, it is time to use the regression model. Use of regression model in this study is prediction. Essentially, prediction entails calculating the values of the dependent variable based on assumed values of the independent variables and their related but previously calculated unstandardized β coefficient. The predicted longitudinal leveling is a function of longitudinal leveling of previous year, number of trips and tonnage for the straight sections of North South line. For the Curved Sections the predicted LL is function of Number of trips, tonnage and Curve Radius. The functions can be written as

Longitudinal leveling (t) = f (number of trips (t-1), tonnage (t-1), Speed, Grade) ………………………………………………………………..…………………… (3.5)

Longitudinal leveling (t) = f (number of trips (t-1), tonnage (t-1), Curve Radius, Speed, Grade)……………………………………………………………………………… (3.6)

3.3.5. Artificial Neural Network

Artificial Neural Network is part of Machine learning which performs as neurons of Human brain. ANN uses the knowledge obtaining through the training of data to predict the unknown values. This makes ANN advantageous to other empirical models also it has ability to capture the effect of different parameters influencing deterioration in predictive model. ANN model has three layers which are input layer in which the influencing parameters are introduced into the system. Hidden layer in which the learning of system is done and weights are signed to the nodes and the Output layer were the estimated values are displayed after the calculation. The influencing variables obtained in Chapter 4 will be used in modeling of track degradation using ANN. The procedures followed in developing the model together with assumptions made are detailed in this Chapter. The ANN model will be developed using MATLAB Artificial Neural tool named nntool. The model will be validated by plotting the predicted values versus the observed values and goodness of fit of the model will be checked.

3.3.5.1 Artificial Neural Network Model Development

i. Introduction

Artificial Neural Network (ANN) is part of artificial intelligence which uses part of data to learn the pattern of process and afterwards perform complex calculation to provide values for unseen data. ANN are biologically inspired and perform the functions as Neurons of Human brain. These neurons are connected to each other using connections. The relationship between input and output parameters are known during learning stage of model and the knowledge in stored in these connectors. The ANN uses the knowledge in these connectors to solve provided problem. A simple ANN architecture has three layers; an input layer, one hidden layer and output layer. ANN has several advantages over the other empirical models such as ANN has ability to approximate large class of functions with high degree of accuracy and the prediction depends on the characteristics of data and no prior assumptions is required. The diagram below shows the simple architecture of ANN.

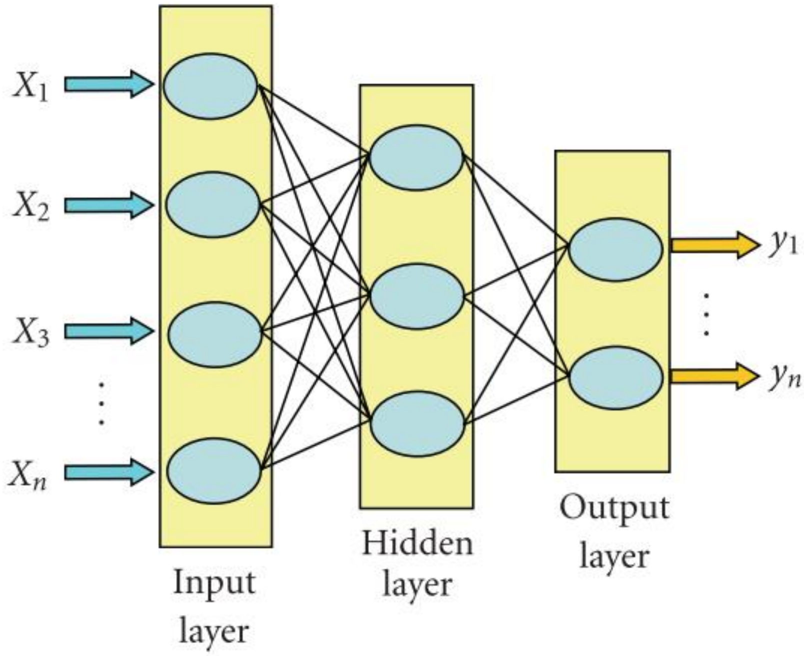

Figure 3.5: Artificial Neuron Network Architecture(Shahin, Jaksa and Maier, 2009)

3.3.5.2 ANN Model Build up

Artificial Neural Network is applied to predict the track geometry degradation using the same sample data applied in the regression model. The aim is to compare the accuracy of these models in degradation prediction. ANN model was selected based on its advantage of not requiring any prior assumption and prediction is solely based on the data characteristics. Furthermore, ANN models were developed by Engineers and computer scientists whose main target were to obtain prediction accuracy and methodology which provide better results.(Maier and Dandy, 2000)

Four models are developed using ANN, two models are for curved section of North south line and the remaining two for the straight sections of N-S line. The model build up was based on different steps followed which are explained in detailed below. The steps include selection of input, data splitting, selection of model architecture, determination of model structure, model training (and model validation. The steps followed are illustrated in the figure 5.7 below.

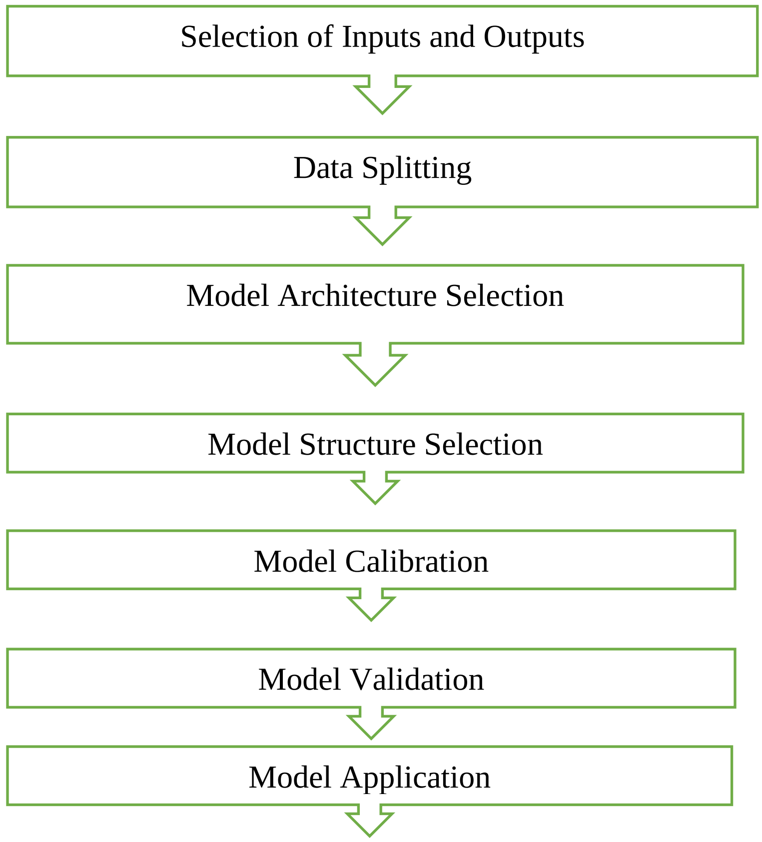

Figure 3.6: Steps in ANN Model Development (Maier *et al.*, 2010)

i. Choice of Potential Inputs and Outputs

This is the first step in the development of ANN model in which the variables to be modeled (appropriate output variables) and the selected input variables are selected by the modeler. Selection of input and output variables are based on prior knowledge of the modeler or the available data. There are several techniques employed in the selection of input variables. These techniques can be divided into two including Model free techniques and Model based techniques. Model free techniques include analytical methods based on correlation, Mutual information and ad-hoc methods based on available data or domain range while pruning method, sensitivity analysis method, global optimization methods and other ad-hoc methods such as trial and error are part of Model based techniques. Selection of appropriate inputs is very important as inclusion of many inputs may cause to the over fitting of the model and exclusion of important input may cause development of model which fails to fully describe input and output relationship. (Maier *et al.*, 2010)(Wu, Dandy and Maier, 2014)

For this particular model, correlation analysis was conducted and relationship between the variables were determined in Chapter 4. Hence the input variables for the straight section of North South line are Number of trips, Tonnage, Speed and Grade while for the curved sections are Number of trips, tonnage, Speed, Curve radius and Grade. The output variable is the change in longitudinal leveling for both straight and curved sections of North South line.

i. Data Splitting

The selected data which includes the input variables and output variables are called the Selected data (Unprocessed). In this part the selected data is processed (scaled, lagged) to transform the data into a suitable form for application to the next step which is input selection. The Processed Data are divided into three categories Training data, testing data and Validation data. The training data is used in the learning the characteristics of data and assigning the synaptic weights to the connectors. Testing data is used to determine the optimum structure of the network and to determine when to stop training of data. Validation data is used to check the generalization ability of the developed model. Approach followed in the division of data can be either supervised approach or unsupervised approach. Supervised data division ensures that the statistical properties of data within a group are similar. Trial and error and data splitting based on optimization are examples of Supervised approach. Unsupervised approach does not take into account the similarities of statistical properties of data within a group. Dividing data based on underlying physical domain or process, Random data splitting and Self organizing maps are examples of Unsupervised approach. In this particular study the Supervised approach was selected so as to retain the statistical properties within the group. Trial and Error method was selected to ensure that optimum division of data is achieved.

ii. Model Architecture Selection

In this stage the overall structure of the model and information flow in the model are determined. The mostly used ANN structure is Feed forward which information is only fed in one direction from input layer to output layer. An example of Feed forward architecture is Multi-layer Perceptron (MLP). MLP consists of one input layer, one or more hidden layer and output layer.

The nodes in the input and output variables solely depends on the number of input and output. Other Feed Forward architecture includes Generalized Regression neural network (GRNN), cascade forward networks (CFN), radial basis function networks (RBF), modular neural networks (MNN), associative memory networks (AMN), reformulated neural network (ReNN) and probabilistic neural networks (PNN). Transfer of information from one layer to another is done by Activation function or transfer function which performs the network calculations. Activation functions can be classified in four categories namely piece-wise linear, unit step or threshold, sigmoid and Gaussian. The output activities are proportional to total weighted output in the piece-wise linear activation function. For threshold units the outputs are set at one of two levels, depending on whether the total input is greater than or less than some threshold value. Sigmoid functions have more function resemblance to neurons than linear or threshold. The output varies according to the variation of input in the non-linear fashion. Gaussian functions are bell shaped curves which are continuous. The most common Gaussian function is hyperbolic 'tansig' function. TanH function is a zero centered activation function and it is used to model inputs which are strongly positive or strongly negative even neutral.(Maier *et al.*, 2010; Wu, Dandy and Maier, 2014)

For this study Multi-layer Perceptron was selected as model architecture and Tan sigmoid (TANSIG) function was selected as activation function for hidden layers while a linear transfer function (PURELIN) is used in output layer.

iii. Model Structure Selection

Model Structure Selection; aim at finding balance of network complexity and generalization ability of the model. This part focuses on selection of number of hidden nodes, number of nodes per layer and number of layers. Different methods are applied in selection of model structure. The methods include global optimization method, step wise method (e.g. Pruning and constructive), trial and error method and regularization method. Selection of model structure should be done with care; a very complex structure calibration of model will be difficult and if a very simple model structure is selected the input-output relationship will not be well captured.

Trial and error method were selected and applied in selection of model structure since this method gives the modeler the ability to try different structures until optimum results are obtained.

iv. Model Calibration

Model Calibration is also known as Model training which deals with determination of model parameter values (weights) which are able to provide correct input-output relationship through given function form. The model calibration is a very complex process which can affect the model outcomes hence optimization algorithms are used to select model calibration. Local and global optimization algorithms are mostly applied in calibration of ANN model. Local method includes back propagation and Newton method while Global method includes generic algorithm. The Back propagation is commonly used in the training of MLP. It works in two phases; the first phase is called the forward phase in which the connection weights of network are fixed and the input signal is passed through each layer the output of this phase is error calculation. In second phase named Backward phase the calculated error is passed through the network in backward direction. In this phase the error is minimized through adjustment of the connection weights. Back propagation is selected in this study since it fits the model requirement and its popularity in many ANN studies

v. Model Validation

Model validation involves checking the produced model does not include flaws and can be applied in the required activity. Different statistical methods can be used in the validation of the model such as goodness of fit (R2), the sum of squared errors (SSE), root mean square error (RMSE), sum of absolute deviations (TSAD) and the mean sum of absolute deviations (MSAD). The goodness of fit and Mean squared error (MSE) are used in the validation of the ANN model.

Mean squared Error (MSE) is the average squared difference between the estimated values and the actual value. Almost always strictly positive (and not zero) is because of randomness and small values indicate better model. The formula governing the MSE value is shown below.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2 \qquad (5.7)$$

Where:

MSE= Mean Squared Error, n= Number of data points, Y_i= Observed values and $\hat{Y}_i$= Predicted Values.

vi. Model Application

ANN model is applied in the prediction of the track geometry degradation by using inputs to this model as Number of trips and tonnage for straight sections of North south line and obtaining the influence of mentioned variables on geometry parameter namely longitudinal leveling. Number of trips, Tonnage and Curve radius are inputs for the curved sections of North south line and are used in prediction of longitudinal leveling changes. At last, 4 ANN models are developed in which 2 model are for straight sections of Uplink and Downlink lines respectively and the other 2 models are for curved sections of Uplink and Downlink lines of North South line of AART.

3.4 Models Comparison

Regression Model and Artificial Neural Network model results were checked and compared to determine the suitable model for the track geometry prediction of Addis Ababa Light rail Transit. The simple model (Regression model) and a complex model (ANN) were critically analyzed on their ability to produce degradation model. The accuracy of each model were critically evaluated and compared with the other. The proposed model can further be applied in the optimization of the inspection and maintenance practice of the AALRT and can also be used as the base to model other parts of track structure.

CHAPTER FOUR

RESULTS AND DISCUSSION

4.1. Results

In this particular chapter factors influencing track geometry degradation followed with development of degradation models' results are discussed. Two models were developed to predict the track geometry degradation of North South line of AALRT. Since the N-S line is double line model development was done separately for the Uplink line and downlink line. Hence for each method used four models were developed to specifically describe the behavior of the track degradation. Such models are divided into two for straight sections of the lines and two model for curved sections. Correlation analysis and ANOVA analysis results are illustrated in Tables 4-1 to Table 4-11. The results for Multiple regression models are illustrated in Tables 4-13, 4-14 and 4-15 while the results for Artificial Neural Network are illustrated in Tables 4-17 and 4-18. For each model the model performance will be described followed by discussion of the results and finally the comparison of results of two models will be done.

4.2.Factors Influencing Track Geometry Degradation

Light rail systems are characterized by running over the streets, with semi exclusive (or reserved) right-of-way. The fact that the vehicle runs at street level (which is very beneficial for accessibility), leads to the necessity of adapting the track geometry to that existing in the streets, which sometimes gives rise to great challenges: strong vertical grades (rising or falling), horizontal circular curves with very small radius without super-elevation and transition curves, vertical curves with very small minimum radius, and so forth.(Novales, Orro and Bugarín, 1887)

Understanding the degradation mechanism of railway track is a crucial step in planning and optimizing the maintenance activities. Before development of degradation pattern of track there is a need of studying the factors causing such degradation. Various factors have been identified by a number of researchers to contribute to track degradation and some of them are evaluated by this study. Factors discussed in this study are number of trips, curve radius, track surface and tonnage.

4.2.1. Number of trips

Number of trains passing through the track segment is termed as number of trips. This is one of operational data which affects the long term behavior of the track(Mundrey, 2003). 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. Five Months data for each year 2018 and 2019 were collected and are applied to analyze the effect on such in track geometry degradation. Number of trips 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 number of trips and change of longitudinal leveling defect. The correlation measures how the variables relate. SPSS Software was used to perform correlation analysis for straight and curved section of N-S line of AALRT. The correlation was conducted separately on Uplink and Downlink line of North South line. Table 4-1 and Table 4-2 shows the correlation between changes in longitudinal leveling and number of trips for straight section and Curved section of line for most recent year 2019.

Table 4-1: Correlation between changes in longitudinal leveling value and number of trips for straight section of line for 2019

Location	Correlation value	Significance level (P-value)
Uplink line	0.147	0.014
Downlink line	0.481	0.000
North South line	0.362	0.000

From Table 4-1 shows that there is positive correlation between number of trips and longitudinal leveling change for Straight section of line. Strong correlation between the variables is observed on Downlink line of 0.48 with significance level of 0. In total obtained correlation values imply that as number of trips increases the change of longitudinal leveling also increases. The P- value for straight section shows to be less than 0.05 hence the number of trips and change in longitudinal leveling are statically significant and are linearly correlated.

Table 4-2: Correlation between changes in longitudinal leveling value and number of trips for Curved section of line for 2019.

Location	Correlation value	Significance level (P-value)
Uplink line	0.218	0.000
Downlink line	0.392	0.000
North South line	0.236	0.000

From Table 4-2 shows that there is positive correlation between number of trips and longitudinal leveling change for Curved section of line. As all correlation values are positive, it implies that as number of trips increases the change of longitudinal leveling also increases. All obtained P-value are 0 which implies their correlation between variables is significant. Hence number of trips and change in longitudinal leveling are statically significant and are linearly correlated.

4.2.2. Curve Radius

Addis Ababa Light Rail transit has a total of 87 curved section, 41 curves in 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 line with a total length of 6.59Km or 39.8% of total line. The maximum and minimum curve radius in North South line is 2004m and 50m respectively. The maximum and minimum curve radius in East West line is 3004m and 154m respectively. The critical curve radiuses observed in N-S line are considered in this study and are listed in table 4-3 below.

Table 4-3: Critical Curves in North South Line of AALRT

S/N	**Location**	**Curve Radius (m)**
1	Kaliti	150
2	Abo Junction	150
3	Stadium	50
4	St. Lideta	50
5	Darmar left turn	100
6	Darmar right turn	100
7	Autobistera	50

8	Menilik II Square underground	65

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 longitudinal leveling correlation analysis was conducted. The analysis was conducted only on the curved section of AALRT where the effect can be observed. The results obtained in the analysis are shown in table 4-4and table 4-5 below.

Table 4-4: Correlation between changes in longitudinal leveling value and curve radius for curved section of line for 2019.

Location	Correlation value	Significance level (P-value)
Uplink line	-0.109	0.050
Downlink line	-0.236	0.009
North South line	-0.147	0.002

Curve radius is only applicable on the curved section of N-S line. The results illustrated in Table 4-4 shows that there is negative correlation between curve radius and change in longitudinal leveling. This implies that as curve radius increases the change in longitudinal leveling decreases. The obtained P- values are less than 0.05 which shows the correlation between the variables is significant. Relationship between Change in longitudinal leveling and Curve radius is statically significant and are linearly correlated

4.2.3. Tonnage

Tonnage has been found by different researchers to be the main contributor on track geometry deterioration. Tonnage is a total load which passes on a track including the weight of the trains and passengers. Daily data of number of passengers transported in each line was collected from AALRT offices and was used to generate the total tonnage passing through the North South line of AALRT. The calculations of daily tonnage passing through North South line corresponding to the inspection data dates was done so as to perform correlation between the two. Example of such calculation is shown in Table 4-5 for the month of June 2018.

Table 4-5: Calculation done to obtain tonnage passing on North South line on June 2018.

Item	Value
Passenger flow (a)	1207639
Average weight of passenger (b)	65Kg= 0.065tons
Passenger flow in tons $c = a * b$	78,496.54 tons
Weight of trains (d) • Consider empty vehicle	= 44 tons
Total weight on the line $g = c + d$	78,540.535 tons
Total weight in MGT	0.078541 tons

The above calculation was done on all the available data to obtain the tonnage accommodated on the line for each month for the six-month data from June 2018 to September 2019. The correlation analysis between the obtained Tonnage (Million gross tons) and change in longitudinal leveling was carried out. The obtained results are shown in tables below.

Table 4-6: Correlation between changes in longitudinal leveling value and tonnage for straight section of line for 2019.

Location	Correlation value	Significance level (P-value)
Uplink line	0.126	0.034
Downlink line	0.696	0.000
North South line	0.376	0.000

Table 4-6 shows the correlation between the change in longitudinal leveling and Tonnage for straight sections of line for 2019. There is positive correlation between the variables with higher correlation observed in Downlink line with a value above 0.6. Significant values for all lines are less than 0.05 which implies that correlation between variables is significant. In conclusion the relationship between change in longitudinal leveling and Tonnage are statically significant and are linearly correlated for straight sections of line.

Table 4-7: Correlation between changes in longitudinal leveling value and tonnage for curved section of line

Location	Correlation value	Significance level (P-value)
Uplink line	0.128	0.021
Downlink line	0.217	0.018
North South line	0.135	0.000

Table 4-7 shows obtained correlation values range between 0.1 and 0.2. In summary there is positive correlation between change in longitudinal leveling and tonnage. The significance levels are less than 0.05. In conclusion, the change in longitudinal leveling and tonnage are linearly correlated and statically significant for straight and Curved Section of Uplink line.

4.2.4. Track Surface

The track surface of a railway track can be classified into two ballast track surface and slab track surface. With consideration of cost ballast track is most common type of track surface. AALRT incorporates both types in which ballast track is used on ground level and slab track is used on underground and elevated section. AALRT uses ballasted track bed of 300mm thick single layer gravel ballast with pre-stressed concrete sleeper with minimum concrete grade of C60 while in elevated section, the track bed is monolithic track bed on vertical support rail-bed, made of C40 concrete, 800mm wide. In underground section, the track bed is fully-laid monolithic track bed made of C30 concrete. Concrete short sleeper has a minimum concrete grade of C50. The N-S Line has a total length of 16.693km, from which 10.075km lies on the subgrade section, 5.944km elevated section, and the rest 0.655km in the underground section. This line has 22 stations, 8 elevated (including 5 common-track stations), 1 underground and the rest on the ground level.

There are total of 12 highway-rail grade crossings, 7 in the East-West direction (at ayat, summit roundabout, cmc, salitemihret roundabout and gurdshola) and 5 in the North-South direction (at kaliti, saris, adey abeba and meshualekia round about).

Under load, the track structure gradually deteriorates due to dynamic and mechanical wear effects of passing trains. Improper drainage, unstable roadbed, inadequate tamping, and deferred maintenance can create surface irregularities. Track surface irregularities can lead to serious consequences if ignored.

Investigation of influence of track surface on changes of longitudinal leveling was done using SPSS Software. Ballasted and concrete track surface are the only consideration in this study. Track surface is considered as a categorical variable in this analysis. Categorical variable is a variable which has fixed number of possible values and allows assigning these values in categories. One-way analysis of variance (ANOVA), which analyses if there is statically significance difference between the means of two or more independent groups. The relationship between the track surface and change in longitudinal leveling for curved and straight section of N-S line of AALRT are shown in table 4-5.

Table 4-8: ANOVA for changes in longitudinal leveling and Track surface for straight section of line for 2019.

Location	F- Value	Significance level (P-value)
Uplink line	0.042	0.999
Downlink line	6.208	0.000
North South line	2.296	0.020

Table 4-8 illustrate ANOVA for change in longitudinal leveling and track surface between the changes in longitudinal leveling and track surface for Straight section of line. Large F-value of 6.2 is observed on the Downlink line. The significance level for Downlink line and North South line are less than 0.05 while for Uplink line is greater than 0.05. There is significant correlation between change in longitudinal leveling and track surface for Down link line and North South line. Opposite is observed for Uplink line in which F-value less than 0 and P-values greater than 0.05 is obtained.

Table 4-9: ANOVA for changes in longitudinal leveling and Track surface for Curved section of Uplink line for 2019.

Location	F- Value	Significance level (P-value)
Uplink line	4.81	0.029
Downlink line	0.672	0.414
North South line	1.789	0.182

ANOVA results between the changes in longitudinal leveling and track surface for curved section of line are illustrated in Table 4-9. Track surface has impact in modeling of longitudinal leveling change for Uplink line for curved section with F-value of 4.81 and significance level of less than 0.05. Significance levels for Downlink line and North South line are greater than 0.05 which implies there is no statically association between these two variables for these two lines. Due to the existence of significance of variables in one line and absence of significance of variables in another line makes its hard to apply track surface in track geometry degradation modeling.

4.2.5. Speed

Long term behaviour of the railway track is mainly affected by speed and tonnage (Mundrey and Consultant, 1993). 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 rail vehicle allowed to pass through the section is used and correlation with change in longitudinal leveling are calculated. Table 4-10 and Table 4-11 shows correlation between longitudinal leveling and Speed for straight and curved sections of line respectively.

Table 4-10: Correlation between changes in longitudinal leveling value and Speed for Straight section of line for 2019.

Location	Correlation value	Significance level (P-value)
Uplink line	0.264	0.00
Downlink line	0.489	0.00
North South line	0.353	0.00

Table 4-10 shows correlation between longitudinal levelling change and speed for straight sections of the line. Downlink line have shown better correlation value compared to others (Uplink and North South line) with the value of 0.4. All obtained values have significance level of 0 which is lower than 0.05. The conclusion drawn is there is positive correlation between longitudinal leveling and speed and these values are statically significant.

Table 4-11: Correlation between changes in longitudinal leveling value and Speed for curved section of line for 2019.

Location	Correlation value	Significance level (P-value)
Uplink line	0.637	0.00
Downlink line	0.280	0.02
North South line	0.453	0.00

Positive correlation is observed between longitudinal leveling and Speed for the curved sections of line as illustrated in table 4-11 above. Uplink line and North South line have depicted better correlation values of above 0.4 compare to Downlink line. All obtained significance levels are less than 0.05 which implies the correlation values are linearly significant.

4.2.6. Grade

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.(Novales, Orro and Bugarín, 1887). Addis Ababa Light rail is designed with maximum gradient of 50‰. The gradient of level grade in the section should not be less than 3‰. The North South line profile drawing was used to correlated the grade of sections of line with the inspection data. Table 4-12 and Table 4-13 investigate the correlation between longitudinal leveling changes and grade for straight and curved sections respectively.

Table 4-12: Correlation between changes in longitudinal leveling value and Grade for Straight section of line for 2019

Location	Correlation value	Significance level (P-value)
Uplink line	0.235	0.00
Downlink line	0.460	0.00
North South line	0.319	0.00

Table 4-12 shows correlation between changes of longitudinal leveling values and Grade for the straight section of the line. Positive correlation is observed between variables and better correlation is observed on Downlink line (0.46) compared to other directions. All P- values are less than 0.05 hence it can be concluded that the two variables are positively correlated and the correlation values are statically significant.

Table 4-13: Correlation between changes in longitudinal leveling value and Grade for curved section of line for 2019.

Location	Correlation value	Significance level (P-value)
Uplink line	0.470	0.00
Downlink line	0.560	0.00
North South line	0.371	0.00

From Table 4-13 it is observed that there is positive correlation between changes in longitudinal leveling values and grade for Curve sections. All obtained correlation values are greater than 0.3 with highest value observed on Downlink line (0.6). All P-values are 0 which is less than 0.05. It can be concluded that the two variables are positively correlated and the correlations are statically significant.

The tables 4-14 and 4-15 provide the summary of the relationship between the dependent and independent variables. Table 4-14 gives the summary of correlations between changes in longitudinal leveling and independent variables (number of trips, tonnage, Speed, Grade and curve radius) for straight and curved sections of North South line of AART. Table 4-15 gives the summary of ANOVA between changes of longitudinal leveling and Track surface for straight and curved sections of North South line of AART.

Table 4-14: Correlations between changes in longitudinal leveling and independent variables for straight and curved section for years 2018 and 2019

Direction	Variable	Correlation Values				P-Values			
Uplink Line	Year	2018		2019		2018		2019	
	Track geometry	Straight (S)	Curved (C)	Straight (S)	Curved (C)	S	C	S	C
	Number of Trips	0.148	0.214	0.147	0.218	0.024	0.00	0.014	0.00
	Tonnage	0.229	0.215	0.126	0.128	0.00	0.00	0.034	0.021
	Speed	0.227	0.529	0.264	0.637	0.01	0.00	0.00	0.00
	Grade	0.392	0.458	0.235	0.470	0.00	0.00	0.00	0.00
	Curve Radius	--	-0.137	--	-0.09	--	0.012	--	0.017
Downlink Line	Number of Trips	0.525	0.356	0.481	0.392	0.00	0.002	0.00	0.00
	Tonnage	0.213	0.375	0.696	0.217	0.048	0.001	0.00	0.018
	Speed	0.669	0.533	0.489	0.280	0.00	0.00	0.00	0.02
	Grade	0.507	0.404	0.460	0.560	0.00	0.00	0.00	0.00
	Curve Radius	--	-0.317	--	-0.236	--	0.006	--	0.009

Table 4-15: ANOVA between changes in longitudinal leveling and Track Surface for straight and curved section for years 2018 and 2019

Direction	Variable	F Values				P-Values			
Uplink Line	Year	2018		2019		2018		2019	
	Track geometry	Straight (S)	Curved (C)	Straight (S)	Curved (C)	S	C	S	C
	Track Surface	1.755	14.386	0.042	4.810	0.187	0.00	0.999	0.029

Downlink Line	Track Surface	16.383	0.939	6.208	0.672	0.00	0.336	0.000	0.414

Table 4-14 shows that Number of trips, Speed, Grade and Tonnage are positive correlated to changes in longitudinal leveling and are statically significant. There is negative correlation between the changes in longitudinal leveling and curve radius and the variables are statically significant but can only be applied in curved sections of line. Table 4-15 shows there is correlation between changes in longitudinal leveling and Track surface for the curved section of uplink line are statically significant while there is correlation between variables for straight sections of Downlink line are statically significant. Figure below shows the influencing variables on changes in longitudinal leveling for straight and curved sections of Uplink and Downlink lines.

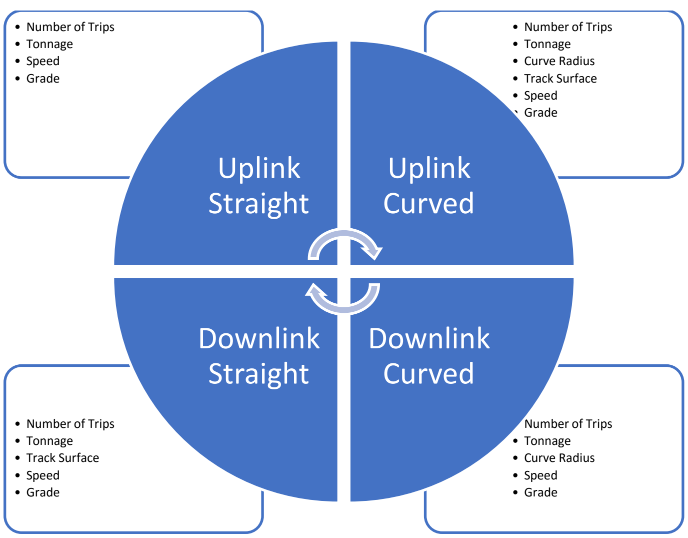

Figure 4.2: Summary of Influencing variables on longitudinal leveling Degradation

4.3 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. Collinearity of independent variables checks for correlation between the independent variables. For this study Number of trips, Tonnage and Curve radius are independent variables used. Table 4-16 illustrate the analysis of collinearity between the independent variables.

Table 4-16: Collinearity diagnosis between independent variables.

Variables	Tolerance	VIF
Number of Trips	0.668	1.496
Tonnage	0.487	2.052
Curve Radius	0.630	1.588
Speed	0.518	1.932
Grade	0.623	1.605

The Tolerance value obtain is greater than 0.1 and VIF value is less than 10. The analysis shows that there is neither collinearity nor multi-collinearity between independent variables.

For testing assumption of regression analysis five assumptions have to be satisfied by the regression analysis to get valid results. These assumptions are linearity, expected mean error is zero, Homoskedasticity, No autocorrelation and Error distribution. 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 N-S line. These relationships were checked using scatter plots and are illustrated in Figures below.

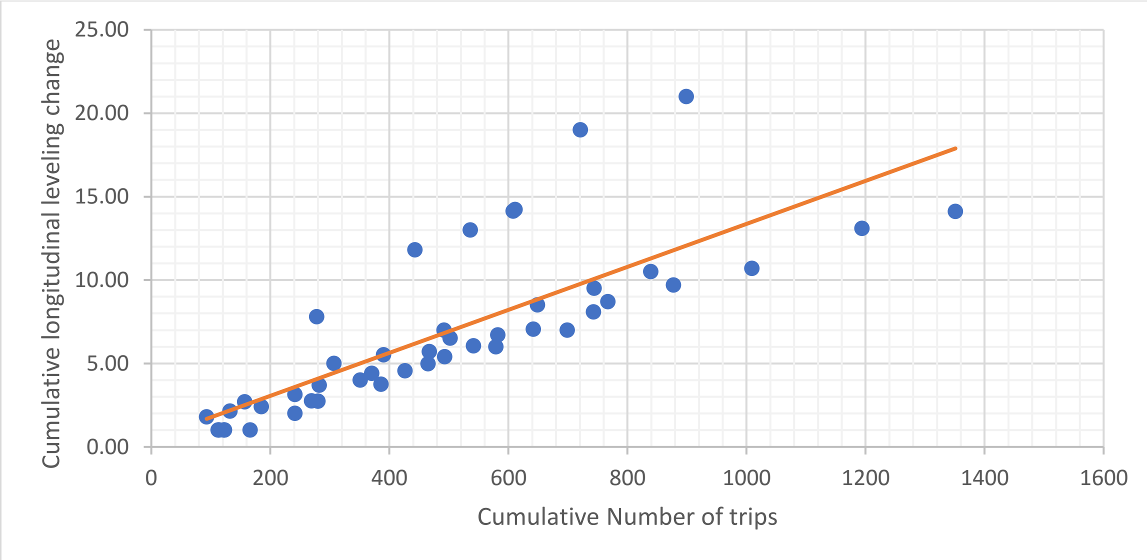

Figure 4.3: Relationship between longitudinal leveling and Number of trips for straight section of N-S Line

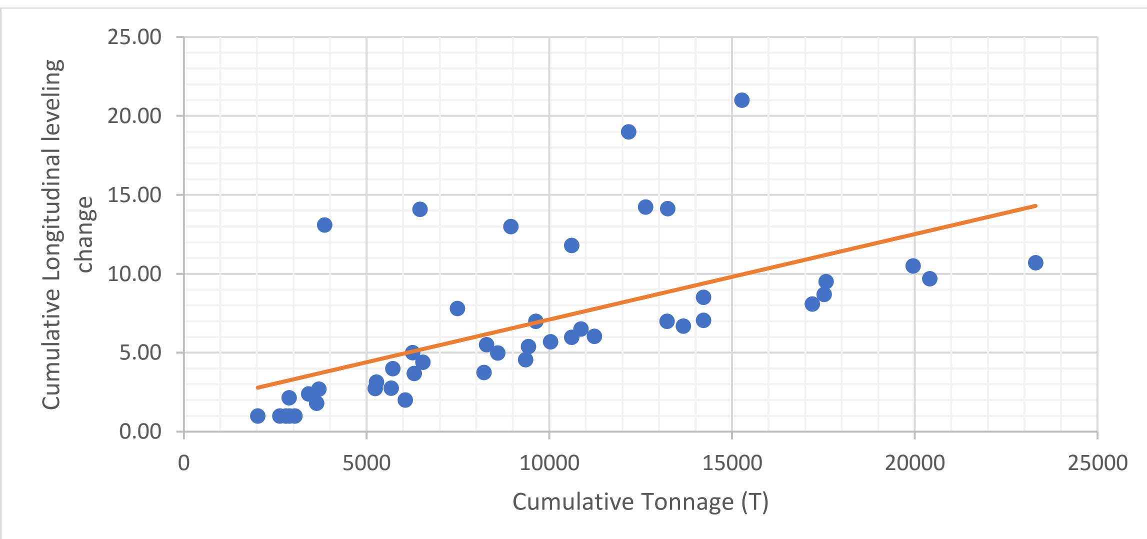

Figure 4.4: Relationship between longitudinal levelling and Tonnage for straight section of N-S Line

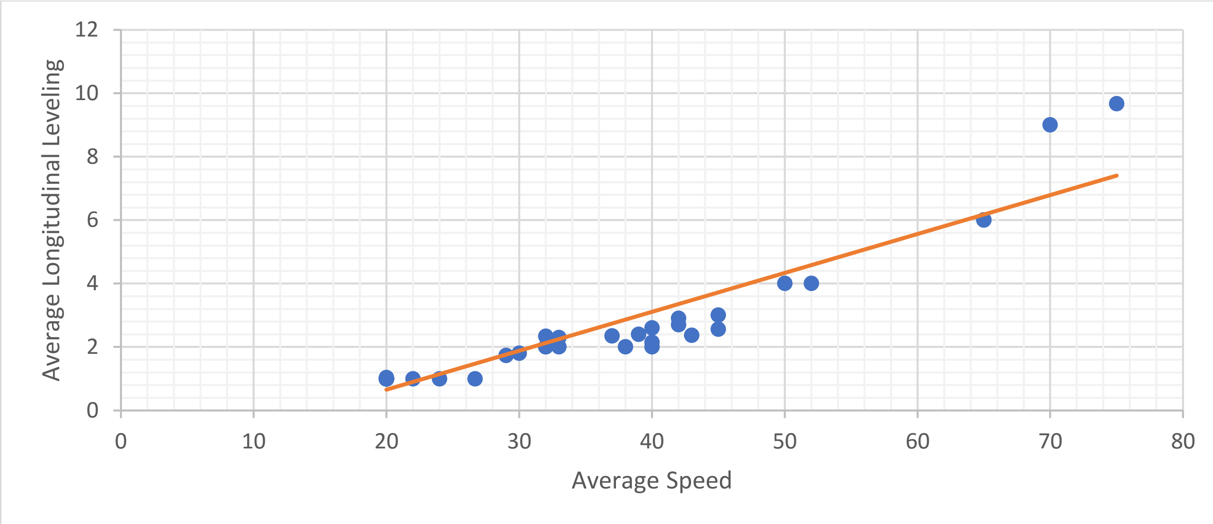

Figure 4.5: Relationship between longitudinal levelling and Speed for straight section of N-S Line

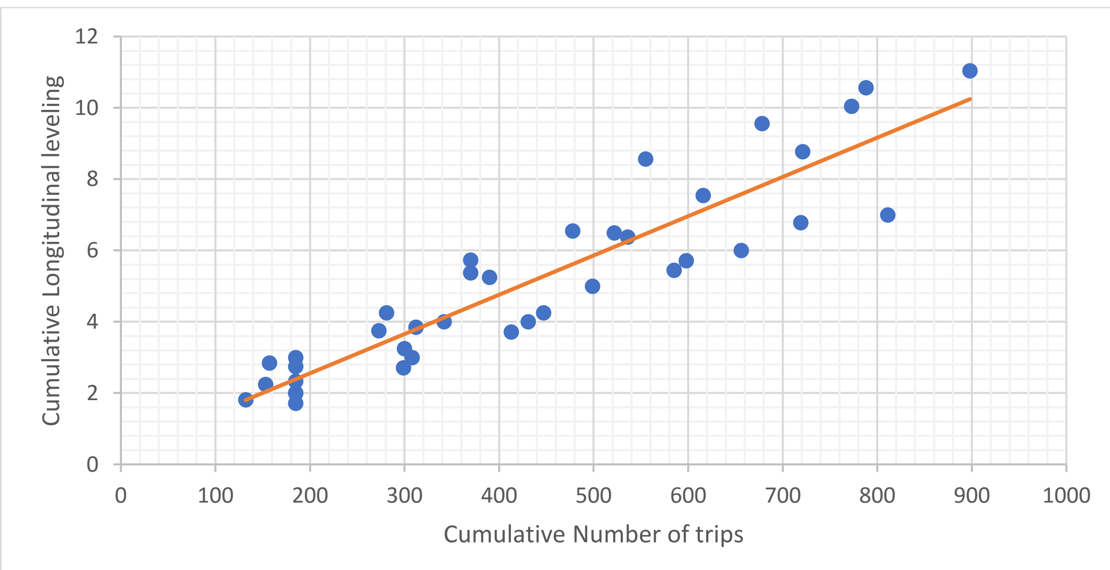

Figure 4.6: Relationship between longitudinal leveling and Number of trips for Curved section of N-S Line

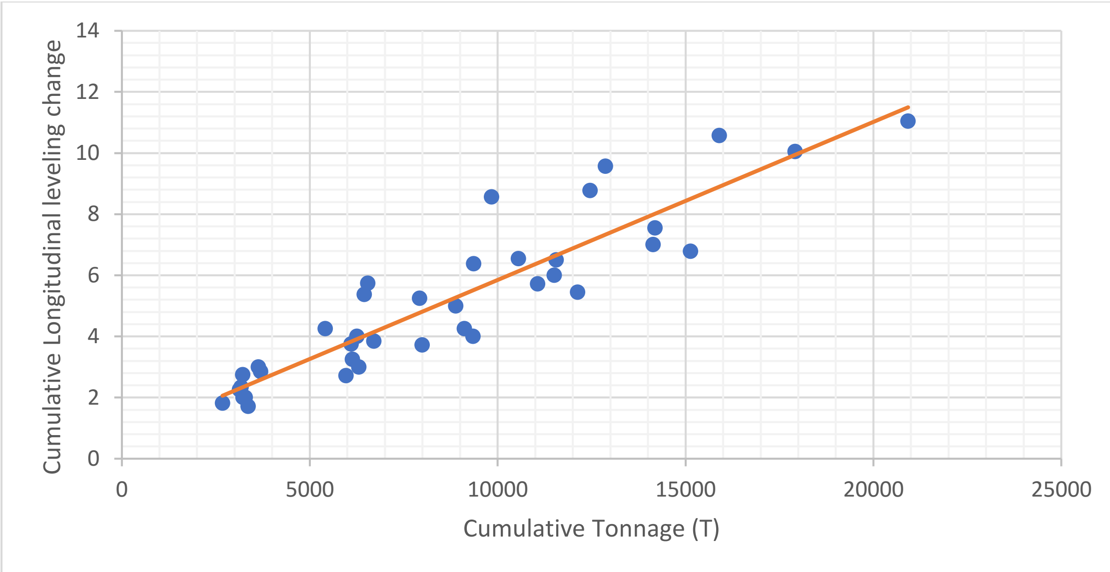

Figure 4.7: Relationship between longitudinal levelling and Tonnage for Curved section of N-S Line

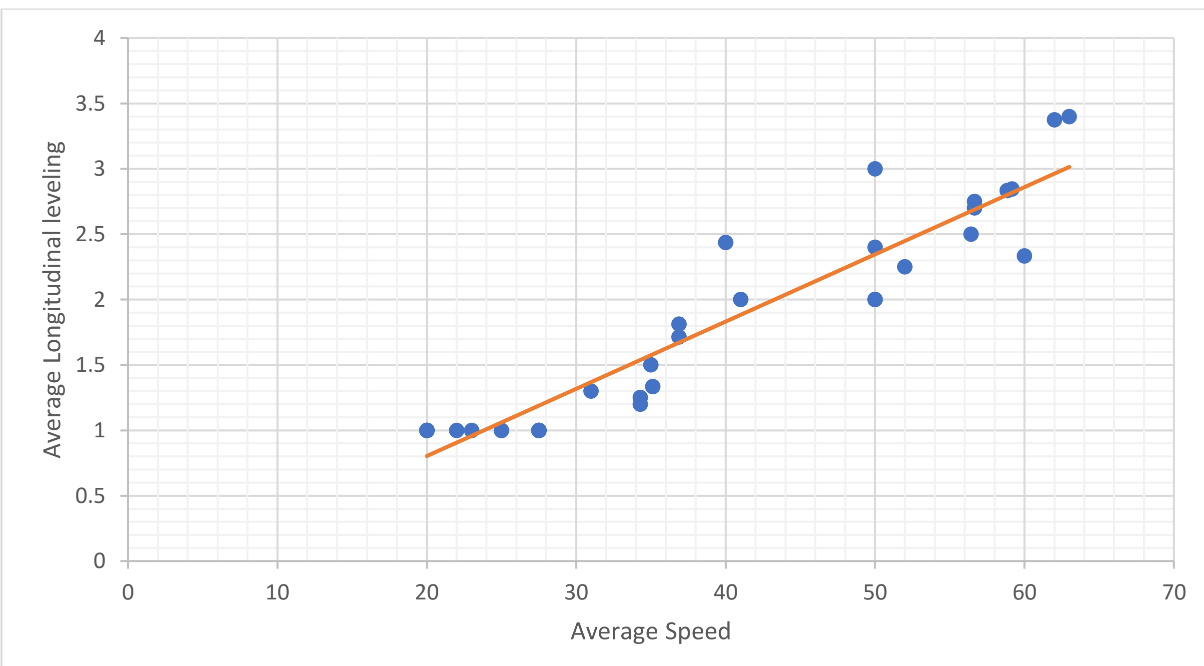

Figure 4.8: Relationship between longitudinal levelling and Speed for Curved section of N-S Line

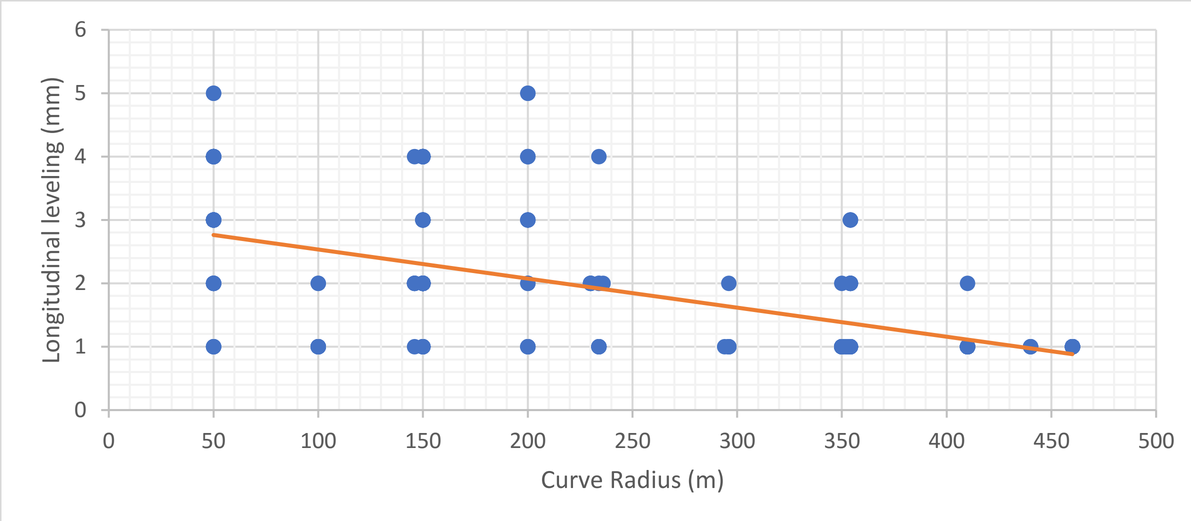

Figure 4.9: Relationship between longitudinal levelling and Curve Radius for Curved section of N-S Line

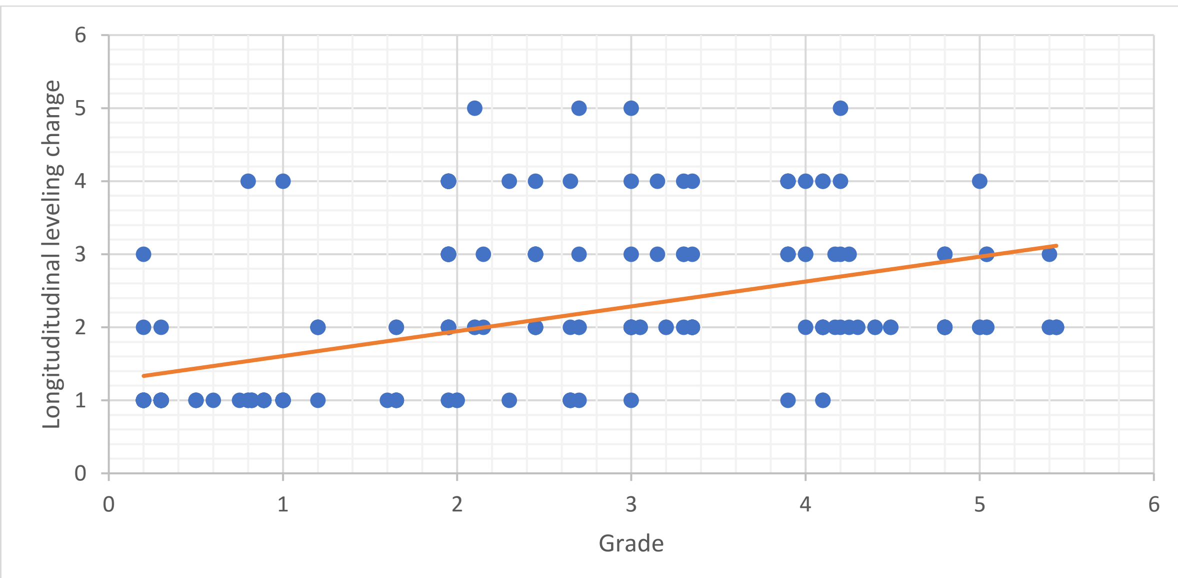

Figure 4.10: Relationship between longitudinal levelling and Grade for Curved section of N-S Line

There is linear relationship between the independent variables (i.e., number of trips, tonnage, Speed, Grade) and dependent variable (longitudinal levelling) for the straight and curved sections of N-S line of AALRT. With consideration of only the Curved Sections it was also observed that there is linear relationship between longitudinal leveling and Curved Radius as illustrated in Figure

4.7. Regression analysis can be applied in this study since there is linear relationship between independent and dependent variables.

After following all six procedures for developing regression model, the multiple regression model results for the track geometry degradation are illustrated in Table 4-17, Table 4-18 and Table 4-19 below. Table 4-18 shows the regression coefficients for the two developed models for straight section and Table 4-19 shows the regression coefficients for the remaining two models for the curved sections of North South line.

Table 4-17: Estimated variables from regression model for straight section

	Parameters	Values	t-statistics	Significance
Uplink	Constant (α)	-2.514	-9.238	0.00
	Regression coefficient (β_1)	3.53e-4	2.212	0.03
	Regression coefficient (β_2)	0.0356	7.209	0.00
	Regression coefficient (β_3)	0.0128	4.156	0.00
	Regression coefficient (β_4)	0.078	2.540	0.01
Downlink	Constant (α)	-7.702	-14.166	0.00
	Regression coefficient (β_1)	4.82e-4	1.351	0.01
	Regression coefficient (β_2)	0.0721	7.902	0.00
	Regression coefficient (β_3)	0.0158	1.186	0.03
	Regression coefficient (β_4)	0.103	1.184	0.00

Table 4-18: Estimated variables from regression model for Curved section

	Parameters	Values	t-statistics	Significance
Uplink	Constant (α)	-0.189	-0.556	0.009
	Regression coefficient (β_1)	8.15e-4	12.896	0.00
	Regression coefficient (β_2)	0.055	3.265	0.00
	Regression coefficient (β_3)	0.0247	1.728	0.008
	Regression coefficient (β_4)	0.105	3.004	0.00
	Regression coefficient (β_5)	-6.6e-4	-1.495	0.00

Downlink	Constant (α)	-0.732	-1.674	0.02
	Regression coefficient (β_1)	7.58e-4	7.176	0.00
	Regression coefficient (β_2)	0.068	2.347	0.02
	Regression coefficient (β_3)	0.0151	2.141	0.03
	Regression coefficient (β_4)	0.080	1.120	0.01
	Regression coefficient (β_5)	-9.31e-4	-1.324	0.02

For considering only the curved and straight section of line regardless of their direction, the regression coefficients were calculated and are shown in the Table 4-15 below.

Table 4-19: Estimated variables from regression model for North South line

	Parameters	Values	t-statistics	Significance
Straight	Constant (α)	-4.971	-16.507	0.000
	Regression coefficient (β_1)	7.58e-4	5.206	0.000
	Regression coefficient (β_2)	0.057	13.342	0.000
	Regression coefficient (β_3)	0.0118	3.772	0.000
	Regression coefficient (β_4)	0.131	4.130	0.000
Curved	Constant (α)	-0.543	-2.293	0.023
	Regression coefficient (β_1)	7.25e-4	11.780	0.000
	Regression coefficient (β_2)	0.078	5.134	0.000
	Regression coefficient (β_3)	0.055	2.396	0.018
	Regression coefficient (β_4)	0.162	1.484	0.002
	Regression coefficient (β_5)	-1.09e-4	-0.339	0.034
Full Line	Constant (α)	-0.521	-2.141	0.034
	Regression coefficient (β1)	6.88e-4	12.457	0.00
	Regression coefficient (β2)	0.073	4.945	0.00

	Regression coefficient (β3)	0.092	2.958	0.004
	Regression coefficient (β4)	0.085	1.272	0.017
	Regression coefficient (β5)	-2.6e-4	-1.768	0.038

The estimation equations obtained including the dependent variable (change in longitudinal leveling, independent variables (Number of trips, tonnage and Curve radius), regression coefficients and constants are shown below.

The model equation for straight sections;

Longitudinal leveling (t) = $\alpha + \beta_1 \cdot Tonnage\ (t-1) + \beta_2 \cdot Number\ of\ trips(t-1) + \beta_3 \cdot Speed + \beta_4 \cdot Grade$ (4.1)

After insertion of regression coefficients and constants the equations become;

Uplink of N-S Line

LL (t) = -2.514+3.53e-4 Tonnage (t-1) + 0.036 Number of trips (t-1) +0.013 Speed + 0.078 Grade

(4.2)

Downlink of N-S Line

LL (t) =-7.702+4.82e-4 Tonnage (t-1) + 0.072 Number of trips (t-1) +0.016 Speed + 0.103 Grade

(4.3)

Total straight section of North South line

LL (t) = =-4.971+7.58e-4 Tonnage (t-1) + 0.057 Number of trips (t-1) +0.012 Speed + 0.131 Grade

(4.4)

The model equation for Curved sections;

Longitudinal leveling (t) = $\alpha + \beta_1 \cdot Tonnage\ (t-1) + \beta_2 \cdot Number\ of\ Trips(t-1) + \beta_3 \cdot Speed + \beta_4\ Grade + \beta_5 Curve\ Radius$ (4.5)

After insertion of regression coefficients and constants the equations become;

Uplink of N-S Line

LL (t) = -0.189+8.15 e-4 Tonnage (t-1) + 0.055 Number of trips (t-1)+ 0.025 Speed+ 0.105 Grade - 6.64e-4 C.R(t-1) (4.6)

Downlink of N-S Line

LL (t) = -0.732+7.58 e-4 Tonnage (t-1) + 0.068 Number of trips (t-1)+ 0.015 Speed+ 0.080 Grade - 9.31e-4 C.R(t-1) (4.7)

Total Curved section of North South line

LL (t) -0.543+7.25 e-4 Tonnage (t-1) + 0.078 Number of trips (t-1)+ 0.055 Speed+ 0.162 Grade - 1.09 e-4 C.R(t-1) (4.8)

Full North South line

LL (t) -0.521+6.88 e-4 Tonnage (t-1) + 0.073 Number of trips (t-1)+ 0.092 Speed+ 0.085 Grade - 1.768 e-4 C.R(t-1) (4.9)

4.3.1 Model Performance

The model performance is checked by plotting the observed values against the predicted values for both straight and curved sections of N-S line and checking the model accuracy. Figure 4.8 and Figure 4.9 shows the graphs of observed against predicted longitudinal leveling values for straight and curved section respectively.

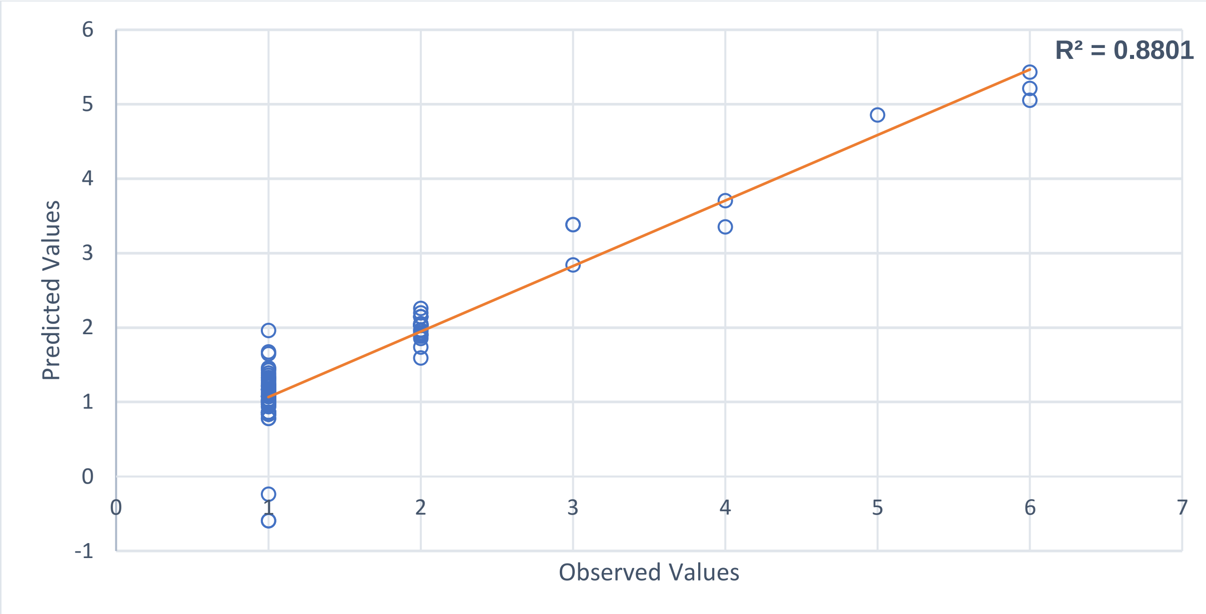

Figure 4.10: Observed values against Predicted values of Longitudinal leveling for straight sections

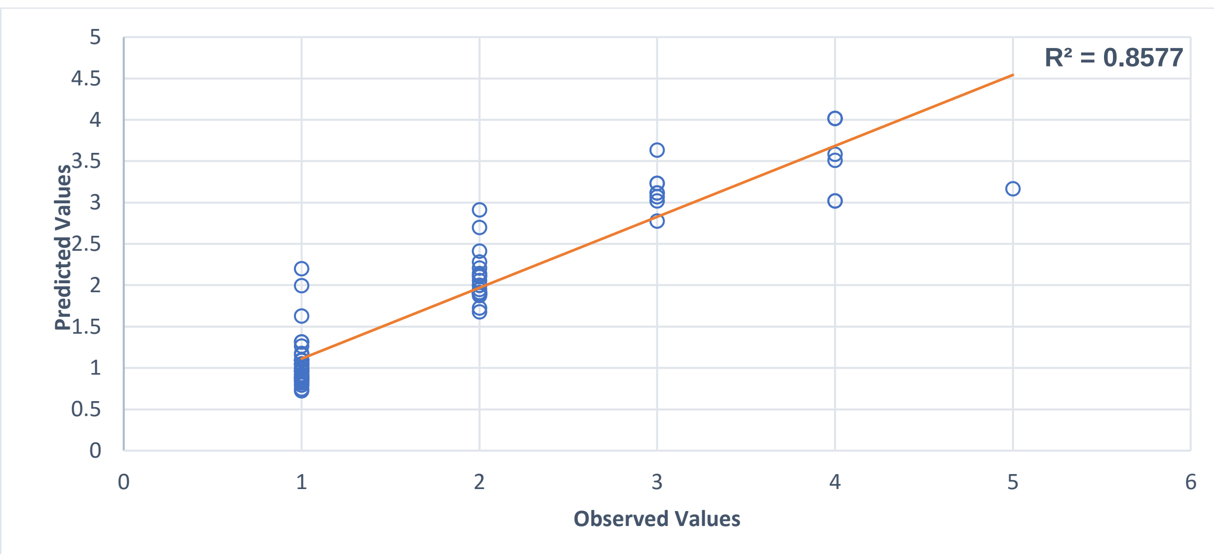

Figure 4.11: Observed values against Predicted values of Longitudinal leveling for curved sections

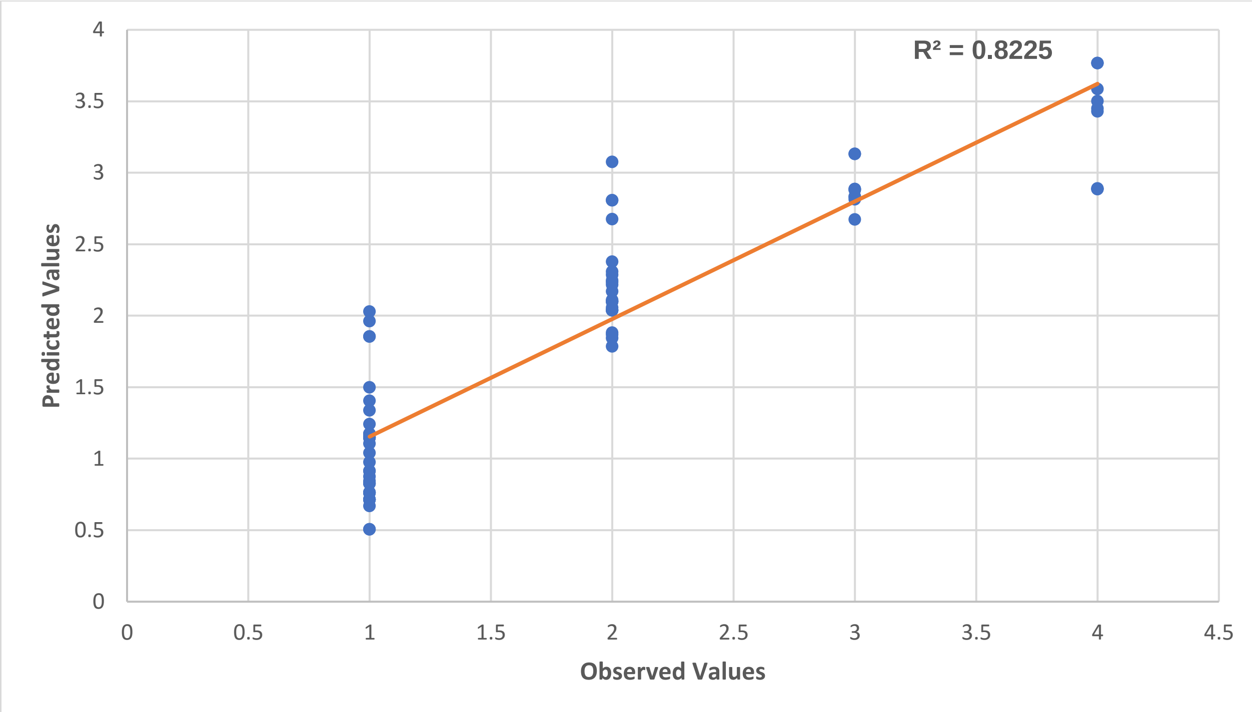

Figure 4.12: Observed values against Predicted values of Longitudinal leveling for North South line

Mean absolute error (MAE) and R- square were used to measure the developed multiple regression model performance for straight and curved sections. The results obtained are shown in table 4-16 below.

Table 4-20: Statistical Variables of Regression Model

Location		Variables		
		R^2	Adjusted R^2	MSE
Uplink	Straight	0.853	0.846	0.042
	Curved	0.906	0.896	0.088
Downlink	Straight	0.984	0.979	0.039
	Curved	0.904	0.876	0.078
North South line	Straight	0.880	0.875	0.134
	Curved	0.858	0.848	0.150
	Full line	0.823	0.810	0.183

4.4. Artificial Neural Network Results

A four-layer Artificial Neural network model including an input layer, two hidden layers and output layer was developed to predict track degradation of North South line. The optimum model was obtained by alternating the number of neurons in hidden layers. For first hidden layer (10,15,20) were used while for the second hidden layer (7,10,15) were used. Four ANN models were developed per direction (Uplink or Downlink line) in which two ANN models were for straight section and remaining for curved section. Lastly two ANN models were developed for North south line as whole with one for straight section and the other for curved sections. The ANN model results for straight sections of North South line are shown in Table 4-22 and for curved sections are shown in Table 4-23.

Table 4-21: Results of ANN model for Straight section of the line

Direction	First Hidden layer	Second hidden layer	R^2	MSE
Uplink line	10	7	0.945	0.0276
	15	**10**	**0.971**	**0.0170**
	20	15	0.923	0.0420
Downlink line	10	7	0.929	0.605
	15	10	0.897	0.853
	20	**15**	**0.956**	**0.274**

Table 4-22: Results of ANN model for Curved section of the line

Direction	First Hidden layer	Second hidden layer	R^2	MSE
Uplink line	10	7	0.959	0.056
	15	**10**	**0.970**	**0.033**
	20	15	0.931	0.097
Downlink line	10	7	0.835	0.199
	15	**10**	**0.927**	**0.073**
	20	15	0.841	0.193

Table 4-23: Results of ANN model for Curved section of the line

Direction	First Hidden layer	Second hidden layer	R^2	MSE
North South line (Straight)	10	7	0.943	0.134
	15	**10**	**0.958**	**0.077**
	20	15	0.913	0.173
North South line (Curved)	10	7	0.896	0.117
	15	**10**	**0.917**	**0.093**

	20	15	0.901	0.115
North South line	**10**	**7**	**0.962**	**0.060**
	15	10	0.861	0.222
	20	15	0.940	0.095

4.4.1 ANN Model Performance

Model accuracy or performance is check by plotting the observed values against estimated values. From table 4-21, 4-22 and 4-23 it can be deduced that best results are obtained when 15 neurons are used in first hidden layer and 10 neurons in the second layer for both straight and curved sections of North South line. Figure 4.12 and 4.13 shows the Observed values against predicted values of Longitudinal leveling for straight and curved sections of North South line respectively.

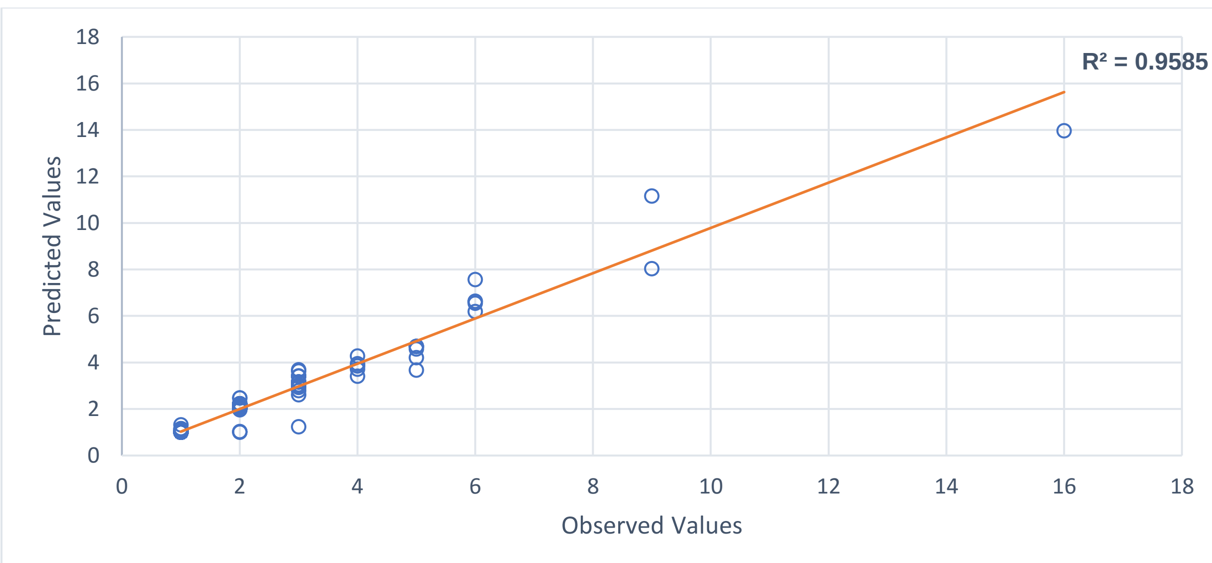

Figure 4.13: Observed values against Predicted values of Longitudinal leveling for straight sections

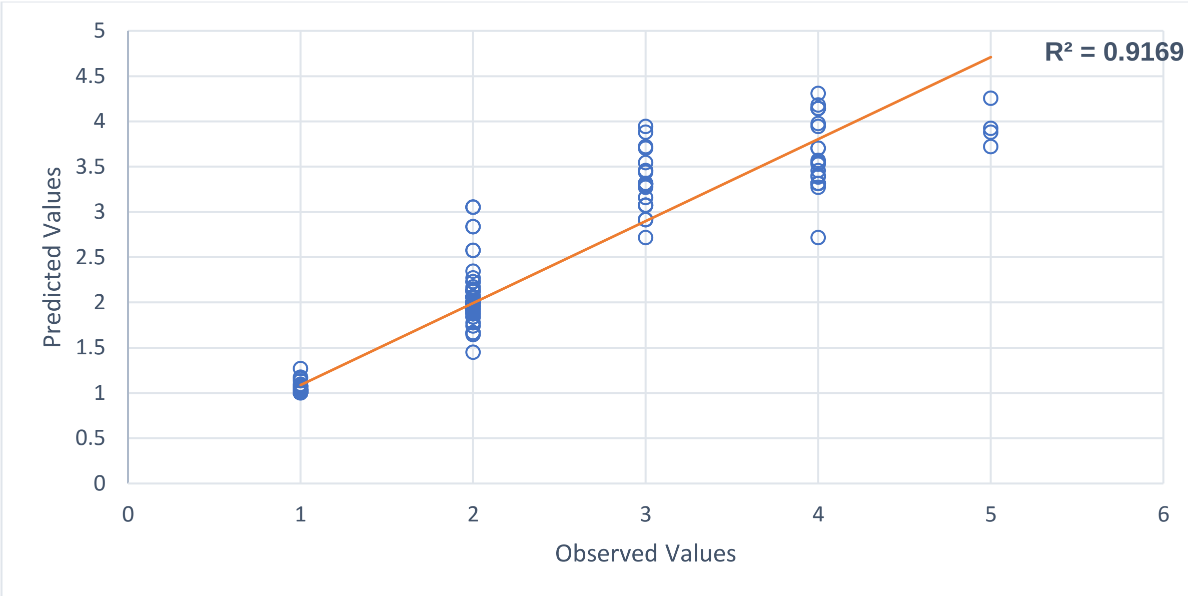

Figure 4.14: Observed values against Predicted values of Longitudinal leveling for curved sections

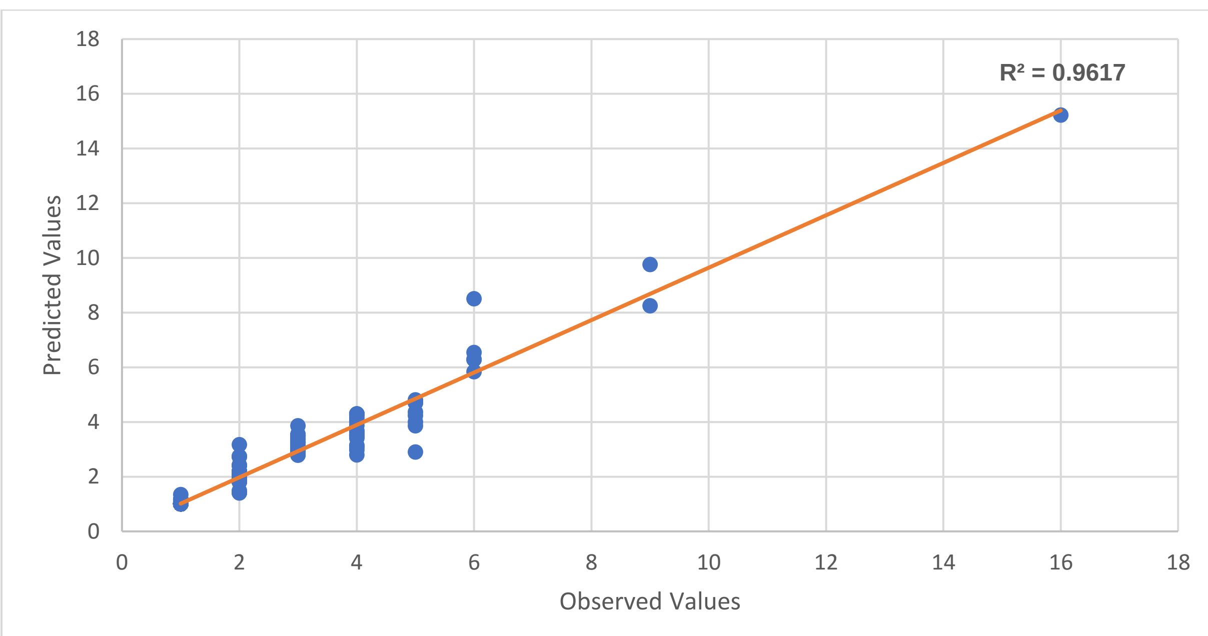

Figure 4.15: Observed values against Predicted values of Longitudinal leveling for North South line

4.5. Discussion

Two models are used in this study to predict the track geometry degradation of North South line of AALRT. Two statistical measures namely R-squared and Mean Squared Error are applied to check the accuracy of developed models. North South line is a double track line, (Uplink and Downlink line) development of models were done separately for each direction and as the whole line.

Regression analysis: Six models were developed in which three for straight section and remaining three for curved section. For Uplink line two models were developed where better performance is observed in curved section (R^2=90%) compared to straight section (R^2=85%). Contrary to Uplink line model, better performance is observed in straight section (R^2=98%) compared to curved section (R^2=90%) of Downlink line. Further, two models are developed for North South line regardless of direction. The better performance was observed on the straight section (R^2=88%) of line compare to curved section (R^2=86%). Lastly, A model for full North South line was developed in which the performance of 82% was observed. Variation in performance can be attributed by variation in number of independent variables used in the modeling. Also, presence of large amount of data for Curved sections compare to Straight section with noise data may also be another reason. In totality the developed models have good performance of above 85%.

Mean square error (MSE) is measures the average squared difference between estimated values and actual values. MSE is a non-negative value and values closer to zero are better and are used to measure the performance of the model. Refer to Table 4-20 obtained MSE values for Uplink and Downlink are very low (less than 0.1) compared to North South line (less than 0.2). In totality all obtained valued are small and well explains the variation of estimated values from the actual values.

Artificial Neural Network Analysis: Four-layer ANN model was used to develop track geometry degradation of North South line. To obtain optimum model number of neurons in two hidden layers are altered until better performance is achieved. 15 Neurons in First and 10 Neurons in Second hidden layer is an optimum ANN model for the straight and curved sections of Uplink line. The performance of 97% is achieved in these models. Downlink line show different optimum models where for straight section 20 Neurons in First and 15 Neurons in Second hidden layers while for curved section 15 Neurons in First and 10 Neurons in Second Hidden layers. Better performance is observed in Straight section (96%) compared to Curved sections (93%) of Downlink line.

North South line ANN optimum model included 15 Neurons in First hidden layer and 10 Neurons in Second hidden layer. Straight sections have shown performance of 96% while performance of 92% was observed on Curved Sections. For the full North South line optimum model included 10 Neurons in First hidden layer and 7 Neurons in Second hidden layer. This model had the performance of 96%.

4.6.Comparison of Models

The Statistical measurements are used to compare the performance of developed models. Through comparison a better model to be employed in track geometry degradation modeling will be known. Table 4-18 shows the comparison of statical variables between Regression model and ANN model.

Table 4-24: Comparison of Models

Location		Regression Model		ANN Model	
		R^2	MSE	R^2	MSE
Uplink	Straight	0.853	0.042	0.97	0.017
	Curved	0.906	0.088	0.97	0.033
Downlink	Straight	0.984	0.039	0.96	0.274
	Curved	0.904	0.078	0.93	0.073
North South	Straight	0.880	0.134	0.96	0.077
	Curved	0.858	0.150	0.92	0.093
	Full	0.823	0.183	0.96	0.060

Downlink regression models have better performance compared to ANN Models for straight and curved section of the line. Better performance is observed on Uplink and North South line ANN models compared to regression models. The ANN models have shown R-squared values greater than 90% which is implies the variation of predicted values and actual values is very minimum. MSE values for ANN models are lower compared to Regression model. This can be attributed by the ability of ANN to develop model based on data characteristics and no prior assumptions is required. Both Models have shown good performance of above 85% can both be applied to predict the future condition of track geometry degradation.

CHAPTER FIVE

CONCLUSIONS AND RECOMMENDATIONS

5.1 Conclusions

Longitudinal leveling defects are used by different railway organization around the world to determine the condition of the railway track. Optimization of inspection and maintenance practice will improve the operation efficiency of railway lines and keep the track geometry in acceptable levels. Maintenance actions are used to control the degradation of the track, reduce or eliminate the likelihood of failures, and restore a failed part to an operational state. Effective maintenance policy can be determined only when track degradation behavior is correctly modeled. Modeling the track behavior requires in-depth knowledge of track geometry degradation. In addition, there is still need of research on degradation pattern of light rails with application of different models due to most of developed models focus on heavy rails.

This study focused on analyzing and predicting the track geometry degradation of AALRT. Factors influencing the track geometry degradation for curved and straight sections of N-S line were evaluated first. SPSS Software was used to conduct correlation analysis between the change in longitudinal leveling and influencing factors. The following conclusions were drawn during development of models;

- Six influencing factors were analyzed in this study namely; Number of trips, Tonnage, Speed, Grade, Curve Radius and Track surface. The analysis was done separately for Uplink line and Downlink line. For the straight sections it was observed that Number of trips, Tonnage, Speed and Grade are influencing variables for Uplink line while Number of trips, Tonnage, Speed, Grade and Track surface were for Downlink line. All analyzed factors were found to influence longitudinal leveling change for curved sections of Uplink line. Number of trips, Tonnage, Speed, Grade and Curve radius were influencing variables for curved sections of Downlink line. For model development number of trips, Tonnage, Speed and Grade were used for Straight sections of lines while Number of trips, Tonnage Curve radius were employed for track geometry degradation modeling for curved sections of the line.

- Multiple regression model and Artificial Neural Network model were developed to predict future track geometry behavior. Multiple regression models were developed in SPSS while MATLAB was used to develop ANN models. Six regression models were developed with two for every direction (e.g., Uplink line and Downlink line) and remaining two models for North South line regardless of direction. Models' performance differs per direction. For Downlink line and North South line better performance was observed in Straight sections compared to Curved sections while opposite was discovered for Uplink line.
- ANN model was used as complex model to predict future condition of track. Six models were developed per direction. To obtain optimum model number of neurons in two hidden layers are altered until better performance is achieved. 15 Neurons in First and 10 Neurons in Second hidden layer is an optimum ANN model for the straight and curved sections of Uplink line. Downlink line show different optimum models where for straight section 20 Neurons in First and 15 Neurons in Second hidden layers while for curved section 15 Neurons in First and 10 Neurons in Second Hidden layers. Better performance is observed in Straight section compared to Curved section of Downlink line. North South line ANN optimum model included 15 Neurons in First hidden layer and 10 Neurons in Second hidden layer. All developed models had performance of greater than 90% for straight and curved sections.
- Comparison of two models showed that both can predict the longitudinal leveling changes with very low error, although the ANN model has better performance than Regression model.

In this study maintenance activities were not considered in the development of degradation model. Future studies, the proposed model can be extended to accommodate the maintenance activities and include other sections of track such as stations, terminus and turnout.

5.2. Recommendations

The degradation model was developed to predict the future track geometry behavior for the curved and straight section of AALRT. The following are recommendations based on the results of investigation in this research

- Addis Ababa LRT need to automate the inspection process by employing use of track inspection car which provide wide range of data with high precision. The track inspection car provides a detailed report on track quality, list of critical errors and diagrams showing trend of each track parameter (alignment, twist, gauge etc.). Infrastructure managers will be able to select most effective maintenance policy with detailed data set.
- Data management and compilation system of Addis Ababa LRT need to be improved. Future reference of track defect and accident data becomes difficult if data is stored in undigitized form. It is very hard to find any future plans for our new railway lines.
- Research department or section should be established at AALRT which will focus on evaluating the operation of track which includes forecasting the future condition of railway track hence selection of appropriate maintenance policy. Furthermore, this center will research appropriate technologies to be adopted to contain track deterioration. Inspection, Maintenance and restoration costs of the part of track can be determined through developed degradation model.

5.3. Future Work

- Maintenance optimization model can be developed for better maintenance planning.
- Regression model can be developed using other parameter such as gauge, twist or alignment etc.
- Development of track quality index for Addis Ababa Light Rail transit.
- Development of a maintenance plan model that will allocate available free slots of the track for maintenance with respect to RAMS parameters.
- Development of track structure deterioration model.

References

Agrawal, R. . and Ratnadip, A. K. (2013) ‘An Introductory Study on Time Series Modeling and Forecasting Ratnadip Adhikari R. K. Agrawal’, *arXiv preprint arXiv:1302.6613*, 1302.6613, pp. 1–68.

Ahmad, R. and Kamaruddin, S. (2012) ‘An overview of time-based and condition-based maintenance in industrial application’, *Computers and Industrial Engineering*. Elsevier Ltd, 63(1), pp. 135–149. doi: 10.1016/j.cie.2012.02.002.

American Railway Engineering and Maintenance of Way Association (2006) *Manual for Railway Engineering*. USA.

Andrade, A. R. and Teixeira, P. F. (2015) ‘Statistical modelling of railway track geometry degradation using Hierarchical Bayesian models’, *Reliability Engineering and System Safety*, 142, pp. 169–183. doi: 10.1016/j.ress.2015.05.009.

Andrews, J., Prescott, D. and De Rozières, F. (2014a) ‘A stochastic model for railway track asset management’, *Reliability Engineering and System Safety*, 130, pp. 76–84. doi: 10.1016/j.ress.2014.04.021.

Andrews, J., Prescott, D. and De Rozières, F. (2014b) ‘A stochastic model for railway track asset management’, *Reliability Engineering and System Safety*. Elsevier, 130, pp. 76–84. doi: 10.1016/j.ress.2014.04.021.

Audley, M. and Andrews, J. D. (2013) ‘The effects of tamping on railway track geometry degradation’, *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, 227(4), pp. 376–391. doi: 10.1177/0954409713480439.

Bai, L. *et al.* (2016) ‘Classification-learning-based framework for predicting railway track irregularities’, 230(2), pp. 598–610. doi: 10.1177/0954409714552818.

China, M. of R. of P. R. of (2010) *CODE FOR DESIGN OF RAILWAY TRACK*. Beijing.

Elkhoury, N. (2018) ‘Analysis and Prediction of Tram Track Degradation’, (April), p. 97.

Fekadu, D. (2017) ‘Addis Ababa Institute of Technology Addis Ababa Institute of Technology’, (June).

Francis, W. (2018) ‘Inspection and Maintenance of the Addis Ababa Light Rail Transit’, (November).

Guler, H., Jovanovic, S. and Evren, G. (2011) ‘Modelling railway track geometry deterioration’, *Proceedings of the Institution of Civil Engineers: Transport*, 164(2), pp. 65–75. doi: 10.1680/tran.2011.164.2.65.

Hasan, M. S. (2015) ‘Deterioration Prediction of Concrete Bridge Components Using Artificial Intelligence and Stochastic Methods’, (March).

Hawari, H. M. (2007) ‘Minimising Track Degradation through Managing Vehicle / Track Interaction’, *Development*, (February).

He, Q. *et al.* (2015) ‘Track geometry defect rectification based on track deterioration modelling and derailment risk assessment’, *Journal of the Operational Research Society*. Nature Publishing Group, 66(3), pp. 392–404. doi: 10.1057/jors.2014.7.

Higgins, C. and Liu, X. (2018) ‘Modeling of track geometry degradation and decisions on safety and maintenance: A literature review and possible future research directions’, *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, 232(5), pp. 1385–1397. doi: 10.1177/0954409717721870.

Iwnicki S, G. S. and K. W. (2000) ‘Track settlement prediction using computer simulation tools’, *Vehicle System Dynamics*, 33(13), pp. 37–46.

Kawaguchi, A., Miwa, M. and Terada, K. (2005) ‘Actual data analysis of alignment irregularity growth and its prediction model’, *Quarterly Report of RTRI (Railway Technical Research Institute) (Japan)*, 46(4), pp. 262–268. doi: 10.2219/rtriqr.46.262.

Kumar, U. *et al.* (2016) *Current Trends in Reliability , Availability , Maintainability and Safety*.

Lyngby, N., Hokstad, P. and Vatn, J. (2008) ‘RAMS Management of Railway Tracks’, *Handbook of Performability Engineering*, pp. 1123–1145. doi: 10.1007/978-1-84800-131-2_68.

Maier, H. R. *et al.* (2010) ‘Methods used for the development of neural networks for the prediction of water resource variables in river systems: Current status and future directions', *Environmental Modelling and Software*. Elsevier Ltd, 25(8), pp. 891–909. doi: 10.1016/j.envsoft.2010.02.003.

Maier, H. R. and Dandy, G. C. (2000) 'Application of Artificial Neural Networks to Forecasting of Surface Water Quality Variables: Issues, Applications and Challenges', pp. 287–309. doi: 10.1007/978-94-015-9341-0_15.

Mooi, E. (2016) *Regression Analysis*. doi: 10.1007/978-3-642-53965-7.

Muinde, M. S. (2018) 'Railway track geometry inspection optimization', p. 61. Available at: http://ltu.diva-portal.org/smash/get/diva2:1257286/FULLTEXT01.pdf.

Mundrey, J. . (2003) *Railway track engineering*. New Delhi: McGraw–Hill.

Mundrey, J. S. and Consultant (1993) *Th* McGrow KM Componm Railway Track Engineering Fourth Edition*.

Novales, M., Orro, A. and Bugarín, M. R. (1887) 'Track Geometry for Light Rail Systems'. doi: 10.3141/2146-03.

Sadeghi, J. and Askarinejad, H. (2010) 'Development of improved railway track degradation models', *Structure and Infrastructure Engineering*, 6(6), pp. 675–688. doi: 10.1080/15732470801902436.

Sadeghi, J. and Askarinejad, H. (2012) 'Application of neural networks in evaluation of railway track quality condition †', 26(1), pp. 113–122. doi: 10.1007/s12206-011-1016-5.

Sato, Y. (1995) 'Japanese Studies on Deterioration of Ballasted Track', *Vehicle System Dynamics*, 24(sup1), pp. 197–208. doi: 10.1080/00423119508969625.

Shafahi, Y., Masoudi, P. and Hakhamaneshi, R. (no date) 'Track Degradation Prediction Models, Using Markov Chain, Artificial Neural and Neuro-Fuzzy Network', *Railway-Research.Org*. Available at: http://www.railway-research.org/IMG/pdf/i.1.1.1.3.pdf.

Shahin, M. A., Jaksa, M. B. and Maier, H. R. (2009) 'Recent Advances and Future Challenges for Artificial Neural Systems in Geotechnical Engineering Applications', *Advances in Artificial*

Neural Systems, 2009, pp. 1–9. doi: 10.1155/2009/308239.

Soleimanmeigouni, I. *et al.* (2018) 'Modelling the evolution of ballasted railway track geometry by a two-level piecewise model', *Structure and Infrastructure Engineering*. Taylor & Francis, 14(1), pp. 33–45. doi: 10.1080/15732479.2017.1326946.

Vale, C. and M. Lurdes, S. (2013) 'Stochastic model for the geometrical rail track degradation process in the Portuguese railway Northern Line', *Reliability Engineering and System Safety*, 116, pp. 91–98. doi: 10.1016/j.ress.2013.02.010.

Wu, W., Dandy, G. C. and Maier, H. R. (2014) 'Protocol for developing ANN models and its application to the assessment of the quality of the ANN model development process in drinking water quality modelling', *Environmental Modelling and Software*. Elsevier Ltd, 54, pp. 108–127. doi: 10.1016/j.envsoft.2013.12.016.

Yousefikia, M. *et al.* (2014) 'Modeling Degradation of Tracks for Maintenance Planning on a Tram Line', *Journal of Traffic and Logistics Engineering*, 2(2), pp. 86–91. doi: 10.12720/jtle.2.2.86-91.

Zakeri, J. A. and Shahriari, S. (2012) 'Developing A Deterioration Probabilistic Model for Rail Wear', *International Journal of Traffic and*, 1(2), pp. 13–18. doi: 10.5923/j.ijtte.20120102.02.

Zio, E., Marella, M. and Ã, L. P. (2007) 'Importance measures-based prioritization for improving the performance of multi-state systems : application to the railway industry', 92, pp. 1303–1314. doi: 10.1016/j.ress.2006.07.010.

APPENDICES

APPENDIX A: Regression model plots

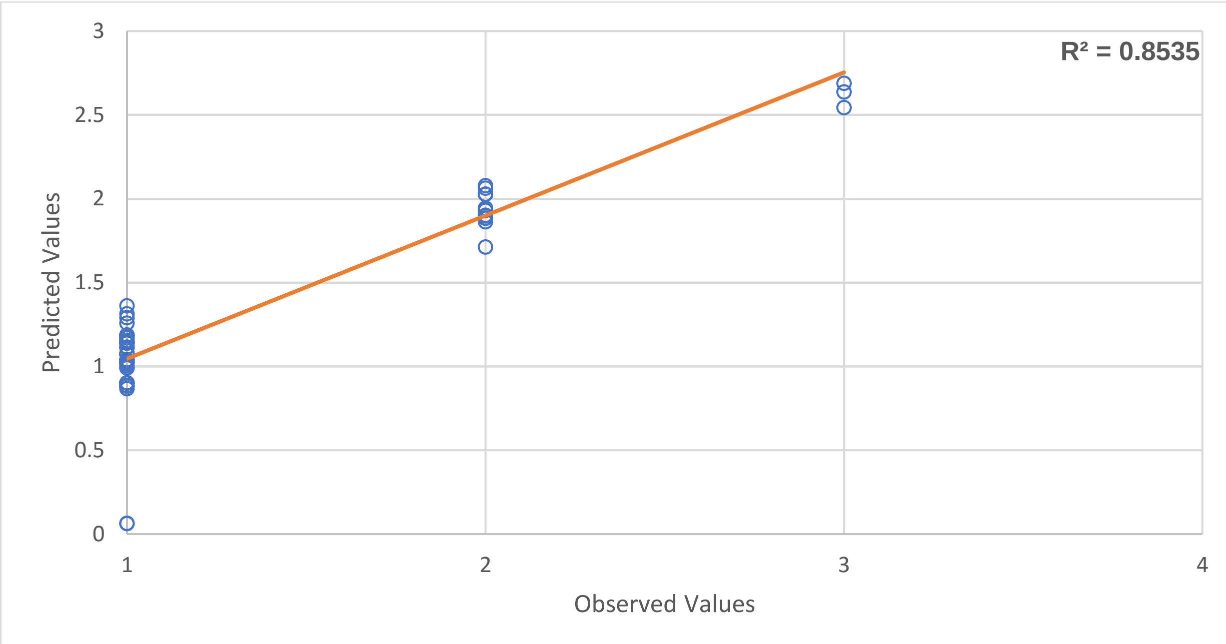

Figure A.1: Observed values against Predicted values for straight sections of uplink line

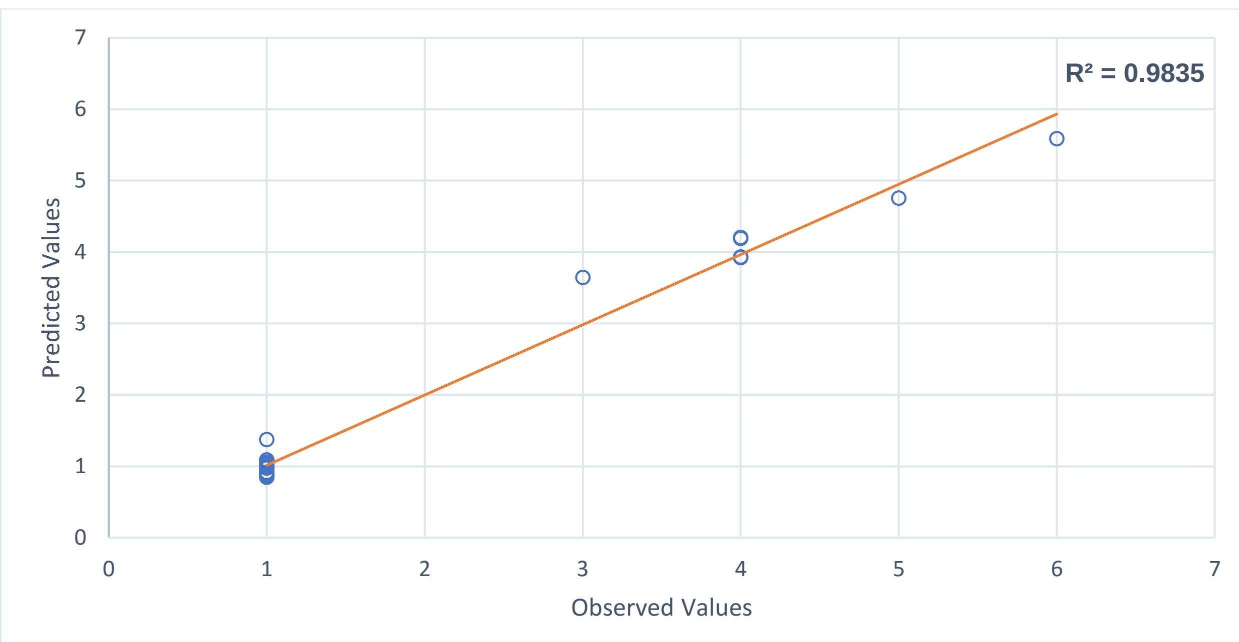

Figure A.2: Observed values against Predicted values for straight sections of downlink line

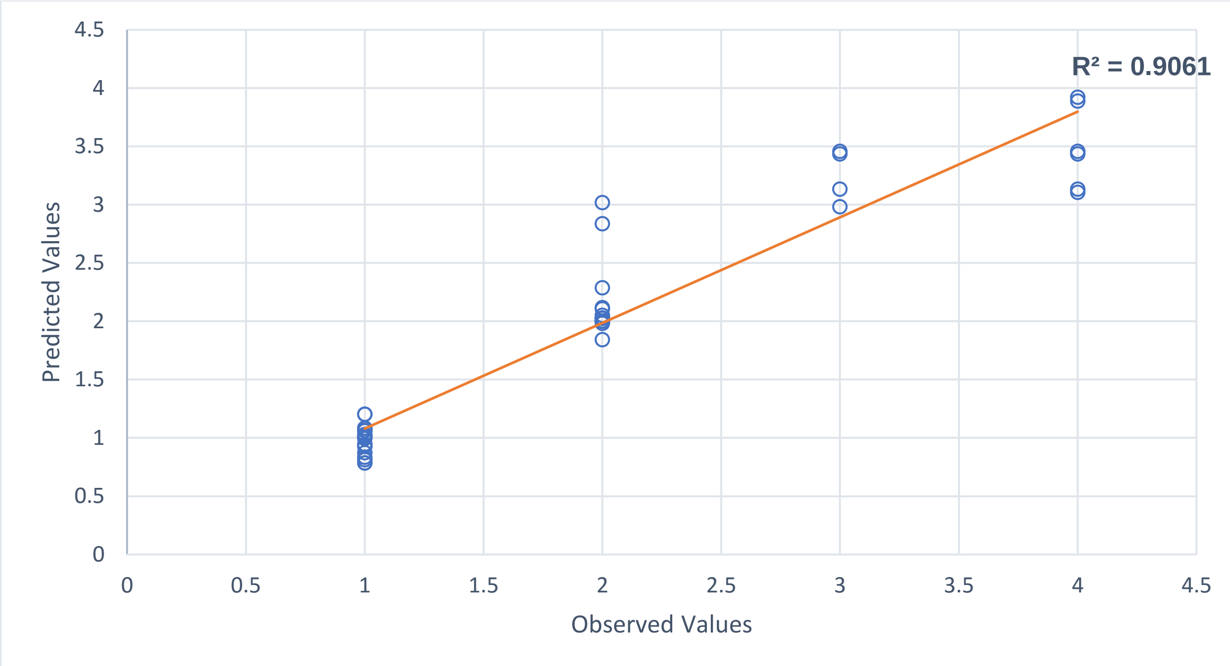

Figure A.3: Observed values against Predicted values for Curved sections of Uplink line

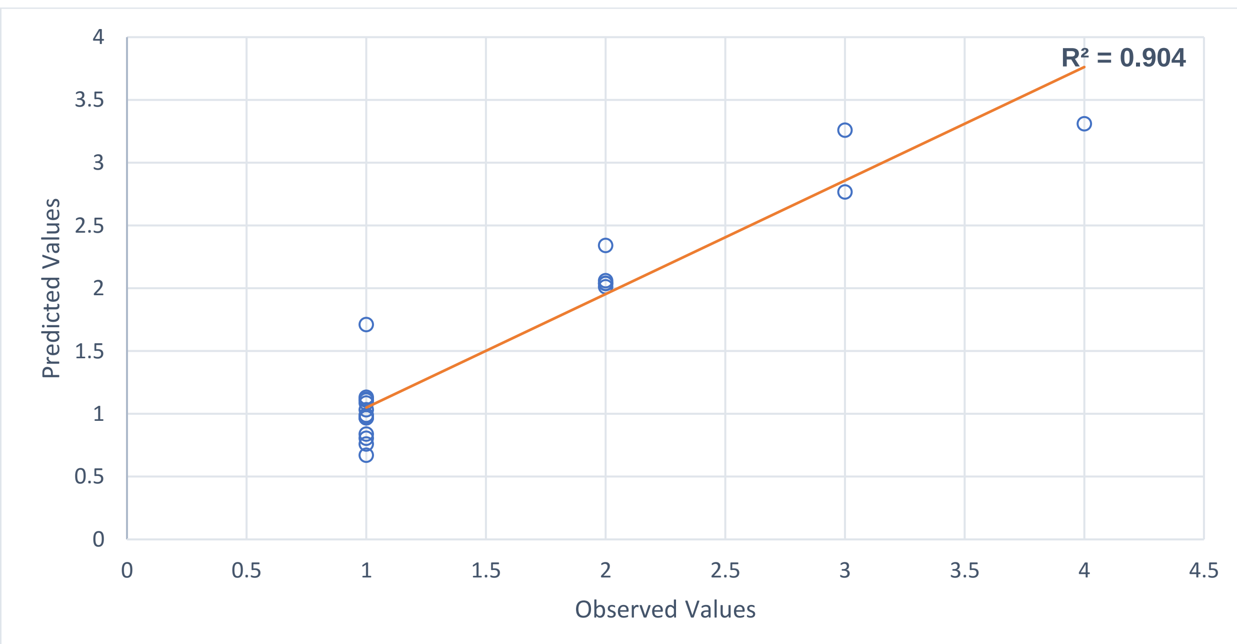

Figure A.4: Observed values against Predicted values for Curved sections of Downlink line

APPENDIX B: ANN Model plots

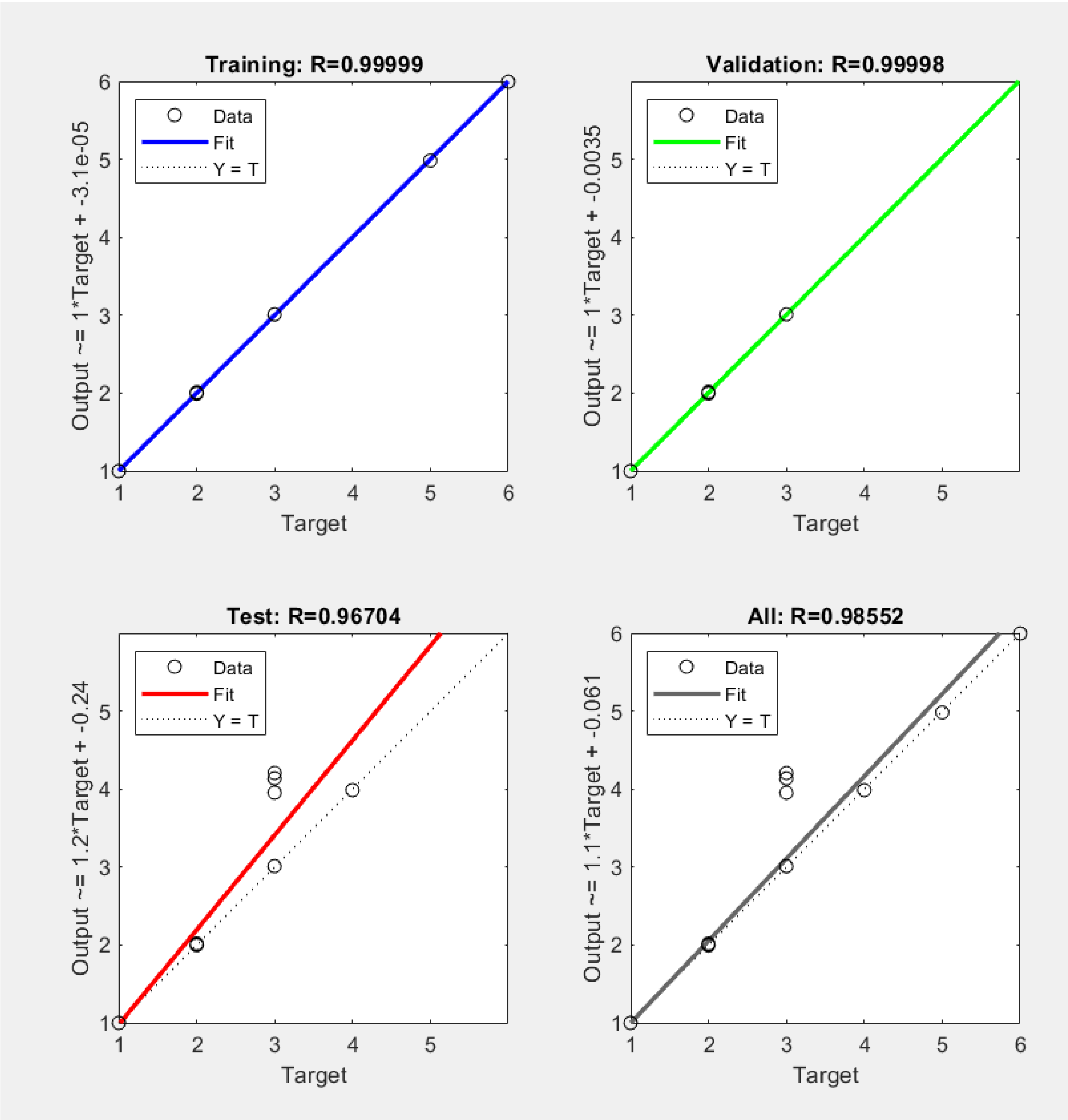

Figure B.1: Coefficients of correlation for straight sections of Uplink line

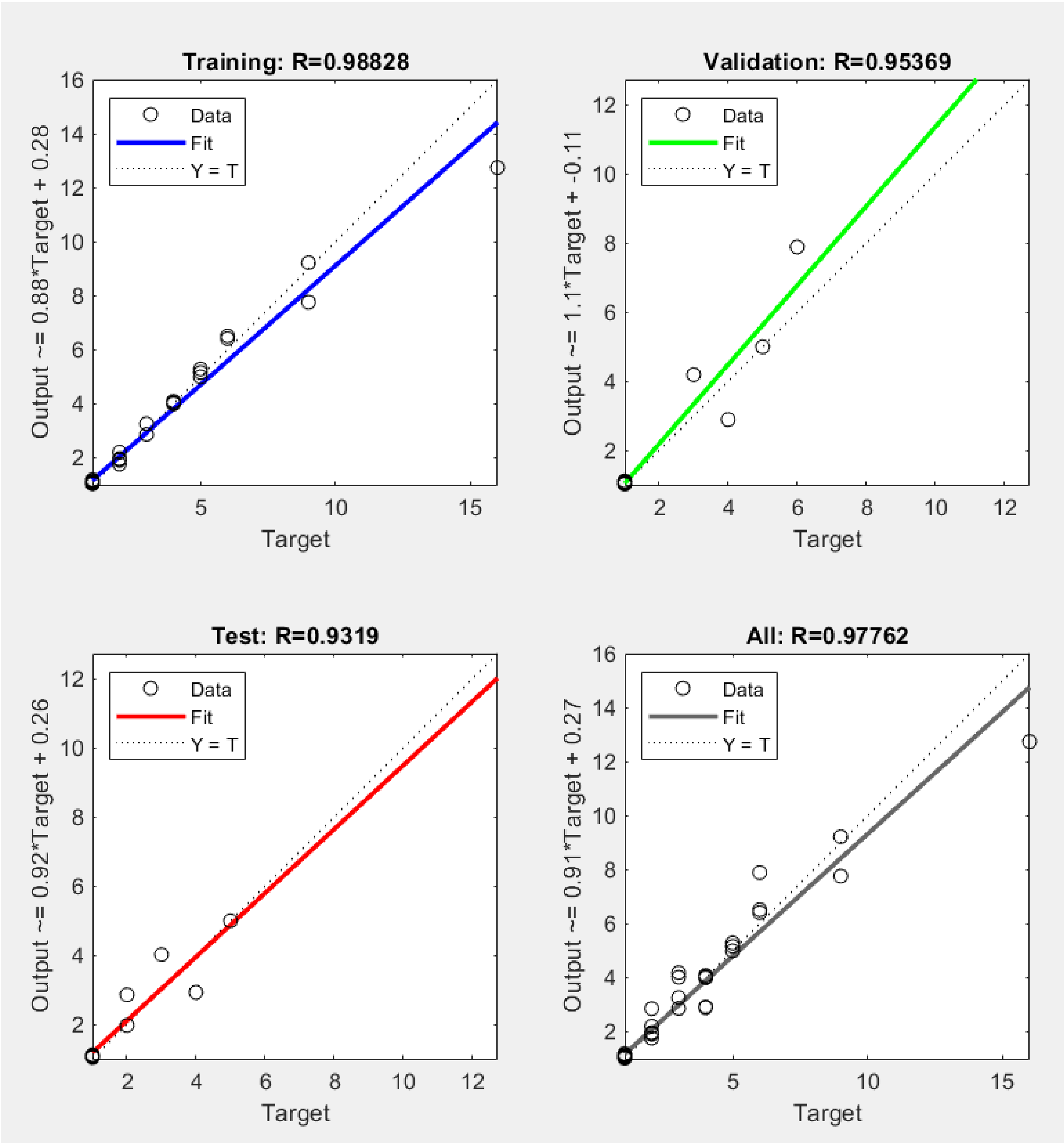

Figure B.2: Coefficients of correlation for straight sections of Downlink line

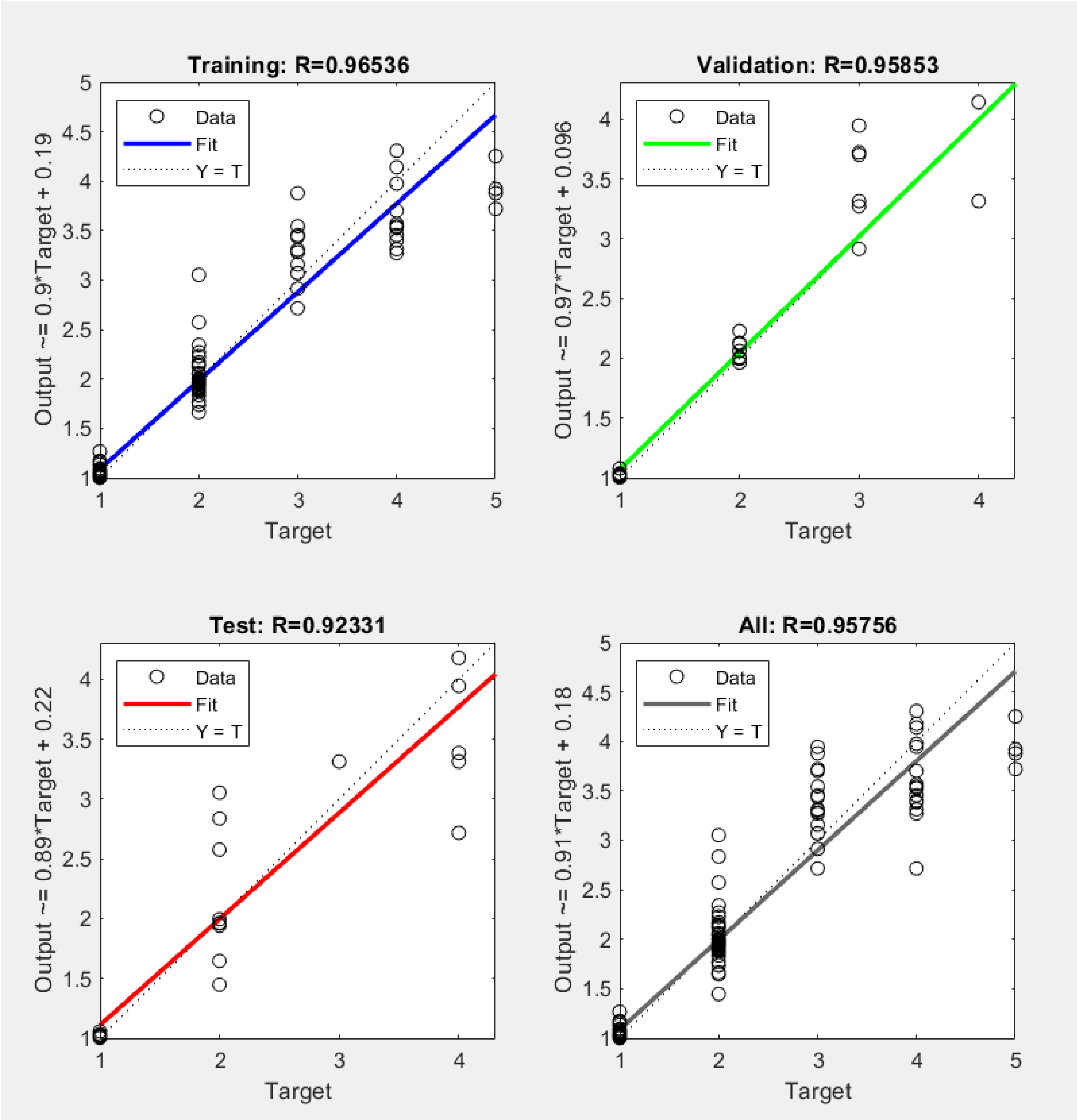

Figure B.3: Coefficients of correlation for straight sections of North South line of AALRT

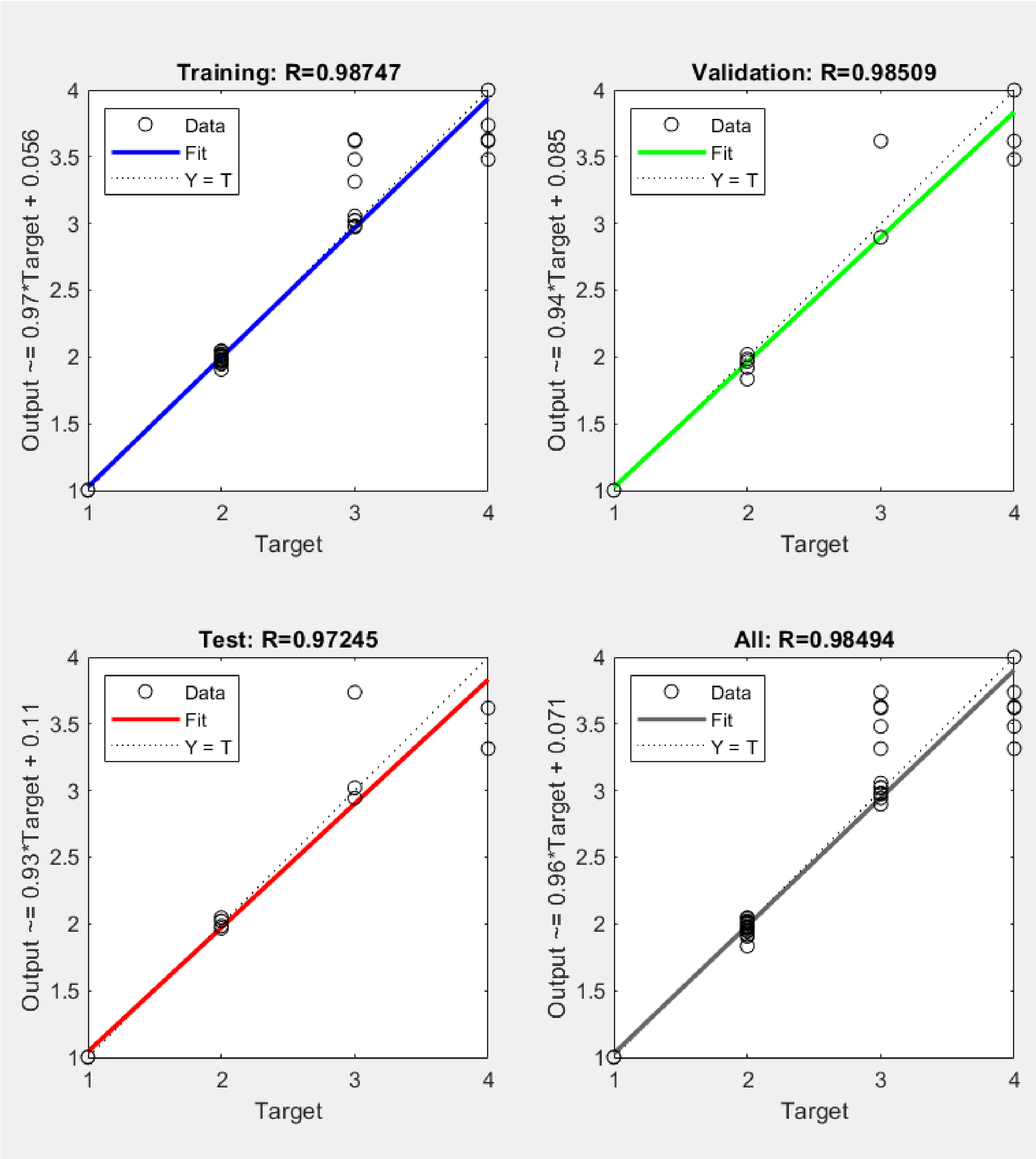

Figure B.4: Coefficients of correlation for Curved sections of Uplink line

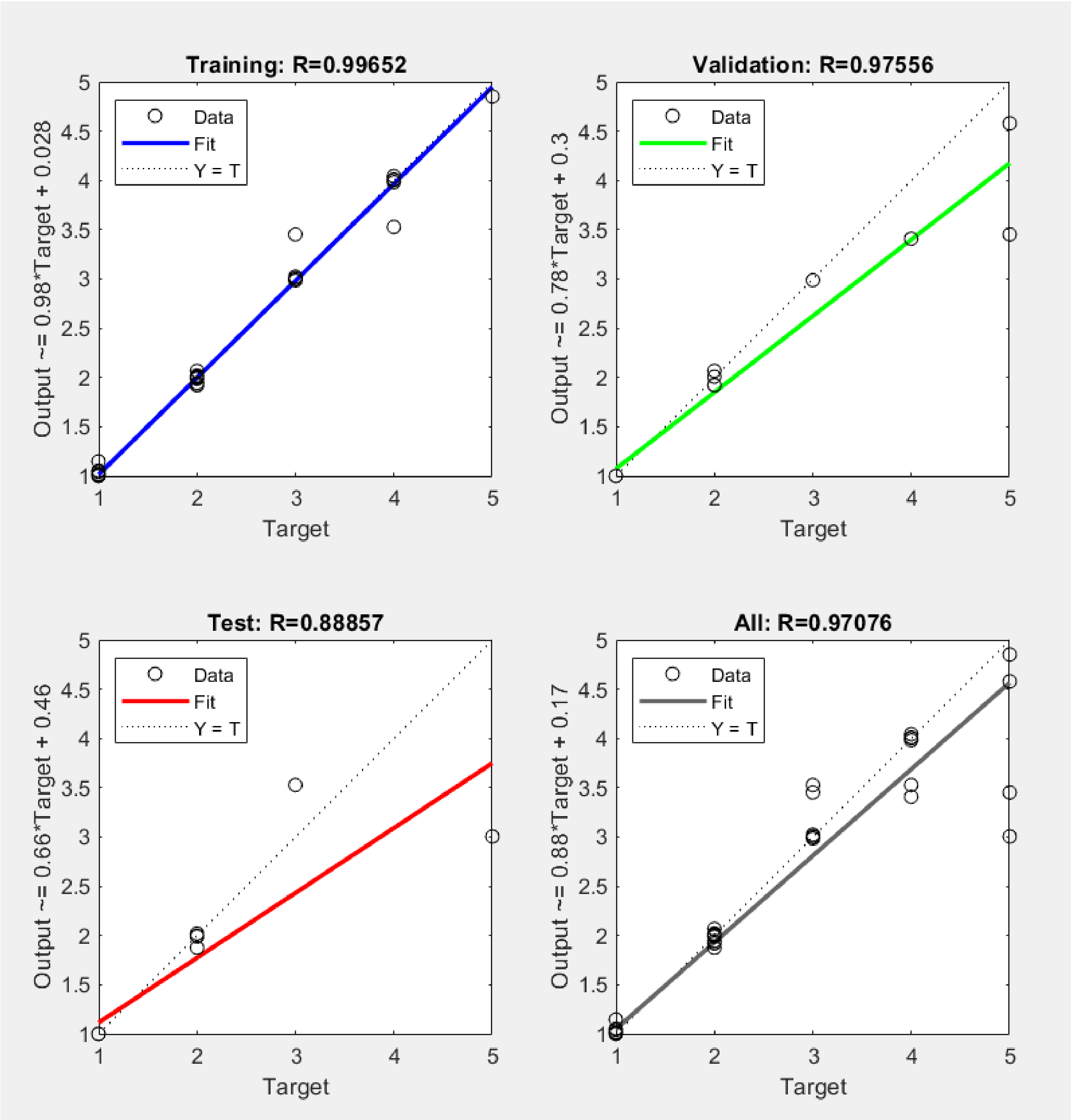

Figure B.5: Coefficients of correlation for Curved sections of Downlink line

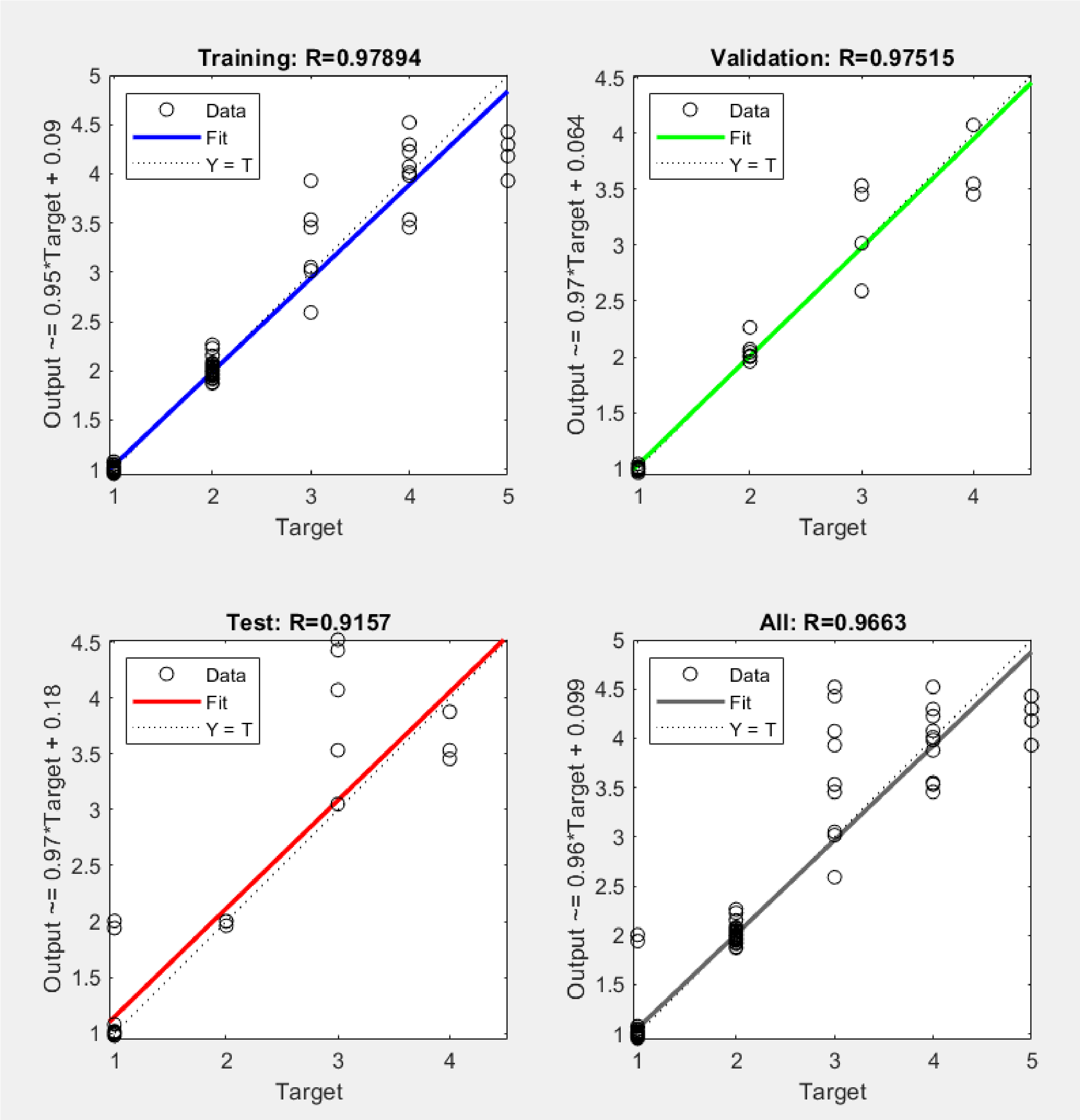

Figure B.6: Coefficients of correlation for Curved sections of North South line of AALRT

APPENDIX C: Summary of Sample Data

- Summary of Sample data for Straight section in 2019

Date	Location	Average Longitudinal leveling	Cumulative Longitudinal leveling	Number of trips	Tonnage	Cumulative N.T	Cumulative Tonnage	Average Speed
6/17/2019	Uplink	2.14	2.14	132	2878.52	132	2878.52	60.00
6/18/2019	Uplink	1.00	3.14	109	2394.38	241	5272.9	38.08
6/19/2019	Uplink	2.36	5.51	149	3006.92	390	8279.82	65.00
6/20/2019	Uplink	1.00	6.51	112	2582.78	502	10862.6	35.00
6/21/2019	Uplink	2.00	8.51	147	3355.88	649	14218.48	55.00
6/22/2019	Uplink	1.00	9.51	95	3355.88	744	17574.36	34.44
6/23/2019	Uplink	1.00	10.51	95	2376.86	839	19951.22	35.00
8/1/2019	Uplink	2.70	2.70	157	3693.14	157	3693.14	60.00
8/2/2019	Uplink	1.00	3.70	125	2609.54	282	6302.68	35.49
8/3/2019	Uplink	2.00	5.70	185	3731.66	467	10034.34	50.00
8/4/2019	Uplink	1.00	6.70	115	3630.8	582	13665.14	20.00
8/5/2019	Uplink	2.00	8.70	185	3853.16	767	17518.3	50.00
8/6/2019	Uplink	1.00	9.70	110	2892.5	877	20410.8	35.00
8/7/2019	Uplink	1.00	10.70	132	2892.5	1009	23303.3	35.00
8/8/2019	Uplink	2.40	13.10	185	3853.16	1194	3853.16	55.00
8/9/2019	Uplink	1.00	14.10	157	2609.54	1351	6462.7	37.19
9/4/2019	Uplink	1.00	1.00	114	2021.24	114	2021.24	45.83
9/5/2019	Uplink	1.74	2.74	166	3210.38	280	5231.62	51.67
9/6/2019	Uplink	2.25	4.99	185	3355.88	465	8587.5	57.50
9/8/2019	Uplink	1.00	5.99	114	2021.24	579	10608.74	40.00
9/9/2019	Uplink	1.00	6.99	120	2609.54	699	13218.28	43.75
10/10/2019	Uplink	2.40	2.40	185	3410.96	185	3410.96	65.00
10/11/2019	Uplink	2.00	4.40	185	3128.48	370	6539.44	60.00
10/12/2019	Uplink	1.00	5.40	123	2892.5	493	9431.94	35.00
11/11/2019	Uplink	1.00	1.00	122	2892.5	122	2892.5	20.00
11/12/2019	Uplink	4.00	5.00	185	3368.48	307	6260.98	70.00
11/13/2019	Uplink	2.00	7.00	185	3368.48	492	9629.46	55.00
6/21/2019	Downlink	1.00	1.00	112	2793.5	112	2793.5	36.11
6/22/2019	Downlink	1.75	2.75	157	2878.52	269	5672.02	53.46
6/23/2019	Downlink	1.00	3.75	117	2541.68	386	8213.7	36.04
6/24/2019	Downlink	2.30	6.05	155	3016.82	541	11230.52	58.33
6/25/2019	Downlink	1.00	7.05	101	2984.42	642	14214.94	35.48
6/26/2019	Downlink	1.03	8.08	101	2984.42	743	17199.36	35.78
8/1/2019	Downlink	1.80	1.80	93	3630.8	93	3630.8	49.00

8/2/2019	Downlink	6.00	7.80	185	3853.16	278	7483.96	75.00
8/3/2019	Downlink	4.00	11.80	165	3128.48	443	10612.44	65.00
8/7/2019	Downlink	2.33	14.13	165	2625.54	608	13237.98	45.00
9/5/2019	Downlink	1.00	1.00	166	2625.54	166	2625.54	20.00
9/6/2019	Downlink	3.00	4.00	185	3095.58	351	5721.12	61.67
9/7/2019	Downlink	9.00	13.00	185	3226.38	536	8947.5	75.00
9/9/2019	Downlink	6.00	19.00	185	3226.38	721	12173.88	70.00
9/10/2019	Downlink	2.00	21.00	178	3095.58	899	15269.46	55.00
11/13/2019	Downlink	1.00	1.00	123	3036.68	123	3036.68	26.67
11/14/2019	Downlink	1.00	2.00	118	3018.5	241	6055.18	45.00
11/15/2019	Downlink	2.56	4.56	185	3291.38	426	9346.56	65.00
11/18/2019	Downlink	9.67	14.22	185	3291.38	611	12637.94	70.00

- Summary of Sample data for Curved section in 2019

Date	Location	Average Longitudinal leveling	Cumulative Longitudinal leveling	Number of trips	Tonnage	Cumulative N.T	Cumulative Tonnage	Average Speed
6/18/2019	uplink	1.81	1.81	132	2676.62	132	2676.62	52.81
6/19/2019	uplink	2.44	4.25	149	2732.96	281	5409.58	56.56
6/20/2019	uplink	1.00	5.25	109	2511.8	390	7921.38	35.11
6/21/2019	uplink	1.25	6.50	132	3630.8	522	11552.18	31.25
8/1/2019	uplink	2.85	2.85	157	3693.14	157	3693.14	58.85
8/2/2019	uplink	1.00	3.85	155	3006.92	312	6700.06	40.23
8/5/2019	uplink	2.70	6.55	166	3853.16	478	10553.22	58.25
8/6/2019	uplink	1.00	7.55	138	3630.8	616	14184.02	32.50
8/7/2019	uplink	2.50	10.05	157	3731.66	773	17915.68	56.25
8/8/2019	uplink	1.00	11.05	125	3006.92	898	20922.6	36.88
9/4/2019	uplink	1.71	1.71	185	3355.88	185	3355.88	52.14
9/6/2019	uplink	1.00	2.71	114	2609.54	299	5965.42	33.00
9/8/2019	uplink	1.00	3.71	114	2021.24	413	7986.66	27.50
10/18/2019	uplink	2.00	5.71	185	3079.58	598	11066.24	55.00
10/19/2019	uplink	2.75	2.75	185	3210.38	185	3210.38	56.43
11/12/2019	uplink	1.00	3.75	88	2892.5	273	6102.88	32.50
11/13/2019	uplink	2.00	2.00	185	3282.02	185	3282.02	55.00
11/14/2019	uplink	1.00	3.00	123	3025.45	308	6307.47	34.29
6/21/2019	downlink	1.00	4.00	123	3036.68	431	9344.15	20.00
6/22/2019	downlink	2.25	2.25	153	3124.64	153	3124.64	51.25
6/24/2019	downlink	1.00	3.25	147	3006.92	300	6131.56	42.50
6/25/2019	downlink	1.00	4.25	147	2984.42	447	9115.98	43.46
6/26/2019	downlink	1.20	5.45	138	3006.92	585	12122.9	61.79
8/1/2019	downlink	1.33	6.78	134	3006.92	719	15129.82	53.33
8/2/2019	downlink	3.00	3.00	185	3630.8	185	3630.8	50.00
8/5/2019	downlink	1.00	4.00	157	2625.54	342	6256.34	50.00
8/6/2019	downlink	1.00	5.00	157	2625.54	499	8881.88	55.00
8/7/2019	downlink	1.00	6.00	157	2625.54	656	11507.42	60.00
9/5/2019	downlink	1.00	7.00	155	2625.54	811	14132.96	42.50
9/6/2019	downlink	2.00	2.00	185	3226.38	185	3226.38	20.00
9/7/2019	downlink	3.38	5.38	185	3226.38	370	6452.76	35.63
9/10/2019	downlink	1.00	6.38	166	2908.2	536	9360.96	56.67
11/13/2019	downlink	2.40	8.78	185	3095.58	721	12456.54	40.00
11/14/2019	downlink	2.33	2.33	185	3172.64	185	3172.64	60.00
11/15/2019	downlink	3.40	5.73	185	3368.48	370	6541.12	63.00
11/16/2019	downlink	2.83	8.57	185	3291.38	555	9832.5	59.17
11/18/2019	downlink	1.00	9.57	123	3036.68	678	12869.18	47.75
11/19/2019	downlink	1.00	10.57	110	3030.68	788	15899.86	47.86