Pavement Life Cycle Assessment:
From Case Study to Machine Learning Modeling

by

Md Rakibul Alam

A thesis

presented to the Memorial University of Newfoundland

in fulfillment of the

thesis requirement for the degree of

Master of Engineering

in

Civil Engineering

Date of Convocation: October, 2020

St. John’s, Newfoundland, Canada, 2020

AUTHOR'S DECLARATION

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners. I understand that my thesis may be made electronically available to the public.

ABSTRACT

Climate change is a global challenge with long-term implications. Human activities are changing the global climate system, and the warming of the climate system is undeniable. According to a roadway construction study, the construction of the surface layer of an asphalt pavement alone generates a carbon footprint of 65.8 kg of CO_2 per km. Therefore, a sensible approach to study environmental impact from road pavement is crucial.

Pavement life cycle assessment (LCA) is a comprehensive method to evaluate the environmental impacts of a pavement section. It features a cradle-to-grave approach assessing critical stages of the pavement's life. Material production, initial construction, maintenance, use and end of life phases exist in an entire pavement life cycle. The thesis consists of three components, which started with finding the environmental impact for different pavement maintenance and rehabilitation (M&R) techniques in the maintenance phase. The second component evaluated the environmental impact due to pavement vehicle interaction (PVI) in the use phase. Finally, the goal of the third component was to develop a set of pavement LCA models.

To evaluate environmental impact for four major M&R techniques: rout and sealing, patching, hot in-place recycling (HIR) and cold in-place recycling (CIR), initially a fractional factorial design approach was applied to determine which factors were significant. Considering those significant factors and other necessary data, a hypothetical LCA case study was performed for the city of St. John's. It was found that the global warming potential (GWP) held the highest values among four M&R techniques. CIR technique produced the lowest percentage of GWP (83.87%), and for asphalt patching, the CO_2 emission resulted in the highest percentage (92.22%) which became the least suitable option.

To understand the PVI effect, the required data and information are collected from the Long-Term Pavement Performance (LTPP) program. Out of 141 Canadian road sections, 22 sections were selected. Several climatic parameters, including annual precipitation, annual temperature, and annual freezing index data, were collected from these 22 sections and further processed for developing clusters using a hierarchical clustering approach. Finally, the Athena Pavement LCA tool was used to measure the environmental impact from the PVI effect for each cluster. It was found that cluster 2 (high annual precipitation, high annual freezing index, and

medium annual temperature) experienced the highest rate of IRI increase and therefore, high GWP value. The LCA result also indicated a relatively higher GWP due to pavement roughness from heavy vehicle traffic compared with light vehicle traffic. For the PVI effect due to pavement deflection, cluster 4 (maximum vehicle load and the minimum subgrade stiffness) emitted the highest GWP among all the clusters.

Pavement LCA tools require an extensive amount of data to estimate the environmental impact. In the first and second studies, all Canadian road pavement sections were not possible to consider because of the large quantity of time consumption for LCA of each section. Therefore, a database management software, Microsoft SQL Server Management Studio, was used for filtering and data manipulation of the LTPP database considering all Canadian road sections. The manipulated data were further used to develop the LCA models using machine learning algorithms: multiple linear regression, polynomial regression, decision tree regression and support vector regression. The models determined the significant contributors and quantified the CO_2 emission in pavement material production, initial construction, maintenance and use phase. Model validation was also performed. The study also revealed the contribution of Canadian provinces' CO_2 emission. The proposed LCA models will help the decision-makers in the pavement management system.

Technical Papers from this Research

Journal Articles:

1. Alam M. R, Hossain K, Butt A A, Caudle T, Bazan C. (2019). Life Cycle Assessment of Asphalt Pavement Maintenance and Rehabilitation Techniques: A Study for the City of St. John's. Canadian Journal of Civil Engineering. http://dx.doi.org/10.1139/cjce-2019-0540

2. Alam M. R, Hossain K, Bazan C. (2020). A Systematic Approach to Estimate GWP from Pavement Vehicle Interaction Using Canadian LTPP Data. Journal of Cleaner Production. https://doi.org/10.1016/j.jclepro.2020.123106

3. Alam M. R, Hossain K, Bazan C. (2020). Life Cycle Analysis for Asphalt Pavement in Canadian Context: Modeling and Application. Submitted to International Journal of Pavement Engineering.

ACKNOWLEDGMENTS

All praise and gratitude go to the Almighty Allah who keeps me in the right path and assists me in completing the thesis work successfully. There are a lot of people whom I have to acknowledge for their care, support and inspiration during my MEng thesis. Sincerely, I apologize if I miss anyone during my acknowledgments.

First and foremost, I am deeply indebted and wish to express my utmost gratitude to my thesis supervisor, Dr. Kamal Hossain. His continuous support, generous help, endless encouragement, constructive comments and invaluable suggestions throughout the progress of the thesis work is priceless. He has addressed my research issues and provided effective direction to resolve those whether it is day or night, weekday or weekend. Also, I would like to express my gratitude to my co-supervisor, Dr. Carlos Bazan, for his enormous support and inspiration during my thesis. Without their assistance, it would have been impossible for me to finish my thesis.

Secondly, I acknowledge the contribution of the Advanced Road & Transportation Engineering Lab (ARTEL) at Memorial University of Newfoundland and the research team members of the lab. I have learned a lot of things from them in developing research ideas, methodology, writing skills and skills for working as a proactive team member.

Lastly, my profound thanks, love and gratitude are due to my parents, sister and dearest friends in the USA and my home country Bangladesh. They have mentally supported me in the last two years until I finish my thesis work. As an international student, their supports have carried high value in my journey of graduate study and research in Canada.

To

My parents, sister,
And all of my close friends,
Without whom my success would not be possible

CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 Background

Addressing the leading factors of climate change is one of the major global issues in today's time. Therefore, substantial research is being conducted to study the environmental impacts of road pavement designs and construction practices using life cycle assessment (LCA) methodologies. LCA proposes a cradle-to-grave approach, assessing critical stages of an asset's life. This methodology can be used to evaluate the environmental impact of an entire road system, considering both project level and network level. As a result, highway and pavement management can make an environmentally friendly decision that has a lower impact on the environment.

Case study or project-based LCA was very common for transportation infrastructure in the early stage of LCA. Mainly case study based pavement LCA was highly focused for research purposes because pavement LCA was appropriate for questions relevant in a procurement situation (Azhar, Toller, & Birgisson, 2015). Initially in the 1900s, case study based pavement LCA aimed to find out which pavement type is better between asphalt and concrete pavement (Häkkinen & Mäkelä, 1996; Horvath & Hendrickson, 1998; Meil, 2006; Mroueh et al., 2000; Stripple, 2001; Yu & Lu, 2012). After that, at the beginning of the twenty-first century, studies focused specifically on asphalt or concrete pavements (Butt et al., 2014; Evangelista & De Brito, 2007; Loijos, 2011; Vidal et al., 2013).

In pavement LCA, a lot of information is required of each stage of the pavement life cycle: material production, construction, maintenance, use and end of life. As a model represents the collection of data and summarizes the data for present needs and predictions, an LCA model can analytically result the environmental assessment. Based on LCA framework, this model can estimate the emission from pavement projects within a shorter time than the conventional calculative method. From 2004 until now, several research have been performed to develop the framework, model and computer tools (Horvath, 2004; Huang, Bird, & Heidrich, 2009; João Santos, Ferreira, & Flintsch, 2015; Zhang, Keoleian, & Lepech, 2008).

Canada is the second-largest country in the world and it has huge provinces and territories. Therefore, within a province, there are geometric regions that have completely different climates. Most of the North American LCA model based studies were performed in the US context. Athena Institute developed LCA tool for the Canadian context (Ahammed et al., 2016). However, only one project can be environmentally assessed at a time and requires a huge amount of processing time. Therefore, a modeling approach, specifically machine learning based modeling approach, can deal with a large amount of data and predict the emission report within a few seconds.

1.2 Pavement LCA

As indicated above, pavement LCA is considered as the most comprehensive way to evaluate the environmental impact in the pavement for a given analysis period. This analysis period is not limited to a fixed year. It can be the lifetime of a pavement section. Material production, construction, maintenance and rehabilitation, use and end of life – all phases can be considered in LCA. Pavement LCA methodology is used to compare the impact of different pavement designs, mix designs, alternative maintenance & rehabilitation schedules, etc.

International Organization for Standardization (ISO) 14040 standard (International Standard, 1997) is the general foundation of LCA for each type of asset. Pavement LCA framework (Harvey et al., 2010) from the University of California Pavement Research Center is highly regarded. Their guidelines are organized and updated frequently. The framework and necessary data required are described elaborately in chapter 2.

1.3 Research Problem

As indicated in the first section, most of the case studies of LCA were performed to compare asphalt pavement and concrete pavement. LCA case studies were also performed for comparing the performance of the addition of different types of material with asphalt binder. However, research related to environmental emission from pavement maintenance and rehabilitation (M&R) using LCA methodology is still rare.

Among pavement life cycle phases, the use phase accounts for significant environmental impacts through various factors including rolling resistance effect in increased fuel demand (i.e., pavement vehicle interaction), albedo (solar radiation reflectibility of pavement surface), carbonation (absorption of carbon dioxide in pavement structure), nighttime illumination of the roads (energy demand) and generation of leachates from pavement materials (emission of chemical substances) (Ziyadi, Ozer, & Al-Qadi, 2017). The energy consumption during the pavement use phase is about 700 times higher than that of the construction phase (Paulo, Araújo, Oliveira, & Silva, 2014). A recent study of LCA in asphalt rubber pavement also revealed the dominant energy consumption in the usage stage (Cao, Leng, Yu, & Hsu, 2019). This high energy consumption in the use phase suppressed the energy-saving advantage of warm mix asphalt (WMA) additives during the entire life cycle of the pavement. Massachusetts Institute of Technology (MIT) research team found that most of the pavement's use phase environmental impact results from pavement vehicle interaction (PVI). 58% of greenhouse gas (GHG) was reported from an urban interstate pavement in Missouri where 93% (out of 58% GHG) was reported from PVI (Gregory, 2017). The MIT research team studied four states to understand the environmental impact over a 50 years analysis period of the pavement life cycle. They considered four states representing four different climates: Missouri (wet freeze), Arizona (dry no freeze), Colorado (dry freeze) and Florida (wet no freeze) (Mack et al., 2018). As Canada is a big country with diverse climates, there was a need to study PVI effect for the Canadian climate condition.

As stated earlier, each phase of the pavement life cycle requires a large number of data inventory to estimate environmental impact. The material production and initial construction phases need material properties, percentage of material quantities and pavement layer thicknesses. The maintenance phase requires the pavement distress survey data, maintenance history and environmental condition. The use phase requires information on traffic conditions, pavement roughness and texture properties. As a result, a particular pavement project requires quite an ample time for LCA. The situation becomes worse and time-consuming when it is time to decide the best alternative from different combinations of pavement designs, construction and maintenance options.

1.4 Research Objectives

The thesis consists of three components. In the beginning, an LCA of pavement M&R was performed. The second component was about LCA for the PVI effect. The third and final component was about the LCA model.

The goal of the first component is to determine the environmental impacts of major M&R techniques for asphalt pavements using LCA. The M&R techniques include asphalt patching, rout and sealing, hot in-place recycling (HIR), and cold in-place recycling (CIR). For the LCA analysis, various project parameters were selected using a statistical sensitivity analysis technique. The parameters include but are not limited to the specifications of road pavement section and the evaluation of emission for each M&R in a 30-year life cycle. For the emission analysis, an Athena Institute's LCA tool called the Athena Pavement LCA was used (Alam et al., 2019).

The overall goal of the second component of the thesis is to obtain a better understanding of the PVI impact on environmental effect in asphalt pavement using the LCA framework in the Canadian climate conditions. The required data and information are collected from the Long-Term Pavement Performance (LTPP) program to conduct this research. To understand the PVI effects for different road sections, the global warming potential (GWP) values are computed and compared.

The goal of the third and last component of the thesis is to develop a set of LCA models for each pavement life cycle phase in the Canadian context using database management tools and particularly machine learning algorithms. Machine learning algorithms were used to develop the model after filtering the LTPP data using Microsoft SQL Server Management Studio. This study also shows the emission across Canadian provinces as well as comparative analysis.

1.5 Organization of the Thesis

This thesis consists of seven chapters. The first chapter introduces the motivation of this study and gives a brief overview of Pavement LCA, the research problem, and the specific objectives of this research. The second chapter provides an extensive literature review on pavement LCA framework, case study based LCA researches, PVI effect based researches and LCA model based researches. In addition, this chapter provides an overview of the tools and approaches that I implemented during research. The third chapter describes the methodology of three core studies of this thesis. In general, each component of study has two primary sections: data preparation and analysis approach. The fourth chapter provides the results from LCA for M&R. The fifth chapter presents the results from the second component, LCA for the PVI effect. The fifth chapter presents the result from the third and final component, the LCA models. Significance factors of model, parameter selection, model accuracy and comparative emission analysis for different provinces are described in this chapter. The seventh and final chapter provides the conclusion of this MEng thesis and outlines the future areas of research.

CHAPTER 2

LITERATURE REVIEW

2.1 Background of Pavement LCA Framework

LCA proposes a cradle-to-grave approach, assessing critical stages of an asset's life. LCA requires inventory data and provides an impact assessment system that reflects on the environmental footprint for each critical stage of the asset. For LCA of any asset or product, three major phases are goal and scope definition, inventory analysis, and impact analysis. Pavement LCA also follows the same phases in its analysis period.

2.1.1 Goal and Scope

According to the International Organization for Standardization (ISO) 14040 standard, an LCA study's goal and scope must be defined at the beginning of LCA. Defining the goal of LCA includes identifying its purpose. In general, three purposes can be listed as follows.

- *Project level*: To take a decision for a particular project
- *Network level*: To take a decision for an entire highway network combining several continuous projects
- *Combination of project and network level:* To take a decision for any zone through a set of discrete projects which are sufficient to identify that zone

The goal is vital because the variables to be used in the assessment are usually dependent on what the intended goal is.

The scope helps to establish the system boundaries and the limits of the LCA. The scope of an LCA study also clarifies whether this will quantify the environmental impacts of one system or will compare alternative systems. In the former situation, all the components of all life cycle phases need to be considered, whereas the latter situation allows the reasonable elimination of some components. The reason for the elimination is to avoid any unnecessary complexity.

Under the goal and scope definition phase, a functional unit can be described. The ISO defines the functional unit as "quantified performance of a product system for use as a reference unit" (International Organization for Standardization, 2006). The functional unit for pavement should represent physical dimensions and pavement performance. Performance requirements can include analysis period, traffic type, asphalt mixture composition, etc. The functional units considered in an LCA study can be categorized and presented into physical, structural and annualized functional units (Ziyadi et al., 2017).

- *Physical functional unit*: Physical dimensions of pavements refer to length, width, and the number of lanes for a road. This physical dimension is contingent upon the road classification system. However, when LCA considers special features of a road (i.e., parking lots, intersections etc.), different types of related measurements may be more appropriate.
- *Structural functional unit:* The structural properties of pavement construction material (specific gravity, percentage of material) and traffic loading (both for heavy and light vehicle) can be considered structural functional unit. This functional unit is attributed to structural performance variables.
- *Annualized functional unit:* The analysis period of LCA is considered annualized functional unit. The analysis period is denoted by the time horizon from the inputs to the outputs. As different infrastructure has different functional ages and requires maintenance after different time periods, therefore analysis period is fixed at the beginning of LCA.

2.1.2 Life Cycle Inventory

Life cycle inventory data can be organized in the life cycle phases concept. The life cycle phases of the pavement include pavement design, material production, construction, use, maintenance and rehabilitation, and end-of-life.

- *Pavement design:* Structural design of each pavement in the analysis, including surface, base, subbase, subgrade, shoulder, and drainage.
- *Material production:* Raw material production, mixing of hot mix asphalt (HMA) or portland cement concrete (PCC) in plants, feedstock energy of materials that are used as a fuel, transport of materials from plant to site and vice versa.
- *Initial construction:* Transport of materials and equipment to site, using equipment at the site and construction according to the design

- *Maintenance:* Pavement distress study, selection of maintenance alternatives, transport of materials and equipment to site, using equipment at the site, implementation of the selected alternatives, fuel consumption due to traffic congestion during maintenance
- *Use:* Fuel consumption, rolling resistance effect for excess fuel demand (i.e., pavement vehicle interaction), albedo (solar radiation reflectibility of pavement surface), carbonation (absorption of carbon dioxide in pavement structure), nighttime illumination of the roads (energy demand) and generation of leachates from pavement materials (emission of chemical substances) (Ziyadi et al., 2017).
- *End-of-Life Phase:* Material landfilling, reusing, recycling before landfilling in material production and construction phase (Recycled Asphalt Pavement) along with maintenance phase (Hot In-Place Recycling, Cold In-Place Recycling).

2.1.3 Environmental Impact Assessment

Impact assessment entails determining the environmental relevance of all the inputs and outputs of each phase in the pavement life. This includes the meaningful environmental impacts associated with the production, maintenance, use, and end of life (EOL) phase of the assets. The summary result includes specific environmental impact categories, including acidification potential (kg SO_2 eq.), global warming potential (kg CO_2 eq.), human health respiratory effect potential (kg $PM_{2.5}$ eq.), ozone depletion potential (kg CFC-11 eq.), smog potential (kg O_3 eq.), and eutrophication potential (kg N eq.) (Alam et al., 2019).

For example, to quantify global warming potential (GWP), the emissions in CO_2 can be measured based on the equivalence from the International Panel on Climate Change's 100-year time horizon variables (Meil, 2006) as shown in the following equation.

$$GWP(kg) = CO_2\ (kg) + [CH_4\ (kg) * 23] + [N_2O\ (kg) * 296] \quad \text{... (2.1)}$$

2.2 Review of Case Study Based Research

This section reviews the research articles related to pavement LCA from conference proceedings, journals and technical reports. The review intends to find out the current practices of pavement LCA and identify the gaps from where we can develop. As there is a lot of case study-based pavement LCA research available. The following Table 2.1 summarizes the cited asphalt pavement related literatures which overcame new challenges in their case study. It was found that the goals and scopes are different, and there are different system boundaries in the research. However, the review is summarized in chronological order.

Table 2-1: A brief review of cited literature on pavement LCAs

Serial No.	Author	Year	Pavement Life Cycle Phases Considered	Impacts Considered	Objectives of The Study
1	Häkkinen and Mäkelä	1996	Materials, construction, maintenance, use	Energy, air emissions, raw materials, noise	To compare environmental impacts between concrete and asphalt pavements
2	Horvath and Hendrickson	1998	Materials, construction and EOL	Energy, air emissions, raw materials, water releases, hazardous waste, water use	To compare environmental impacts from asphalt and Steel-Reinforced Concrete Pavements
3	Mroueh et al.	2000	Materials, construction, maintenance	Energy, air emissions, raw materials, leaching water use, noise	To examine the use of industrial by-products in asphalt and concrete roads
4	Stripple	2001	Materials, construction, maintenance, use	Energy, air emissions, raw materials	To examine JPCP* and asphalt pavement using hot and cold production technique

Serial No.	Author	Year	Pavement Life Cycle Phases Considered	Impacts Considered	Objectives of The Study
5	Park et al.	2003	Materials, construction, maintenance	Energy, air emissions	To produce estimates for the materials extraction and production phase
6	Zapata et al.	2005	Materials, construction	Energy	To analyze the energy consumption of a CRCP** and an asphalt pavement
7	Athena Institute	2006	Materials, maintenance	Energy and air emissions	To compares the energy and global warming potential of asphalt and JPCP*
8	Huang et al.	2009	Material, construction, maintenance	Energy and air emission	To cite five reasons that current pavement (until 2009) LCA tools are inadequate
9	Yu and Lu	2012	Materials, construction, maintenance, use and end of life	Energy and air emissions	To compare three overlay systems
10	Vidal et al.	2013	Materials, construction, maintenance, recycling, use and end of life	Energy and air emissions	To compare the impact of zeolite-based WMA pavements and HMA

Serial No.	Author	Year	Pavement Life Cycle Phases Considered	Impacts Considered	Objectives of The Study
11	Butt et al.	2014	Materials, construction, maintenance, use and EOL	Energy and air emissions	To calculate and compare energy consumption of binder and additives
12	Schlegel et al.	2016	Materials, construction, maintenance, recycling, use and end of life	energy consumption	To compare LCA for the use of HMA without hydrated lime and with hydrated lime
13	Farina et al.	2017	Production, construction and maintenance	Energy and global warming potential	To focus on LCA road paving technologies using asphalt mixtures containing recycled materials and reclaimed asphalt pavement
14	Santos et al.	2017	Construction and maintenance	Energy, climate change, acidification, eutrophication, and photochemical ozone creation	To provide the comparison of using American and European LCA tools

Serial No.	Author	Year	Pavement Life Cycle Phases Considered	Impacts Considered	Objectives of The Study
15	Valle et al.	2017	Material production, construction, maintenance, use	GWP	To calculate the life cycle GWP due to climate change
16	Chen and Wang	2018	Material production, construction, maintenance, use and EOL	GHG emission	To quantify GHG emission of asphalt pavements containing RAP
17	Santos et al.	2018	Construction, maintenance and rehabilitation	Air, water and soil emission	To understand the environmental impact of reducing mixing temperature by WMA
18	Wang et al.	2018	Material production, construction, maintenance, use and EOL	Urban flooding, water recycling and water purification.	To quantify the environmental impacts of pervious pavements
19	Samieadel et al.	2018	Material production, construction, maintenance, use and EOL	Global warming potential index and energy consumption	To compare LCA of bio-modified binder and conventional asphalt binder

Serial No.	Author	Year	Pavement Life Cycle Phases Considered	Impacts Considered	Objectives of The Study
20	Cao et al.	2019	Material production, construction, use	energy consumption	To identify the long-term energy-saving role of WMA technologies in AR pavement

*JPCP = Jointed plain concrete pavement

**CRCP = Continuously reinforced concrete pavement

The summary reveals that almost all of the LCA study is for comparison purpose. Initially, it started with regular asphalt and concrete pavement to compare their environmental impact (Häkkinen & Mäkelä, 1996; Horvath & Hendrickson, 1998). There are some research that considered different types of concrete pavement (JPCP, CRCP) for comparative analysis (Meil, 2006; Stripple, 2001; Zapata & Gambatese, 2005). In 2009, Huang et al. studied the gap and inadequacy where pavement LCA can be developed for better life cycle assessment. Though few studies were using industrial by-product and cold production techniques in 2000 and 2001, so many researchers studied environmental impact characterizing pavement material mixture design and construction process (Mroueh et al., 2000; Stripple, 2001). Three overlay system (HMA overlay, PCC overlay and crack, seal & overlay) had been studied to find out their environmental impact (Yu & Lu, 2012). Vidal et al. 2013 and Santos et al. 2018 studied the application of WMA in the pavement life cycle (Santos et al., 2018). Recently, there have been several studies on sustainable asphalt mixture using additives, hydrated lime and biomodified binder (Butt et al., 2014; Samieadel, Schimmel, & Fini, 2018; Schlegel et al., 2016). Besides material characterization, the old recycled pavement application effect recently has also become an important topic of research using the pavement LCA approach (Farina, Zanetti, Santagata, & Blengini, 2017). The environmental impact of a new type of pavement for better drainage, pervious pavement is explained by Wang et al. 2018.

2.3 Review of PVI Effect Based Research

The pavement vehicle interaction (PVI) effect includes three factors: pavement roughness, deflection, and texture depth. Velinsky and White performed the first investigation of fuel consumption due to pavement roughness in 1979. They developed a roughness model to predict fuel consumption based on field data (Velinsky & White, 1979). In their study, they found that vehicle rolling resistance increased with the increase in pavement roughness because of the energy dissipation in tire and vehicle suspension system. After Velinsky and White 1979, a great deal of research was conducted considering different types of vehicle classes (Bester, 1984; Delanne, 1994; Plessis, Visser, & Curtayne, 1990), speed categories (Cenek, 1994; Sandberg, 1990) and experimental methodologies (i.e., laboratory prototype testing, test track, etc.) (Amos, 2006; A. M. A. Soliman, 2006).

In 1984, Bester (Bester, 1984) stated that pavement roughness has a strong correlation with PVI effect and consequent fuel consumption. In 1990, Sandberg (Sandberg, 1990) found that pavement roughness could affect vehicle fuel consumption as much as 12% for the surface condition tested. Cenek (Cenek, 1994) observed in 1994 that an increase in roughness level from 1.4 to 2.3m/km could lead to a rise in PVI by 55%. The remaining researchers also found a linear relationship between energy consumption and pavement roughness.

Besides, some research were done to develop a model of excess fuel consumption because of pavement roughness. In 2006, Soliman, 2006 simulated vehicle motion on two roadway sections with a quarter car model. However, the most extensive study was performed in 2012 by Chatti and Zaabar (Zaabar & Chatti, 2010). They conducted a field investigation on five different roadways sections with five vehicle classes: passenger cars, sport utility vehicles, passenger vans, light trucks and articulated heavy trucks. Tests were performed in both winter and summer environmental conditions. Three vehicle speeds were involved: 35mph, 45mph, and 55mph. It was concluded that the increase in pavement roughness results in an increase in energy consumption using mechanistic model based Highway Development and Management software (HDM-4). For heavy vehicles, the consequences of fuel consumption are prevalent compared with light vehicles.

Deflection occurs after the immediate imposition of the dynamic load from vehicles on the asphalt surface. The tire sinks into the pavement surface, which is visually undeterminable. With the increase in depth of the sink, an uphill slope shows up in front of a tire. Therefore, additional energy is required to compensate for this slope.

Compared with research on roughness-based PVI effect, deflection-related PVI research was less extensive due to less sophisticated technical supports. Zaniewski et al. 1985 conducted quantitative research on this topic. They concluded that 20% more fuel consumption was observed in asphalt pavement compared with concrete pavement (Zaniewski & Butler, 1985). The higher stiffness of concrete pavement compared with asphalt pavement was reported as the reason. They used the data that was developed by the Federal Highway Administration and Texas Research and Development Foundation from 1979-1982 for updating vehicle operating cost.

Recent research was done based on empirical observations and which relate pavement deflection to various factors such as pavement temperature, vehicle classes, speed variation and road grade (Benbow, Iaquinta, Lodge, & Wright, 2007; Hultqvist et al., 2002; Sumitsawan, Ardenkani, & Romanoschi, 2009; Taylor, Farrel, & Woodside, 2002). In 2010, Lenngrenn and Feldner (Lenngren & Faldner, 2011), using a falling weight deflectometer, stated that the energy losses for asphalt pavement deflection are almost four times greater than for concrete pavement. Even light and articulated trucks at low speed can cause a 5% increase in fuel consumption over asphalt pavement in summer conditions (Chatti & Zaabar, 2012).

More recently, Akbarian et al. 2012 conducted a state-of-the-art study of PVI based on a mechanistic approach. They developed a mechanistic based PVI model for pavement deflection (Akbarian et al., 2012). In their study, the viscoelastic pavement was considered on an elastic foundation. A new temperature-dependent factor, relaxation time has been introduced. Currently, several researchers are addressing the PVI effect on pavement LCA. In the United States, there have been studies of life cycle assessment considering the PVI effect in Virginia (Akbarian, Louhghalam, & Ulm, 2014), Florida, Arizona, Colorado and Missouri (Gregory, 2017).

2.4 Review of LCA Model Based Research

This section aims to provide an in-depth literature review on pavement LCA modeling research. As one of the research objectives of this thesis is to develop the model of pavement LCA for Canadian road sections using machine learning approaches, most of the literature reviews below are related to pavement LCA modeling and machine learning algorithm.

Several studies develop LCA models and tools for the life cycle of all or particular phases of the pavement life cycle. Dr. Horvath of the University of California Berkeley developed an LCA framework and computer-based tool for the Recycled Materials Resource Center, USA (Horvath, 2004). Pavement Life-cycle Assessment Tool for Environmental and Economic Effects (PaLATE) is a Microsoft Excel-based tool that estimates the environmental and economic impacts for a single project at a time. The use phase of the pavement life cycle was not considered in this tool. Zhang et al. proposed another model that combined four different external models: material environmental impact model, vehicle emissions model, construction equipment model, and a traffic flow model (Zhang et al., 2008).

Huang et al. described the development of an LCA model for pavement construction and maintenance that accommodates recycling and up-to-date research findings (Huang et al., 2009). Microsoft Excel was selected for the calculation and visualization of emission results in their model.

The studies mentioned above established a framework for LCA of a particular project, although some uncertainty issues became necessary to resolve. Baker and Lepech mentioned several significant uncertainties: database uncertainty, model uncertainty, measurement error and uncertainty in preferences (Baker & Lepech, 2009). To address such uncertainties, Kim et al. developed an artificial neural network model for a Korean project only for the material production and construction phases (Kim, Lee, Park, & Kim, 2013). They considered real-life asphalt pavement projects and input variables after interviews with domain experts. In their study, they examined only the greenhouse gas (GHG) emission from their project. Model accuracy was between -30% and 50%. Noshadravan et al. addressed the uncertainty due to the pavement roughness prediction in the pavement's entire life (Noshadravan, Wildnauer, Gregory, & Kirchain, 2013). This pavement roughness prediction value was retrieved from the pavement ME design tool. Global warming potential (GWP) value was measured using a Monte Carlo simulation to address uncertainty propagation in GWP.

Santos et al. stated that LCA tools using a spreadsheet approach has several limitations including issues with managing and storing a large amount of data, dealing with a variation of data that change over a project analysis period and addressing intrinsic sophistication for vehicle fuel consumption modeling using spreadsheet macros. Therefore, they proposed a pavement LCA model written in visual basic and SQL programming language (João Santos et al., 2015).

Yu et al. found one limitation of the 2015 LCA model research by Santos et al. (Yu, Qiang, & Gu, 2016). Little attention was paid to the reliability of the data source in pavement LCA. As the energy intensity coefficient (EIC) (MJ/kg) of each material has a wide range in different literature, this can cause significant uncertainty in the impact assessment. For example, it was found in previous literature that asphalt's energy intensity coefficient was 0.7-6.0 MJ/kg. Therefore, Yu et al. proposed the Pedigree Matrix method. In this method, each dataset was evaluated based on the data quality indicator (DQI) and converted to the probability density function of the modified beta distribution. Three weighting methods were employed to estimate weights for different datasets. Then, through a Monte Carlo simulation approach, the ultimate probability density function of EIC was determined. This approach helped to quantify the uncertainties in LCA results.

To quantify the uncertainty of input variability, Ziyadi and Al-qadi proposed a simple method in each source to account for model variable and model-form uncertainties (Ziyadi & Al-qadi, 2019). Interval analysis was used to input variability uncertainty analysis. This interval analysis was performed through the Monte Carlo simulation. A Bayesian surrogate model was used to estimate model variable uncertainty. For the evaluation of model-form uncertainty, the orthogonal polynomial basis function concept was implemented. Using the three quantification approaches, uncertainties were analyzed for energy and GWP assessment for the Chicago metropolitan area.

2.5 Tools and Approach

Several tools and mathematical approaches have been implemented to execute the life cycle assessment. Athena Pavement LCA tool was used to estimate GWP emission for M&R and PVI effect studies (Alam et al., 2020; Alam et al., 2019) . In the LCA modeling study, for managing large LTPP database SQL was used. Python language was used for applying different machine learning algorithms on SQL led LTPP data. The fractional factorial design approach was used for sensitivity test in M&R studies to find out the significant factors, which were then used in LCA. The hierarchical clustering method was introduced in the PVI effect study for the clustering of Canadian LTPP sections with different climatic conditions.

2.5.1 Athena Pavement LCA

The Athena Pavement LCA, developed by the Athena Sustainable Materials Institute in Canada, was designed for Canadian conditions and selected US regions, relating to roadway life cycles (Ahammed et al., 2016). The program includes an adequate material database and allows the user to select the design specification of pavement surface, granular sub-base and base materials and shoulder materials. Besides the material properties, the Athena Pavement LCA program comes equipped with a vast library of selectable machinery and practices. It also features quantifiable data, such as project pavement segment length, to adopt a functional unit on which the practices of the case study are compared. Based on this information, Athena Pavement LCA can analyze all the stages in a pavement's life cycle except for the EOL stage.

After assessment of the specific M&R techniques, the project reports result in an array of impact categories including emission factors to air, water, and land. For the absolute value, there are options to select from all the listed measures which include energy consumption, air emissions, water emissions, land emissions, and resource use. The summary result includes more specific environmental impact categories including fossil fuel consumption, acidification potential, global warming potential, human health respiratory effect potential, ozone depletion potential, smog potential, and eutrophication potential. Similarly, pavement vehicle interaction (PVI) effects can be estimated by providing pavement roughness and deflection modulus values between major roadway rehabilitation. Moreover, for the maintenance phase, the sub-columns are labeled to separate material and equipment from transportation. There is also a total value table and the specified units for each impact category.

2.5.2 SQL and Python

SQL stands for Structured Query Language. SQL is used to communicate with a database file. In the LCA modeling study, a large database of all Canadian road sections was filtered, manipulated and prepared for the modeling purpose. Data modeling was performed using python language. For this purpose, a free integrated development environment (IDE), Spyder, was used. Spyder IDE includes editing, interactive testing, debugging and introspection features. Under this IDE, pandas open-source data analysis and manipulation tools were used. During the preparation of the data for the machine learning algorithms, it was important to develop multidimensional arrays. The array development is performed using NumPy library. Scikit-learn machine learning library was used for regression algorithms including multiple linear regression, polynomial regression, support vector regression (SVR) and decision tree regression.

2.5.3 Fractional Factorial Design

A factorial design is one of the experimental design approaches which aims to find out how multiple factors or independent variable affect a dependent variable. A factorial design with two factors that each has two levels is called a 2×2 factorial design**.** There are two types of factorial design: full factorial design and fractional factorial design.

Full factorial design leads to experiments where all possible combinations of factors and levels are considered. Factors are the independent variables. In most of the experiences, we deal with factors. Factors affect the outcome or dependable variable of the experiment. In a factorial design, each treatment factor (the factor that is of interest in a study/experiment) in an experiment will have two levels. Two levels are the minimum and maximum values of the corresponding treatment factor. All factor interactions are considered in full factorial design which makes itself exhaustive, time-consuming and quite expensive approach.

Therefore, the fractional factorial design approach emerges to resolve the drawback of full factorial design. This design method considers only a subset of the possible permutations of factors and levels. The standard notation for fractional factorial designs is l^{k-p}, where, l is the number of levels in a treatment factor, k is the number of treatment factors, p is the number of confounding interactions. Confounding occurs when nobody is sure which factors – or combinations of factors – are affecting the output. A blocking approach can help to minimize confounding. Design-Expert software can be used for the fractional factorial design of a large dataset.

2.5.4 Hierarchical Clustering

Hierarchical clustering is an algorithm that groups similar objects into groups or clusters. In hierarchical clustering, a dendrogram is a tree diagram to illustrate the arrangement of clusters. It shows relationships among similar data. To interpret a dendrogram, it is necessary to examine the heights on the Y-axis at which any two data are joined together and will indicate whether they have similar descriptive characteristics. The height on the Y-axis is based on the Euclidean distance matrix which is estimated from the complete linkage method (when a cluster is formed, its distance to other objects is computed as the maximum Euclidean distance between any object in the cluster and the other object). In the dendrogram, the height denotes the value of this Euclidean distance metric between clusters. As a result, if two clusters merge at a height x, it means that the distance between those clusters is x.

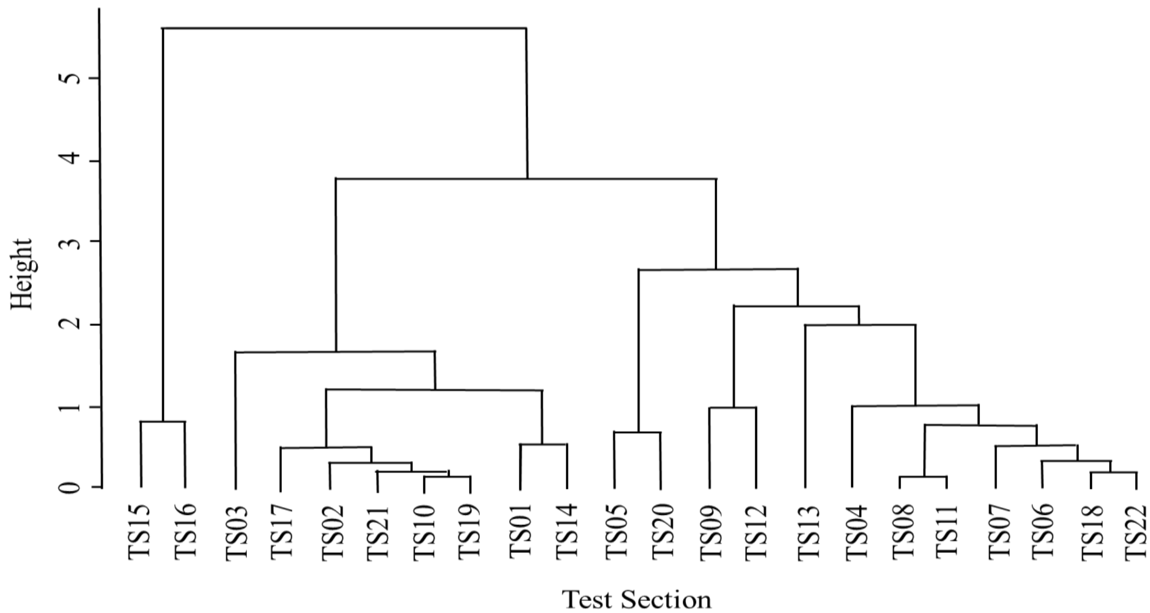


Figure 2-1: Dendrogram

The silhouette width value is a measure of how similar an object is to its own cluster (intra-cluster) compared to other clusters (inter-cluster). The silhouette width value ranges from −1 to +1, where a high value indicates that the objects have similarity to its own cluster and less similarity with neighboring clusters. Silhouette plot is drawn using trial and error method for different numbers of clusters, i.e., 1, 2, 3, 4, 5, and more.

2.5.5 Machine Learning Algorithm

Machine learning algorithms are programs combining math and logic that adjust themselves to learn the dataset and predict values for similar datasets. A machine learning algorithm performs better when it is exposed to a large dataset in the same way human learns over time.

2.5.5.1 Conventional Regression

A multiple linear regression model can have the following mathematical form.

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \varepsilon \qquad \ldots (2.2)$$

Here, y is the response or dependent variable with k number of regressor variables. The parameters β_0 , β_1 , β_2, ……, β_k are called regression coefficients. The least-squares method can be used to estimate the regression coefficients. Models that include interaction effects are called polynomial regression model. As linear and polynomial regressions are old and conventional regression techniques, we skipped the details.

2.5.5.2 Support Vector Machine Regression

In linear regression, it is a goal to minimize the error rate, whereas, in support vector machine regression (SVR), it is the goal to fit the error within a certain threshold. In SVR, the important task is to keep all the points within the boundary line (dashed lines as shown in Figure 4) and the best fit line indicated by the hyperplane (solid one) that has a maximum number of points. These two dashed lines are ε distance away from the reference data. This distance value is chosen by the user.

When there is a nonlinear relationship between the predictor variable and response, we consider to enlarge the feature (variable) space using function of the predictor variable in order to address non-linearity. This enlargement of feature space results in quadratic and cubic terms. The function which is used to do this is called kernels.

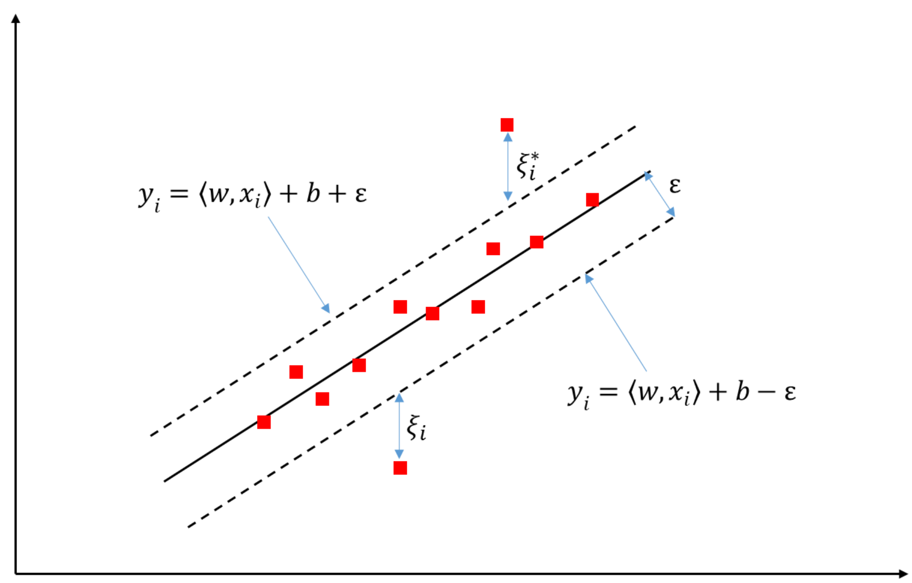


Figure 2-2: Basic concept of support vector regression in 2D

The optimization problem for SVR can be shown in the following mathematical form (Kleynhans, Montanaro, Gerace, & Kanan, 2017).

$$\text{Minimize } \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}(\xi_i^* + \xi_i)$$

$$\text{Subject to } \begin{cases} y_i - \langle w, x_i \rangle - b \leq \varepsilon + \xi_i^* \\ \langle w, x_i \rangle + b - y_i \leq \varepsilon + \xi_i \end{cases}$$

Where w is the learned weight vector, $x_i$is the i^{th} observed variable value, y_i is observed response value and ξ_i is the distance between the boundary line and values outside the boundary line. C is another constraint value that controls the penalty imposed on the observations outside the boundary line. This penalty helps to prevent overfitting.

2.5.5.3 Decision Tree Regression

Decision tree regression segments the predictor into a number of simple regions. In each region, there are observation values. The mean of these regions is decided as the predicted value for that individual region. The systematic method to build a decision tree follows two common steps.

- The set of response variable values for corresponding independent variable values $X_1, X_2, \ldots\ldots, X_j$ are divided into J distinct and non-overlapping regions $R_1, R_2, \ldots\ldots, R_j$
- Though the region R_j could have a set of observations, there is only one prediction value which is the mean of the response values for the observations in R_j

In order to divide the predictor space into high dimensional rectangles (regions), recursive binary splitting approach can be used. The splitting begins at the top of the decision tree at which point all observations can be found in a single region and then successively segments the regions. According to recursive binary splitting, the predictor X_j and the cut point s is selected such that splitting the predictor space into the regions $\{X|X_j < s\}$ and$\{X|X_j \geq s\}$ and leads to the maximum possible reduction in the residual sum of squares (RSS). $\{X|X_j < s\}$ means the region of predictor space in which X_j takes a value less than s. For any j and s, the pair of half-planes and then for lowest RSS, the description can be written in a mathematical form as follows.

$$R_1(j,s) = \{X|X_j < s\}$$
$$R_2(j,s) = \{X|X_j \geq s\}$$
$$min \sum_{i:x_i \in R_1(j,s)} (y_i - \hat{y}_{R_1})^2 + \sum_{i:x_i \in R_2(j,s)} (y_i - \hat{y}_{R_2})^2$$

Here, $\hat{y}_{R_1}$ and $\hat{y}_{R_2}$ are the mean responses for the observations in $R_1(j,s)$ and $R_2(j,s)$, respectively. We split one of the two previously identified regions. We now have three regions. Again, we look to split one of these three regions further, so as to minimize the RSS. The segmentation continues until a stopping criterion is reached. Typically the stopping criterion is to continue segmentation until no region contains more than five observations.

CHAPTER 3

METHODOLOGY

The methodology section includes the steps which were performed to complete the study. Each component of the study is described into two subsections: data preparation and analysis approach.

3.1 Methodology I

Methodology I describes the methodology of environmental emission study for different maintenance and rehabilitation techniques in the city of St. John's context.

3.1.1 Data Preparation

The necessary data is prepared for the LCA study of M&R. This section describes which data has been used and how necessary data for LCA is filtered for the study.

3.1.1.1 Data

To evaluate various M&R techniques in terms of environmental impacts, a statistical technique was considered to determine which factors are significant for environmental impact analysis. A sensitivity test was performed for CIR through fractional factorial design in Design Expert 11. Table 3-2 presents the factors that were considered for sensitivity test whereas Table 3-1 describes the asphalt mixture information. Global warming potential (GWP) was chosen as a response variable. After the sensitivity test, substantial experimental factors were attributed in the Athena Pavement LCA program along with the necessary specification of materials for the analysis. The objective of this hypothetical study is to compare different asphalt pavement M&R techniques using LCA. Therefore, this comparative study does not cover a road network, rather it is focused on an asphalt pavement project only. As a result, road traffic was not considered in the sensitivity analysis.

Table 3-1: Asphalt mixture information

Mix ID	Binder Type	Asphalt Content (% by weight)	VMA (%)	Density (ton/m^3)	Maximum Aggregate Size
NL-1	PG 64-22	6	14	2.46	19mm
HL-3	PG 64-28	4	16	2.24	16mm

3.1.1.2 Sensitivity Test

For the sensitivity analysis, two types of pavement surfaces were considered; one of which was titled HMA PG 64-22 indicated as " NL-1"—a common mix design used in Newfoundland and Labrador and the other was "HL-3"—used in Ontario. Asphalt mix properties are summarized in Table 3-1. NL-1 was considered according to the mixture design specifications of Newfoundland Department of Transportation (Department of Transportation and Works, 2011). Similarly, HL-3 was considered according to the Ontario provincial standard specification (OPSS) 310 (Ministry of Transportation Ontario, 2017). The density of NL-1 is 2.46 ton/m^3 which is greater than that of HL-3 (2.24 ton/m^3) according to Athena Pavement LCA database. Two levels (generally low and high levels which are represented by -1 and +1 respectively), namely Granular A and reclaimed asphalt pavement (RAP) mixtures were considered as base/subbase material.

Table 3-2: Factors for sensitivity test for statistical analysis

Factor		Coding	
Considered LCA Design variable	Letter designation	-1	1
Avg. distance plant to site (km)	A	1km	4km
No. of pavement lift	B	1	3
Pavement surface asphalt mix type	C	NL-1	HL-3
Base/Subbase material type	D	Granular A	RAP
% of affected road	E	5%	20%
Shoulder	F	Unpaved	Paved

Average distance of plant to site was assigned with two levels: 1 km and 4 km. If the factor "distance of plant to site" becomes significant for the small values of distance, certainly for the large value, this factor would be significant. Therefore, the average distance of plant to site was considered small (1 and 4 km) for the sensitivity test. Two levels of pavement lift were used: 1 and 3. For this test, the shoulder of road pavement also had two levels: paved and unpaved. Finally, the percentage of affected road was considered to be 5% and 20%. Global warming potential

(GWP) was chosen as a response variable. All of the considered six factors with two levels represented by -1 and +1 for each factor are shown in Table 3-2. After the execution of projects for 32 combinations of assigned factors through the LCA software, a half fractional factorial design was implemented in Design Expert 11 using GWP values from LCA report results. Based on the half normal probability plot (Figure 3-1) and the *p* values of the ANOVA results (Table 3-3), it was concluded that the LCA system was highly sensitive to the change of average transport distance between the plant to site, the percentage of affected road, and the number of pavement lifts.

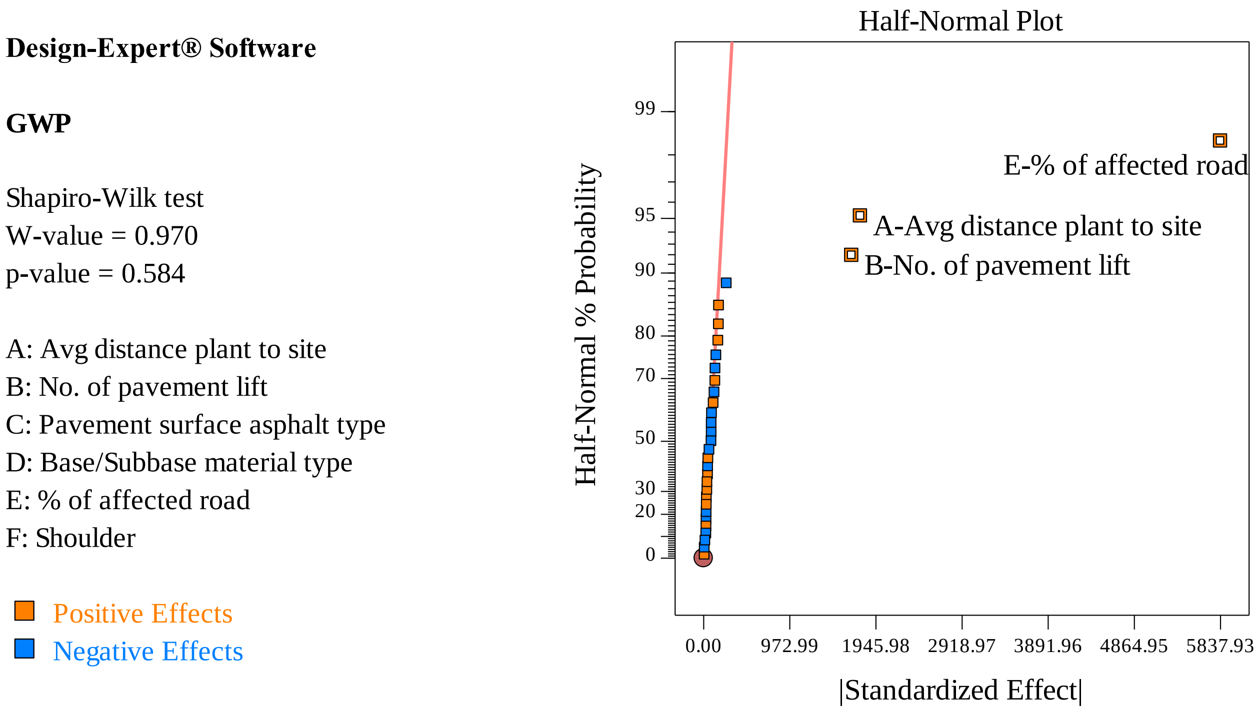


Figure 3-1: Half-Normal percentage probability plot

Table 3-3: ANOVA summary

Source	Sum of Squares	Degree of freedom	Mean Square	F-value	p-value
Model	3.20E+08	3	1.07E+08	1266.31	< 0.0001
A-Avg. distance plant to site	2.51E+07	1	2.51E+07	297.71	< 0.0001
B-No. of pavement lift	2.24E+07	1	2.24E+07	265.75	< 0.0001
E-% of affected road	2.73E+08	1	2.73E+08	3235.49	< 0.0001
Residual	2.36E+06	28	84269.16		
R^2	0.9927				
Adjusted R^2	0.9919				
Predicted R^2	0.9904				
Adequacy of Precision	90.4373				

3.1.2 Analysis Approach

3.1.2.1 Assumptions and Functional Units

To quantify the environmental impacts of various M&R techniques in Athena Pavement LCA, a number of input parameters are required including project size (in terms of road length) and project life. A functional unit of a 1 km two-lane asphalt roadway pavement was considered for a 30-year project life span in St. John's, NL. The pavement included one pavement lift with two granular layers (base and subbase), and an unpaved shoulder on both side of the roadway. For transporting materials, the average distance of plant to site, site to stockpile and equipment depot to site was considered to be 30 km in another Canadian study [Manitoba case study (Ahammed et al. 2016)]. A study in the Netherlands considered distance from plant to site within a range of 44 to 120 km. 30 km distance was considered reasonable to assume, hence was decided to be used for this study as well. According to updated provincial design and construction standard (Highway Design Division, Department of Transportation and Works, Government of Newfoundland and Labrador) issued in April 2017, HMA PG 64-22 (referred as NL1) was considered in the LCA design section as pavement surface material where base material was granular A and subbase material was granular C. Granular A was used for the construction of the unpaved shoulder.

Table 3-4: Road section design dimensions

Element name	Material	Width (m)	Thickness (mm)
Lane 1 Lift 1	NL-1	3.5	60
Lane 2 Lift 1	NL-1	3.5	60
Left Unpaved Shoulder	Granular A	0.5	40
Right Unpaved Shoulder	Granular A	0.5	40
Granular Layer 1 (Base)	Granular A	8	100
Granular Layer 2 (Subbase)	Granular C	8	100

In this component of the study, pavement distress was considered 20% of total surface area. In order to compare the LCA for M&R techniques, expected life of each M&R technique was kept same (5 years). As a result, during 30 years of study period, maintenance and rehabilitation was performed five times.

3.1.2.2 Life Cycle Phase and System Boundaries

As the objective of this LCA study for M&R is attributed to the comparative analysis of different M&R techniques, use phase and end of life phase were exempted. Material production, initial construction and maintenance phase were considered during LCA data inventory. For the analysis purpose, only the emission report for M&R were the points of interest.

3.2 Methodology II

Methodology II describes the methodology of environmental emission study for the PVI effect in the Canadian context.

3.2.1 Data Preparation

The necessary data collection and preparation technique for the study is explained in this section. LTPP road sections of Canada is grouped first using a clustering approach. When the clustering was prepared, LCA was performed in each cluster to understand the PVI effect in LCA.

3.2.1.1 LTPP Data

The Federal Highway Administration's (FHWA) Long-Term Pavement Performance (LTPP) program collects and stores pavement performance data from in-service test sections across the United States and Canada (FHWA, 2018). Out of 2,581 LTPP road sections named as strategic highway research program (SHRP) ID, 141 sections exist in different provinces in Canada. Since it was excessively time-consuming to consider all 141 test sections, in this present component of the study, 22 test sections were selected, as shown in Table . It was also very challenging for us to find a road section with all the necessary information (values of international roughness index or IRI, traffic data, etc.).

Table 3-5: Canadian LTPP road sections

Province	SHRP ID	Road Number/Name	Selected Test Section	Test Section ID
Newfoundland	1801, 1803, 1808	Trans-Canada Highway	1801, 1803	TS09, TS12
New Brunswick	6804	Highway 102	6804	TS18
	1684	Trans-Canada Highway	1684	TS08
	3803, 1802	Highway 11	1802	TS11
Ontario	1622, B310, B320, B330, B340, B360-62, B322	Highway 11	B310	TS22
	AA01-03, BA01-03, BA61-62, AA62	Highway 48		
	A310-11, A320, A330, A340, A350, A311, 1620	Highway 400		
	2811, 2812	Highway 402	2812	TS13
	1680, 1806	Highway 404		
	0901-03, 0960-62	Trans-Canada Highway	0960	TS04
Quebec	9018, 3001	Highway 30		
	1021, 1125, 3015-16, A310, A320, A330, A340, A350	Highway 40	1021	TS06

Province	SHRP ID	Road Number/Name	Selected Test Section	Test Section ID
Quebec	0903, A901-03	Highway 170	0903	TS03
	1127	Highway 73	1127	TS07
	2011	Highway 212		
	3002	Highway 440		
Manitoba	0501-09, 6450-51, 6452, 6454, 1801, A310, A331, A320, A330, A340, A350-51	Trans-Canada Highway	1801	TS10
	3802	Highway 75		
	AA01-03, AA61	Highway 16	AA01	TS21
Saskatchewan	0901-03, 0959-62, 6405, B310, B320, B330-31, B340, B350-51	Trans-Canada Highway	0901	TS02
	6410, 6412	Highway 11	6410	TS17
	A310, A320, A330, A340, A350-52, A6400, A6420, A6801	Highway 9	A310	TS19
Alberta	A901-03	Crownest Highway	A901	TS20
	1803, 0501-09	Highway 16	0501	TS01
	1805	Highway 201		
	8529	Trans-Canada Highway		
	2812	Highway 21	2812 (Alberta)	TS14
	1804	19		
British Columbia	6006	99	6006	TS15
	6007	Trans-Canada Highway	6007	TS16
	9017, 1005	5	1005	TS05

It was found that there were several test sections in each major road and in the same climate conditions. From each road, at least one representative test section was selected which had sufficient data for the LCA. To cover a wide variety of climatic conditions and to reduce unnecessary road sections, representative road sections were selected for this component of the study.

In Table 3-5, it can be seen that all 22 selected test sections were labeled as Test Section (TS) for more easily processing and presenting them visually in the data analysis. Note that in the selection process, it was considered that representative road sections from each province were available. However, Prince Edward Island and Nova Scotia were excluded since they did not have any LTPP sections or lacked the available data needed for this component of the study.

The climate module of the LTPP database contains general environmental information from weather stations located near the test sections. In addition, a road section-specific statistical estimate based on as many as five nearby weather stations is available. These statistical estimates are called "virtual weather stations (VWS)" The following equation is used to weight the influence of operational weather station values based on the distance from the operational weather station to the virtual weather station.

$$V_m = \frac{\sum_{i=1}^{k} \frac{V_{mi}}{R_i^2}}{\sum_{i=1}^{k} \frac{1}{R_i^2}} \qquad \ldots (3.1)$$

Where

V_m= calculated data element for day m for the VWS

V_{mi}= value of data element on day m for operational weather station i

R_i= distance between operational weather station i and pavement project site

k = number of weather stations associated with the project site (up to 5)

To compute the annual freezing index, the following equation is used:

$$FI = \sum_{i=1}^{n}(0 - T_i) \qquad \ldots (3.2)$$

Where

FI = freezing index, degrees Celsius (°C) degree-days

T_i= average daily air temperature on day i, °C

n = days in the specified period when the average daily temperature is below freezing

i = number of days below freezing

When using this equation, only the days where the average daily temperature is below freezing are used. Therefore, the freezing index is the negative of the sum of all average daily temperatures below 0°C within the given period.

Table 3-6: Summary of climate characteristics for the road sections

Test Section ID	SHRP ID	Annual Precipitation (mm)	Annual Temperature (°C)	Annual Freezing Index (°C degree days)
TS01	0501	496.2	2.6	1265
TS02	0901	432.1	2.2	1697
TS03	0903	971.3	2.7	1558
TS04	0960	820.1	4.9	1120
TS05	1005	414.4	5.8	530
TS06	1021	1063.9	5.1	1031
TS07	1127	1188	4.8	1113
TS08	1684	1060.8	5.9	840
TS09	1801 (Newfoundland)	1424.1	5.6	421
TS10	1801 (Manitoba)	513.8	2.5	1722
TS11	1802	1103.3	5.8	795
TS12	1803	1434.6	4	787
TS13	2812 (Ontario)	963.6	8.3	489
TS14	2812 (Alberta)	378.1	3.6	1174
TS15	6006	1342.7	10.5	25
TS16	6007	1676.6	10.3	56
TS17	6410	378.1	3.1	1582
TS18	6804	1115.3	5.3	885
TS19	A310 (Saskatchewan)	499.1	2.3	1660
TS20	A901	420.9	5.2	864
TS21	AA01	507.1	2.8	1677
TS22	B310	1137.6	5.1	976

After selecting road sections based on representation and data availability, these were clustered to create a group of road sections with similar climatic conditions. Again, to create clusters from these road sections, the main consideration was the similarity in climatic conditions between the test sections although they are spatially located in different regions or provinces. To this end, a number of climatic parameters including annual precipitation, annual temperature, and annual freezing index data were collected from the LTPP database and further processed for developing clusters using statistical techniques including dendrogram and silhouette plot.

3.2.1.2 Hierarchical Clustering

For this component of the study, with climatic data from all the test sections, a dendrogram was developed and shown in Figure 3-1. It can be seen that some test sections with the similar climatic condition are in the same province. For example, TS1 and TS14 share the same climatic conditions (same height, 0.8 in the dendrogram) and they are both in the same province of Alberta; TS9 and TS12 also share the same climatic conditions (same height, almost 1 in the dendrogram) and are both in the same province of Newfoundland and Labrador.

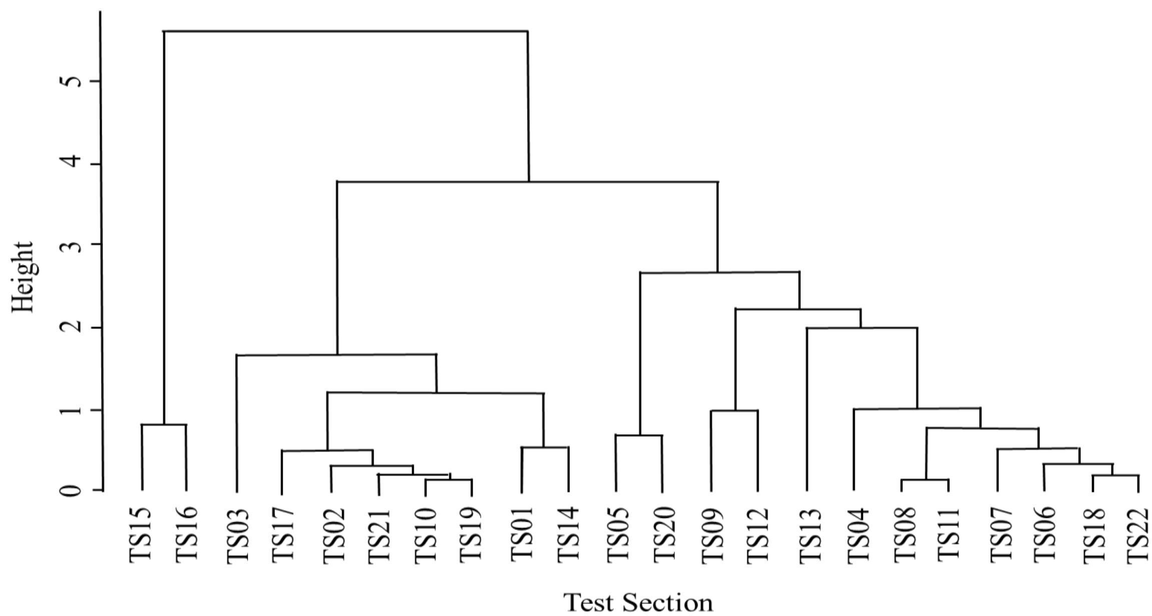


Figure 3-2: Cluster of test sections using a dendrogram

The same climatic conditions were not always found in the same province during this analysis. For example, TS5 and TS20, though they have the same climatic conditions (same height, 0.7). They are located in British Columbia and Alberta, respectively.

Furthermore, the dendrogram provides a hint of a cluster with similar climatic conditions through an approach known as the "cutting approach" (the analyst cuts the tree with an imaginary straight line horizontally along the same height). In our cluster dendrogram, if an imaginary line is drawn horizontally with a height of 5, it can produce two clusters:

- TS15-TS16 are in one cluster (British Columbia)
- The Remaining test sections are in a second cluster (Saskatchewan, Quebec, Ontario, Manitoba, Alberta, Newfoundland, and one section from British Columbia)

This shows that both clusters have a similar condition within their own cluster. But, one cluster has dissimilarity with another cluster. This means that there could be the same climatic condition in different provinces. If another imaginary line with a height of 3 is drawn, it can produce three clusters:

- TS15-TS16 (British Columbia)
- TS3,4 and 17 (Quebec), 2 and19 (Saskatchewan), 21 and 10 (Manitoba), 1 (Alberta)
- TS5 (British Columbia), 20 (Alberta), 9 and 12 (Newfoundland), 13, 4 and 22 (Ontario), 8, 18 and 11 (New Brunswick), 6 and 7 (Quebec)

In the same way, for any two clusters among these above three clusters, one will have different climatic conditions from the others. However, test sections in the same cluster have close similarities among themselves. Using this cutting approach different numbers of the cluster can be produced. The number of cluster will be based on optimum similarities within the cluster. This is estimated by the silhouette plot which has been described in the following subsection.

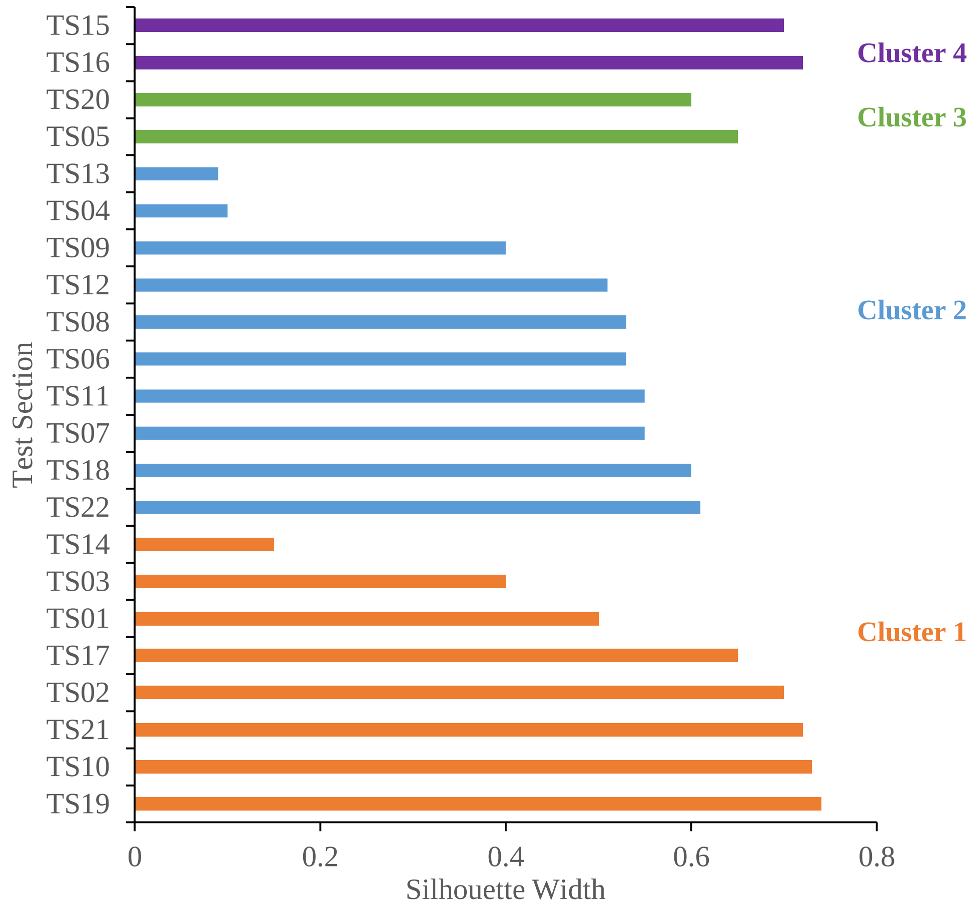


Figure 3-3: Refinement of clusters using Silhouette plot technique

Silhouette plot is drawn using trial and error method for different numbers of cluster i.e., 1, 2, 3, 4, 5 and more. When the number of cluster was 4, the silhouette width was optimum (0.53) as shown in Figure 3-2. Therefore, optimum similarities within cluster was available when four clusters were selected.

A question may arise as to why the silhouette width of TS13 and TS04 (both from Ontario) has less value compared with the rest of road sections in Cluster 2, in other words, TS13 and TS04 has a less intra-cluster similarity. If these two sections make another cluster, then the inter-cluster dissimilarity becomes much weaker which is not desired for good clustering.

The same phenomenon also happened for TS14 (Alberta) of Cluster 1. Therefore, optimum silhouette width was achieved which satisfied the desirable conditions for both inter-cluster and intra-cluster. Table 3-7 shows the threshold values of climate parameters used for classifying each cluster into a severity group based on overall climate conditions of the cluster. Climatic parameters of different clusters are classified in terms of severity (Table 3-8). The Canadian historical climate data from 1976-2005 is available through the climateatlas.ca web portal. Average annual values of precipitation, temperature and freezing index for major cities were collected from the portal. This information pattern helped to choose the severity classification threshold values for this component of the study. The average values of these factors in each cluster are summarized in Table 3-8, while the final severity level of each cluster is shown in Table 3-9.

Table 2-7: Climate severity classification

Climate Parameter	Unit	Low	Medium	High
Annual Precipitation	mm	< 400	400-1000	>1000
Annual Temperature	℃	< 3	3-10	>10
Annual Freezing Index	℃ degree days	< 100	100-800	>800

Table 3-8: Average value of climate parameter for each cluster

	Average Annual Precipitation (mm)	Average Annual Temperature (℃)	Average Annual Freezing Index (℃ degree days)
Cluster 1	523.14	2.76	1516.14
Cluster 2	1132.30	5.55	824.50
Cluster 3	417.65	5.50	697.00
Cluster 4	1509.65	10.40	40.50

Table 3-9: Overall climate conditions for each cluster

	Average Annual Precipitation	Average Annual Temperature	Average Annual Freezing Index
Cluster 1	Medium	Low	High
Cluster 2	High	Medium	High
Cluster 3	Medium	Medium	Medium
Cluster 4	High	High	Low

3.2.2 Analysis Approach

3.2.2.1 Assumptions and Functional Units

In a comparative study, a consistent functional unit must be chosen. The functional unit for pavements should represent physical dimensions and pavement performance. Performance requirements can include analysis period, traffic type, asphalt mixture composition, etc. The functional units considered in this component of the study were categorized and presented into physical, structural and annualized functional units (Ziyadi et al., 2017)

- **Physical functional unit:** In this component of the study, 1 km length of road was considered with a lane width of 3.7 m and all of the roads considered for analysis were major arterial roads (equivalent to US Interstate).

- **Structural functional unit:** Hot-laid asphalt mixture was considered in asphalt bound layers, while Granular A (a well-graded mixture of crushed gravel, sand and fines) was considered for unbound layers. Traffic loading considered for this component of the study included both light-duty vehicles (LV) and heavy-duty vehicles (HV), independently. HV included the vehicles which had a single-unit 2-axle and 6-tire or more, and the vehicles that had less than 6-tire considered as LV.

- **Annualized functional unit:** The time when the road is initially constructed is considered as the beginning year. During road usage, IRI value increases until minor rehabilitation (surface patching, overlay etc.) is performed. When this rehabilitation is finished, IRI value is significantly reduced. The time period from the beginning to immediately before this rehabilitation is considered the analysis period. In summary, the analysis period for the roughness impact from PVI (consequence of GWP) was considered as the elapsed time period from the initial construction to immediately before the rehabilitation, when the IRI values were dramatically reduced. In this component of the study, the range of this analysis period is 7-16 years. The GWP values were measured for each year and per 1000 AADT.

3.2.2.2 Life Cycle Phase and System Boundaries

For this component of the study, the LCA analysis was limited to the use phase. The use phase data was collected only from the Canadian LTPP database. The GWP values were considered as a result of PVI which again were the consequence of pavement roughness and deflection (pavement structural issue). Albedo was not considered because of their lesser significance in asphalt pavement (Kaloush, Carlson, Golden, & Phelan, 2008). Implementation of a carbonation scheme is also not applicable in asphalt pavement, therefore, this was also left out of the analysis.

3.3 Methodology III

Methodology III describes the methodology of the LCA model study in the Canadian context. The machine learning based LCA model for different phases of the pavement life cycle was developed.

3.3.1 Data Preparation

The necessary data collection and preparation technique for the component of the study is explained in this section. All LTPP road sections of Canada was used to find out the CO_2 emission. In order to get emission results, the pavement LCA framework was followed. The necessary calculation of formula is also described here.

3.3.1.1 LTPP Data

There are 141 Canadian road sections available in the LTPP database. This component of the study aims to use data of these road sections to develop the LCA model. Figure 3-4 shows the map of LTPP road sections. Almost all of the LTPP road sections are in the southern part of provinces. Yukon, Northern territories and Nunavut don’t have any LTPP road sections.

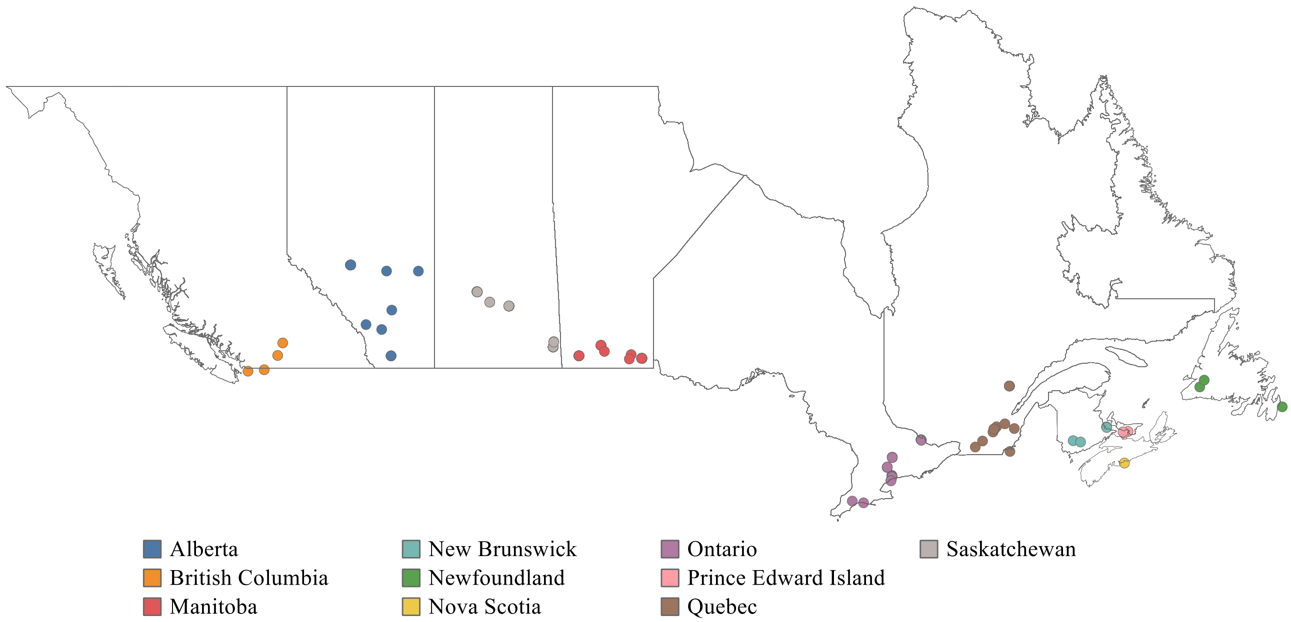


Figure 3-4: All (141) Canadian road sections from LTPP database

3.3.1.2 Formulation

Material Production and Initial Construction

Material production and initial construction is the first step we consider in an LCA study. For the modeling purpose, material production and initial construction phase is again classified in the asphalt layer and granular layer. The predictor variables considered for asphalt layers were: representative thickness, specific gravity of aggregate, asphalt binder, filler, percentage of coarse aggregate, fine aggregate, binder and filler. On the other hand, the predictor variables considered for granular layers were: representative thickness, specific gravity of aggregate, percentage of coarse aggregate and fine aggregate. CO_2 emission (gram) for material production was calculated based on Equation 3.3 which required volume (ft^3), density (lb/ft^3) and CO_2 emission rate (gram/ton).

$$CO_2\ emission_{\substack{material\\production}} = \sum_{i=material} Volume_i * Density_i * CO_2\ emission\ rate_i \quad ...(3.3)$$

Specific gravity from data Table 3-10 was multiplied by 62.4 lb/ft^3 to convert to density. The CO_2 emission rate is measured by the Equation 3.4. Carbon di-oxide emission from each dollar expenditure, $CO_2\ emission/\$_i$ was retrieved from the PaLATE database, which imported the original data from Economic Input-Output (EIO) LCA (Zimmerman, 1997).

$$CO_2\ emission\ rate_i = CO_2\ emission/\$_i * \frac{\$}{material\ production\ mass_i} \quad \ldots (3.4)$$

For construction purposes, the equipment model, fuel consumption rate and utility rate of that equipment model are important components. The density of diesel was considered constant, 852gram/litre. $CO_2\ emission\ rate\ per\ fuel\ mass$ for diesel was also constant, 3.16 gram of CO_2 per 1 gram of diesel. The general Equation 3.5 was used to calculate the emission.

$$\begin{aligned} Total\ CO_2\ emission \\ &= \sum_{\substack{i=pavement\\ layer}} \sum_{\substack{j=construction\\ equipment}} (Volume_i * Density_i * Fuel\ consumption\ rate_j \\ &* Density\ of\ fuel_j \\ &* CO_2\ emission\ rate\ per\ fuel\ mass_j)/\ Utility\ rate_j \quad \ldots (3.5) \end{aligned}$$

Maintenance Phase

As the pavement distress in Canadian road sections was experienced in the surface layer of pavement, only surface layer thickness was considered for the study. Crack filling for pavement cracks was considered. Crack filler material can be applied either hot or cold. Asphalt cement was applied using uncut flush fill configuration. The flush fill configuration was completed by placing into an uncut track. The sealant quantities for crack sealing was measured according to Equation 3.6.

$$Weight\ of\ sealant = Volume * Wastage\ factor * Density\ of\ crack\ sealant \ldots (3.6)$$

A 15% wastage factor and 1.12 specific gravity of crack sealant is usually considered. Another maintenance, patching was implemented in Canadian road sections. Patching mixture from the back of a dump truck is thrown into the distressed area. The usually used mixture is a stockpile patch. In "throw and roll"- the dump truck rolls over the patch one or two times to compact the patching mixtures. Residue asphalt binder content for the patching material was considered as 4%. The residue content is the asphalt binder that is left over after the water or solvent has evaporated from the asphalt emulsion or cutback. According to the typical stockpile mixture gradation, 57.6% coarse aggregate and 38.4% fine aggregate of total mixture was considered. Uniform thickness for all patching areas was 1 inch. Patching material density was the same as the density of asphalt mixture, 1.23 ton/yd^3.

Use phase

The use phase accounts for significant environmental impacts through various variables including rolling resistance effect in increased fuel demand (i.e., pavement vehicle interaction), albedo (solar radiation reflectibility of pavement surface), carbonation (absorption of carbon dioxide in pavement structure), night time illumination of the roads (energy demand) and generation of leachates from pavement materials (emission of chemical substances) (Ziyadi et al., 2017).In the pavement vehicle interaction (PVI) phenomenon, three pavement related variables are considered responsible for the PVI effect: pavement roughness, surface texture, and deflection.

From TRB special report 286, 2 m/km reduction in IRI value can reduce 1 to 2% of fuel consumption (National Research Council (US)., 2006). When considered texture depth, 0.71% of fuel consumption can be decreased from 0.44 mm texture depth reduction for cars with a 20mpg fuel economy (Benbow et al., 2007). Therefore, excess fuel consumption due to roughness and texture depth was measured based on the following Equations 3.7 and 3.8.

$$Excess\ fuel\ consumption_{Roughness} = (IRI_{observed} - Baseline\ IRI) * 0.0075 * Fuel\ consumption \quad \ldots (3.7)$$

$$Excess\ fuel\ consumption_{Texture\ depth} = (TD_{observed} - Baseline\ TD) * 0.0161 * Fuel\ consumption \quad \ldots (3.8)$$

3.3.2 Analysis Approach

3.3.2.1 Assumptions and Functional Units

There are 141 Canadian LTPP road sections which are functionally arterial road sections. Asphalt layer (original surface layer and AC layer below surface layer) and unbound granular layer (base and subbase) on the subgrade were considered in material production and initial construction.

From gradation report, aggregate that passes through 9.5 mm sieve and retained on 4.75 mm were considered coarse aggregate (CA). On the other side, aggregate that passes through 4.75 mm sieve were considered as fine aggregate (FA) and filler is mineral dust passing 0.074mm sieve. The volumetric percentages and specific gravity of the asphalt mixture materials were shown in Table 3-10 (Garber & Hoel, 2009; Kallas, Puzinauskas, & Krieger, 1962; U.S. Department of Transportation, 2009). Volumetric percentage ranges were different for initial construction and maintenance. The different ranges for different phase resulted from a trial and error method, which led to the best performance level (Lavin, 2003).

Table 3-10: Volumetric percentages and the specific gravity of asphalt mixture components

Components of Asphalt Mixture	Volumetric Percentage Limits		Specific Gravity
	Initial Construction	Maintenance	
Asphalt binder	4-6	3.5-4	1
CA	48-55	20-30	3
FA	35-40	60-80	3
Filler (limestone)	5-8	0-2	2.6

Percentage of CA and FA for unbound granular layers (base and subbase) were selected as shown in Table 3-11. The recommended values of aggregate percentage are specified in provincial road construction specifications. The different values are recommended by transportation agencies based on the availability of materials, climatic conditions, and function in corresponding provinces.

Table 3-11: Volumetric percentages of aggregates in unbound granular layers (H. Soliman et al., 2014)

Canadian Provinces	CA(%)	FA(%)
Alberta	58	42
British Columbia	67.5	32.5
Manitoba	75	25
New Brunswick, Newfoundland and Labrador, Nova Scotia, Prince Edward Island	66.1	33.9
Ontario	61.5	38.5
Quebec	66.1	33.9
Saskatchewan	68.5	31.5

In both construction and maintenance phases, equipment were used for asphalt paving, milling, excavation, placing and compaction. For this component of the study, a specific model of diesel-fueled equipment was considered. Engine capacity, productivity and fuel consumption are summarized in Table 3-12.

Table 3-12: Equipment properties for construction and maintenance

Activity	Equipment	Model	Engine Capacity (horsepower)	Utility Rate (ton/hr)	Fuel Consumption rate (liter/hr)
Asphalt Paving	Paver	Dynapac F30C	196	2400	49.11
	Pneumatic roller	Dynapac CP132	100	668	26.12
	Tandem roller	Ingersol Rand DD110	125	285	32.65
Milling	Milling machine	Wirtgen W2200	875	1100	156.20
Excavation, placing and compaction	Excavator	John Deere 690E	131	315	34.22
	Vibratory soil compactor	Dynapac CA 262D	174	1832	27.56

Longitudinal and transverse cracking were very common pavement distress in arterial road sections in Canada. Crack sealing and patching was usually used for low and medium severity. The following Table 3-13 describes the crack width value that was considered for this component of the study. The value was selected as the median of the range of each severity of the crack.

Table 3-13: Pavement distress and considered crack width for this component of the study

Distress Type	Severity Level	Range of Crack Width		Crack Width Considered in This Component of The Study
		Minimum	Maximum	
Longitudinal/transverse cracking	Low	0	6mm	3mm
	Medium	6.1mm	19mm	12.5mm
	High	19.1mm	-	20mm

3.3.2.2 Life Cycle Phase and System Boundaries

Material production, initial construction, maintenance and use phase were considered in this component of the study as shown in Figure 3-5. For the modeling purpose, material production and initial construction phase is again classified in asphalt layer and granular layer. The predictor variables considered for asphalt layers were: representative thickness, specific gravity of aggregate, asphalt binder, filler, percentage of coarse aggregate, fine aggregate, binder and filler. On the other hand, the predictor variables considered for granular layers were: representative thickness, specific gravity of aggregate, percentage of coarse aggregate and fine aggregate. Material and equipment transportation distance from plant (mixing plant) to site (road construction location) was not considered in this component of the study because of data unavailability in the LTPP database.

Time schedule of maintenance activity was different for different road sections. Therefore, frequency of maintenance in 20 years is taken as functional unit in this component of the study. The predictor variables considered in this phase were surface layer thickness, average monthly precipitation, temperature, freezing index, maintenance frequency (number of maintenance per 20 years), pavement age at first maintenance, patching area, length of crack sealing, crack severity and traffic load.

Three predictor variables were considered for the use phase, which are average IRI, average texture depth and traffic load. International roughness index (IRI) predictor variable was considered for pavement roughness. The sensor measured texture depth was considered to represent surface texture. Due to the unavailability of field-measured deflection value in Canadian road sections, deflection was not considered. Albedo measures the ability of a pavement surface to reflect solar radiation. Albedo was not considered because of their lesser significance in asphalt pavement (Kaloush et al., 2008). Carbonation is the process of absorption and storage of carbon dioxide in pavement structure, while forming a bond with calcium oxide or calcium hydroxide. Implementation of a carbonation scheme is also not applicable in asphalt pavement. Therefore, this was also left out in this component of the study. Electrical energy is required to illuminate the roadways at night, for enhancing road safety. If the pavement type and composition can affect the energy demand for illumination, this component can play a significant role in the overall emission footprint of the pavement. Furthermore, some pavements can adversely affect groundwater and soil through the leaching of pollutants (Santero, Masanet, & Horvath, 2010). These two variables are excluded to keep our model simple and due to data unavailability in the LTPP database.

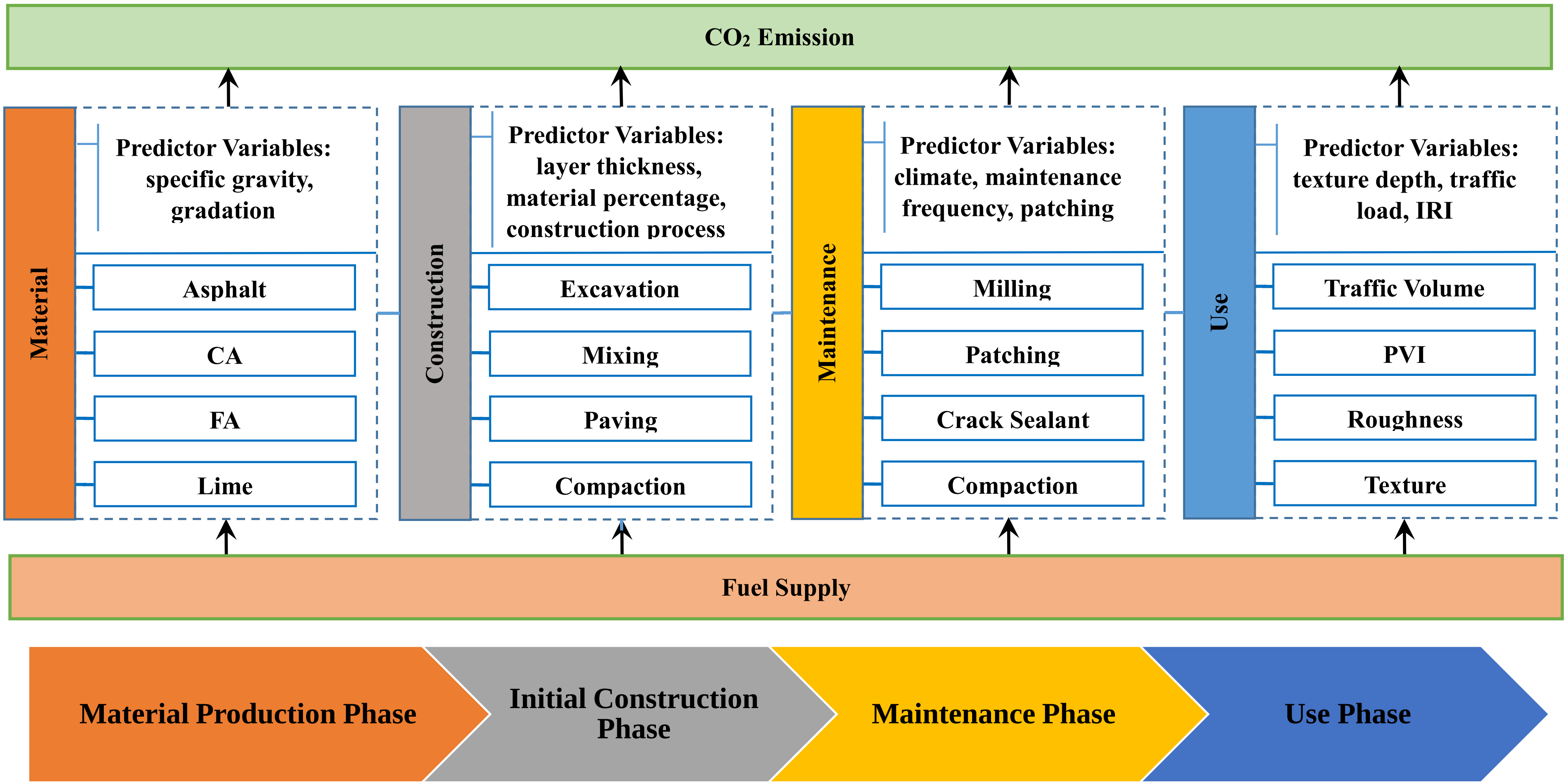


Figure 3-5: Pavement LCA framework considered in the study

3.3.3 Data Modelling

The generated data (predictor variables and CO_2 emission) for 141 LTPP sections through SQL coding were used for data modeling. Before applying the machine learning algorithms, further data preparation was required. In the development of machine learning model, the major task is to preprocess the data.

3.3.3.1 Importing Libraries

The first step of the data preprocessing template is to import the essential python libraries. A library is a tool that can be used to perform a specific job. The first one was numpy. This library contains mathematical tools. Basically, this library was used to include any types of mathematics in the code. The second essential library is matplotlib.pyplot. This library helped to plot graphs. The last library that was essential for the template is pandas library which aims to import data sets and manage datasets.

3.3.3.2 Importing Dataset

Before importing the dataset, there needs to specify the working directory. The data that will be used to code should be stored in that working directory of the computer. The data should be saved as comma-separated value (.csv) file. Pandas library was used to import the dataset.

After importing the dataset, it is important to distinguish the matrix of independent variable (predictor) and the dependent variable (response). A new matrix X was created which had all of the columns of independent variables. Following the similar syntax, the matrix of the dependent variable named y was created. Here, it needs to remember that the indexing in python starts at zero.

3.3.3.3 Handling Missing Values

Missing data is highly expected in real life dataset. To make the model as efficient as possible, it is required to handle missing data. If there are very few observations with missing data compared with a lot of observations, the removal of such observation can be one method. Another idea which can be the better idea is to take the mean of the existing values of the column. In python, using imputer class from scikit learn or sklearn preprocessing library, the missing data can be filled up by the mean of the existing column values. As in this component of the study, the handling of missing values was performed already in SQL, the python code was not required.

3.3.3.4 Encoding Categorical Variable

Typically, any categorical data refers to discrete values. These discrete values are normally a specific finite set of categories or classes. These discrete values can be text or numeric or even unstructured data like images. There are two major classes of categorical data, nominal and ordinal. Movie, music and video game genres, country names, food and cuisine types are few examples of nominal categorical attributes. The ordinal categorical variable could be shoe sizes, education level and employment roles, etc. As machine learning model is one kind of mathematical model, intuitively it can be understood that the categorical value will create a problem. Therefore, it is important to code categorical variables into meaningful numbers.

In order to encode the categorical variables, LabelEncoder class from sklearn.preprocessing library was imported. After that an object named labelencoder_X using LabelEncoder class was created. Using this object, the column of the categorical variable was transformed into numbers. For example, if the pavement layer column had four categorical values: surface, base, subbase and subgrade. The object would transform it to 0,1,2 and 3, respectively. However, the model would understand that the subgrade has a higher value than the subbase or surface has a lower value than base. But this was not the real case, these were just four categories that didn't have any relational order among them. In order to resolve this issue, a dummy variable concept was used. As a result, four columns would be generated instead of one column. Each column would have binary values. OneHoteEncoder class was used to create such dummy variables.

3.3.3.5 Splitting Dataset into Train and Test Set

The dataset which is used for developing machine learning model needs to be split into training set and test set. As the name of machine learning itself refers to the machine which will learn something. In other words, the model will learn from the data to make predictions. Machine learning model will try to understand the correlations of the data using the training data set. After that using the slightly different test set users eventually test the performance of the model. Therefore, there needs two datasets. The performance of the test set should not be much different from the training set performance. Train_test_split class was imported from sklearn.model_selection library. The test set was chosen 20% of the entire dataset and 80% as training data set. In this component of the study, we compared the performance of the different models. For comparison purpose, the test dataset needs to be similar. Therefore, 20% of dataset was kept similar for developing different models for the same primary dataset. That's why, in the code, the random state value was zero.

3.3.3.6 Creating Model Template

Particular classes for linear regression, decision tree regression and support vector regression were imported from sklearn machine learning library. For polynomial regression, linear regression class was used. Before using the linear class for polynomial regression, the numpy array was reset according to second-order polynomial. For decision tree regression, supported criteria was "mse". Hence, the mean squared error was kept minimum for feature selection. Besides, the minimum number of samples at a leaf node was considered 5 which is widely accepted. For SVR, linear kernel was taken because of better performance for pavement LCA rather than rbf kernel. Regularization value C was kept 1 and epsilon value was 0.1. After modeling operation was performed, coefficient values using the ordinary least square method was calculated for linear and polynomial regression. The weight values of a hyperplane in SVR was also calculated. Thus, the template for different algorithms were set up which was used further with little modification for each phase of pavement life cycle.

CHAPTER 4

ENVIRONMENTAL IMPACT ASSESSMENT FOR M&R

Environmental impact assessment for four maintenance and rehabilitation (M&R) is described in this chapter. Environmental emissions from LCA for each M&R are initially reported from section 4.1-4.4 and in the following section 4.5 shows and explains the comparative assessment among M&R techniques.

4.1 Emission for Rout and Sealing

The first M&R practice analyzed was the rout and sealing technique which produced 35,186 kg of CO_2 equivalent as GWP, as summarized in Table 4-1. As the compound-measured impact category, GWP became the highest emissions and the lowest value was for ozone depletion potential, 6.94×10^{-6} kg, measured as released Chlorofluorocarbon-11 (CFC-11) in kg.

4.2 Emission for Asphalt Patching

Similar to the rout and sealing technique, the compound-measured environmental impact category with the highest emission value was GWP for asphalt patching (50,396 kg). As it can be seen in Table 4-1, the lowest impact category was ozone depletion potential, which resulted in 3.367×10^{-6} kg of Chlorofluorocarbon-11 (CFC-11).

4.3 Emission for HIR

The third M&R technique analyzed for the case study was HIR (Table 4-1). When reviewing the results in the compound-measured environmental impact categories, the GWP produced the highest emissions, and the lowest emission was the ozone depletion potential. As the representative of GWP, 14,416 kg of CO_2 was emitted out whereas ozone depletion potential occupied 3.83×10^{-7} kg of CFC-11.

4.4 Emission for CIR

The fourth M&R technique to be analyzed for the case study was CIR. Similar to all methods studied, the CIR produced higher values for the GWP category and lowest values for the ODP category as shown in Table 4-1. The GWP emissions for CIR technology equated to 10,450 kg of CO_2 equivalent. For the ODP, the CIR technology equated to 2.24×10^{-7} kg of CFC-11.

Table 4-1: Assessment of the environmental impact categories for different M&Rs

Impact Category	Rout and sealing			Asphalt patching			HIR			CIR		
	Materials and Equipment	Transport	Total	Materials and Equipment	Transport	Total	Materials and Equipment	Transport	Total	Materials and Equipment	Transport	Total
GWP (kg CO_2 eq.)	34,931	255	35,186	48,453	1,943	50,396	12,244	2,173	14,416	8,135	2,315	10,450
Acidification Potential (kg SO_2 eq.)	376	2.5	378	2267	19	245	108	21	129	97	22	119
HH Particulate (kg $PM_{2.5}$ eq.)	29.36	0.13	29.50	14.69	1.03	15.72	7.59	1.16	8.75	6.97	1.23	8.21
Eutrophication Potential (kg N eq.)	14.75	0.16	14.91	10.24	1.16	11.40	5.17	1.30	6.47	4.54	1.38	5.92
Ozone Depletion Potential (kg CFC-11 eq.)	6.94×10^{-6}	8.93×10^{-9}	6.94×10^{-6}	3.299×10^{-6}	6.78×10^{-8}	3.367×10^{-6}	3.08×10^{-7}	7.58×10^{-8}	3.83×10^{-7}	1.44×10^{-7}	8.08×10^{-8}	2.24×10^{-7}
Smog Potential (kg O_3 eq.)	3,771	80	3,851	3,391	585	3,980	1,422	659	2,082	1,173	703	1,876

4.5 Comparative Analysis among M&R

The results show that the GWP and ODP were the highest and lowest impacts, respectively, for all the M&R techniques. Since the project used diesel as the energy source for operation, CO_2 equivalent showed a significant amount in the results because of 2.68 kg CO_2 production per liter diesel consumption. However, the emissions of CFC-11 were minimal in the project.

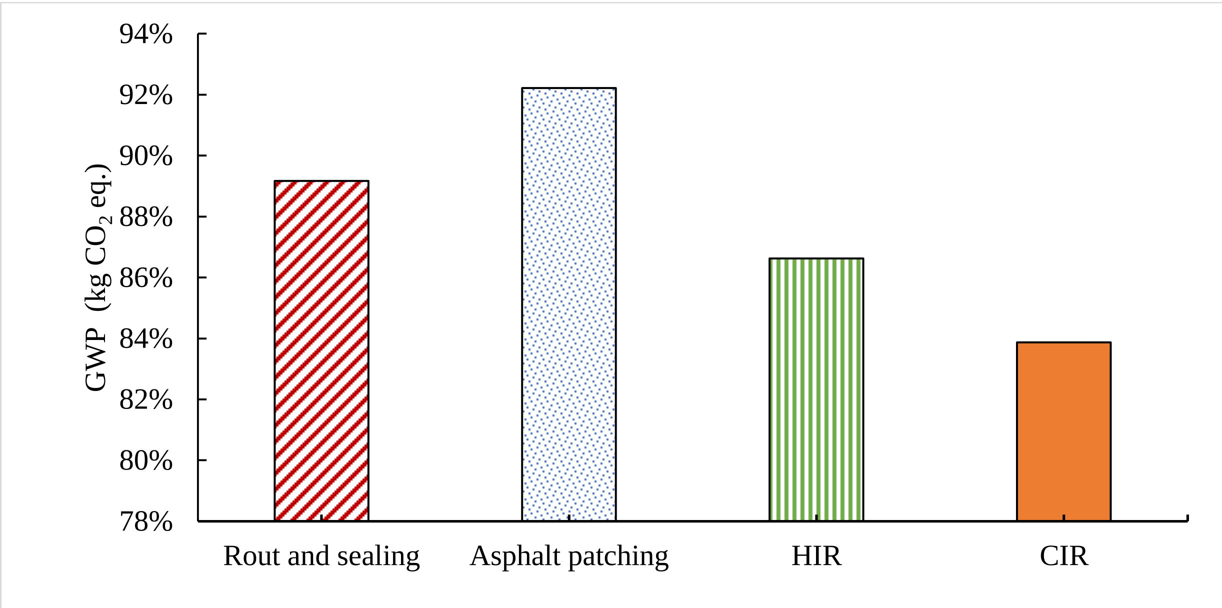


Figure 4-1: Global Warming Potential percentage values of the four analyzed M&R practices

Figure 4-1 presents GWP contribution as the highest compound-measured emission category among the selected four maintenance processes. The CIR technique produced the lowest CO_2 eq. emissions, 83.87% during its project life closely followed by HIR technique, which produced 86.63% of CO_2 eq. emissions. For asphalt patching, the CO_2 emission resulted the highest percentage (92.22%) and thus became the least suitable option among four studied M&R methods in terms of GWP.

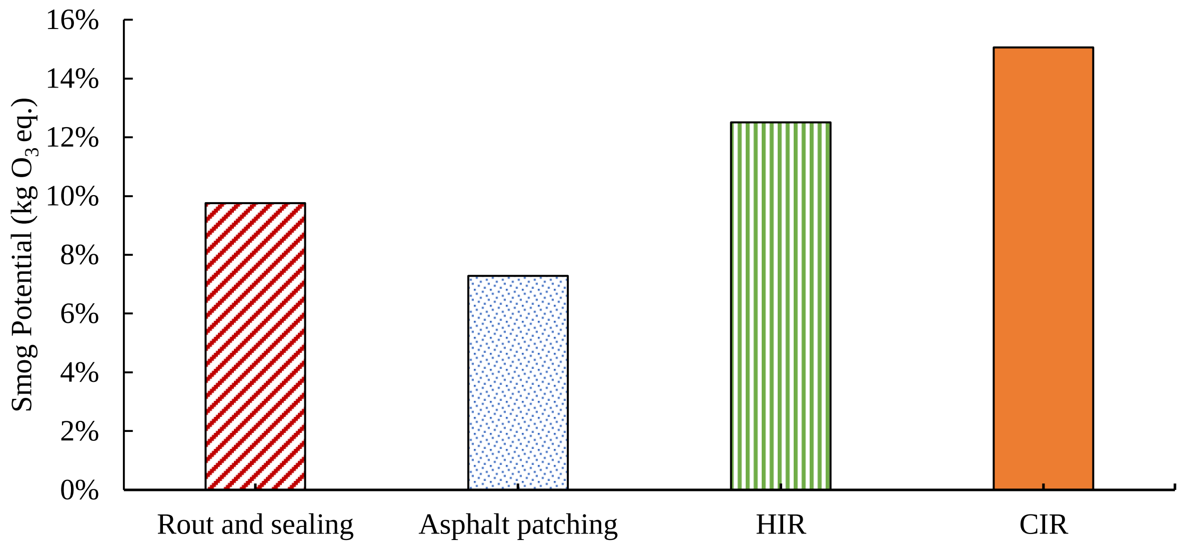


Figure 4-2: The Smog Potential percentage values of the four analyzed M&R practices

The results from the pavement patching and HIR techniques can be explained based on the total number of equipment and equipment time used. Both included technology that used more diesel as fuel consumption and produced high temperatures during manufacturing materials and thus resulted in a higher emission of CO_2. For the CIR methods, less machinery was used and no on-site heating machinery was required, leading to less diesel fuel required for operation, hence producing lower CO_2 emissions. Smog potential was the second largest contributor of emissions for each practice. In addition, 15.06% emission of smog potential was from CIR followed by HIR (12.51%), rout and sealing (9.76%) and finally asphalt patching (7.28%) as shown in Figure 4-2. The reasons behind the higher smog potential of CIR rather than HIR need to be investigated in further research.

Besides GWP and smog potential, the other emission factors combined to carry approximately 1% of the environmental burden, where rout and sealing and CIR had the greatest impact on the percentage of acidification potential (0.96%) followed by HIR (0.78%) and asphalt patching (0.45%). HH particulate and eutrophication potential percentages were very low for all of the four M&R techniques (less than 0.1%).

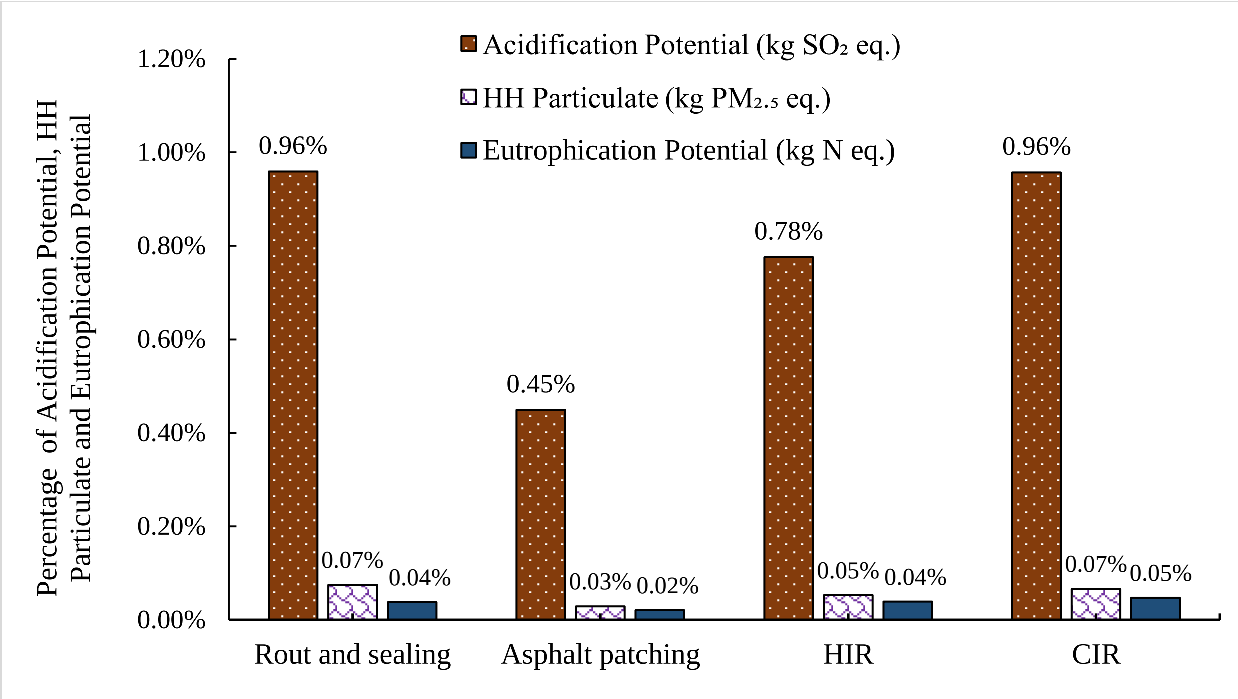


Figure 4-3: Acidification Potential, Human Health Particulate, and Eutrophication Potential Emissions: The percentages of three compound-measured impact categories

CHAPTER 5

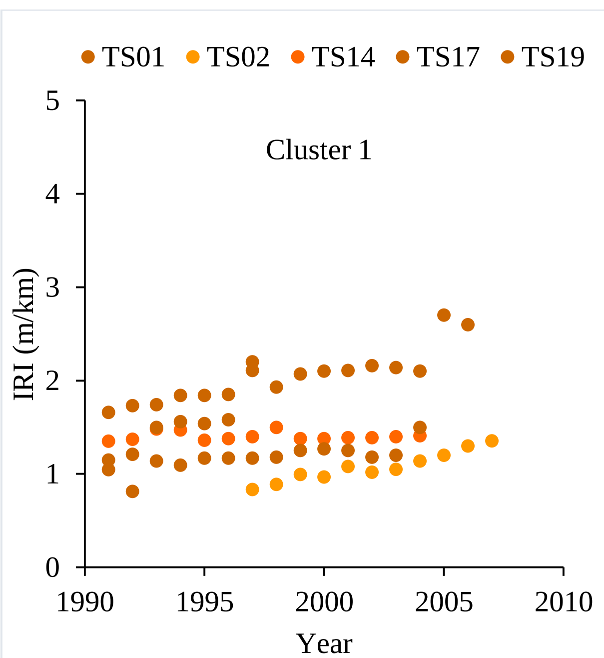
ENVIRONMENTAL IMPACT ASSESSMENT FOR PVI

Environmental impact assessment for pavement vehicle interaction (PVI) of the pavement use phase is described in this chapter. Two major components for PVI are pavement roughness and deflection during traffic load. Environmental emission, particularly GWP emission from LCA for pavement roughness and pavement deflection are described from section 5.1-5.2. PVI effect is explained through the clustering of Canadian LTPP sections.

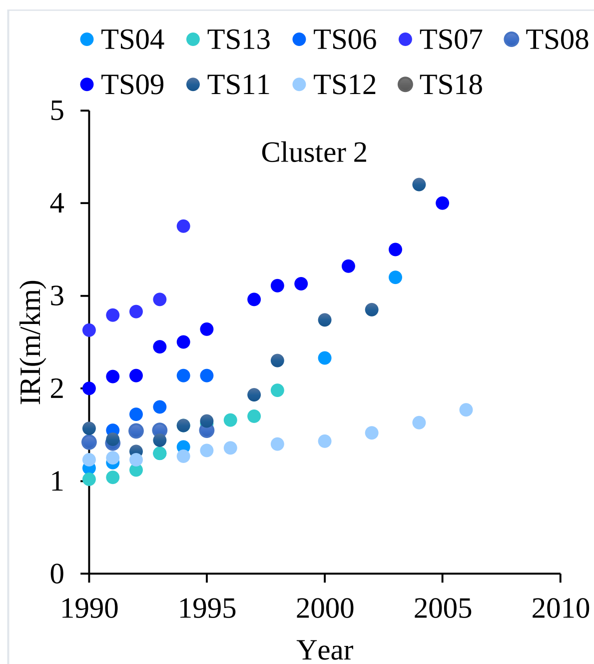
5.1 Pavement Roughness Effect

Pavement roughness is an important factor in determining the PVI effect. GWP value was selected to measure and compare the PVI effect. Usually, pavement IRI rises gradually after initial construction.

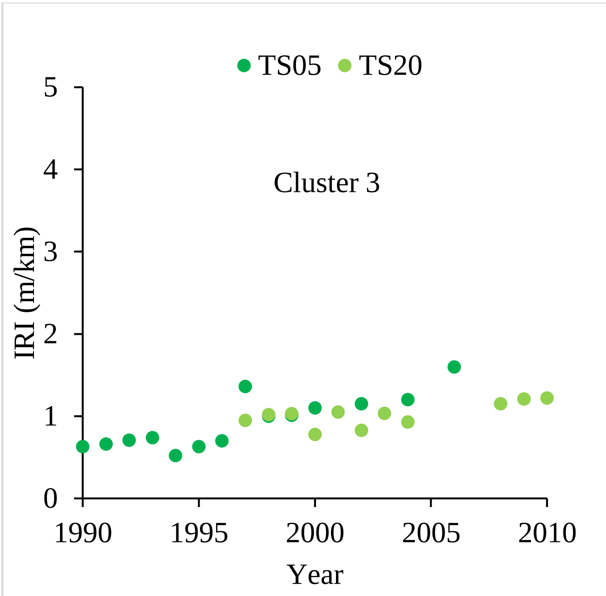
Figure 5-1 (a-d) shows the IRI value for the road sections of each cluster. It was noticed Cluster 2 had, in general, the highest IRI values, followed in this mathematical form: Cluster 2 > Cluster 1 > Cluster 3 > Cluster 4



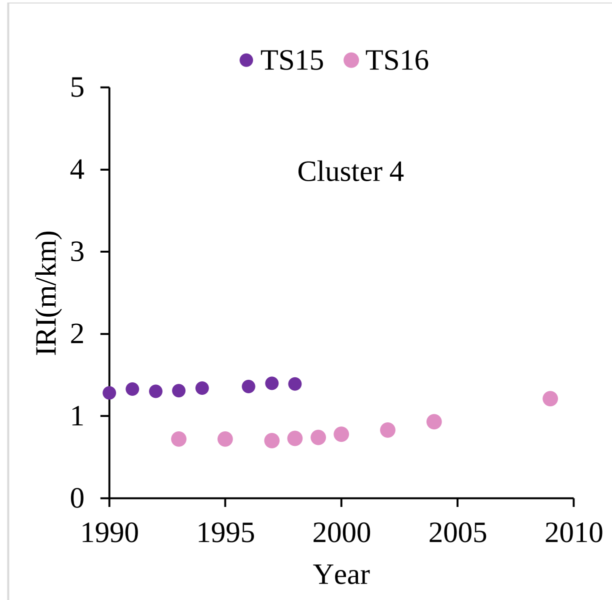
(a)



(b)



(c)



(d)

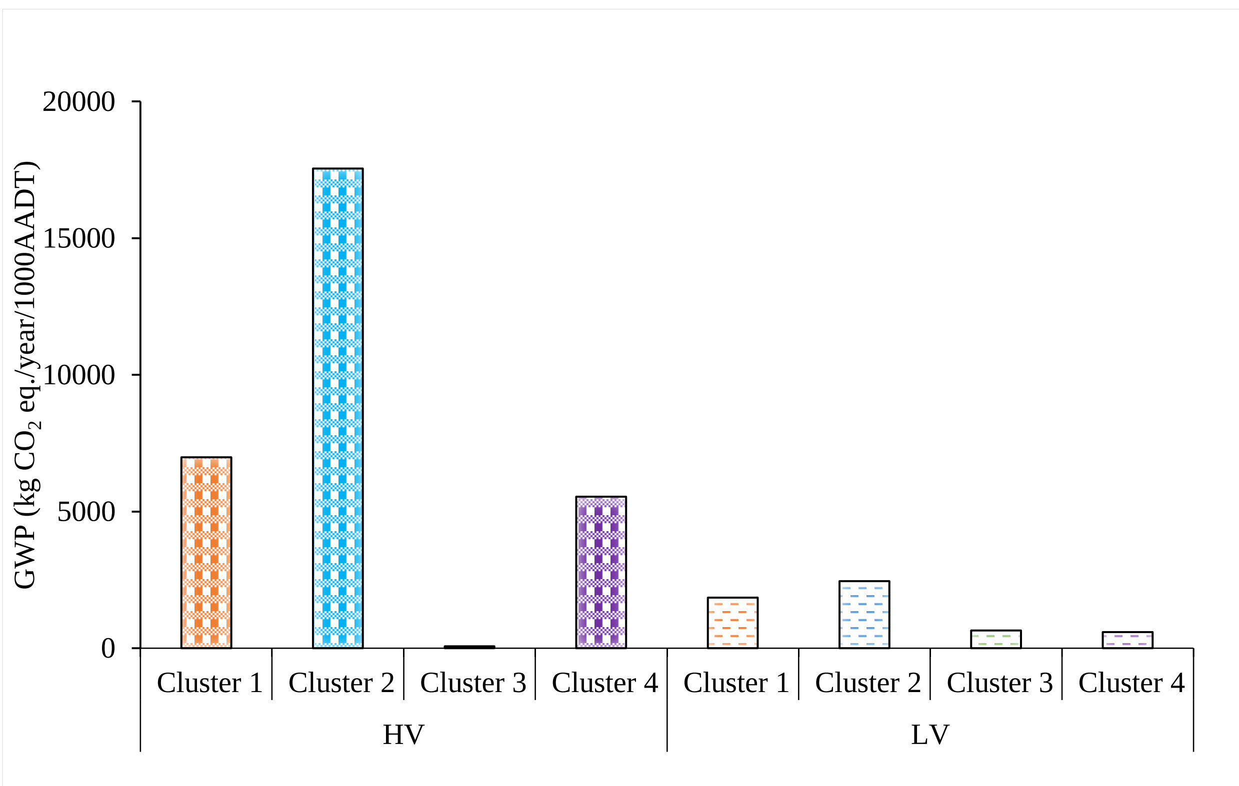


Figure 5-1: IRI over time for each cluster (a-d) and GWP emission due to roughness based PVI (e)

In Cluster 2 regions, please recall that the climate parameter included high annual precipitation, high annual freezing index, and medium annual temperature. It means that there was a high probability that the air void in the soil layers in the pavement structure was filled and saturated with water because of high annual precipitation. Afterward, in the winter months, these waters became frozen and caused frost heave. This frost heave can lead to an expansion of volume by 9% due to phase change from water to ice and can affect the smoothness of pavement and ride quality. In the spring season, these frosts melted and could cause other pavement distresses such as pothole and differential settlement. Furthermore, these melted waters from the top layers went to the bottom layers. These trapped liquid waters weakened the base and subgrade layers which eventually deteriorated the performance of the entire pavement structure. This deterioration was accelerated by the dynamic loading of vehicles and resulted in an increase of IRI, significantly.

Cluster 1 road sections were located in a region where comparatively there was less probability of water saturation in the pavement layer because of medium annual precipitation. However, these road sections were located where the annual average temperature was relatively low. This probably led to having a high freezing index in this cluster. Therefore, the IRI increase rate in Cluster 1 is relatively less than Cluster 2.

On the other hand, because of medium precipitation, temperature and freezing index, the increase rate of IRI for Cluster 3 was relatively less compared with Cluster 3 and 4.

The IRI was increased at a very slow rate which was found in Cluster 4. The annual temperature was high and freezing index was low. This indicates that the road sections in this cluster experienced longer spring and summer seasons. As a result of the long sunny periods, the precipitated rain waters could easily either evaporate or drain out from the pavement structure as there was less chance of freezing.

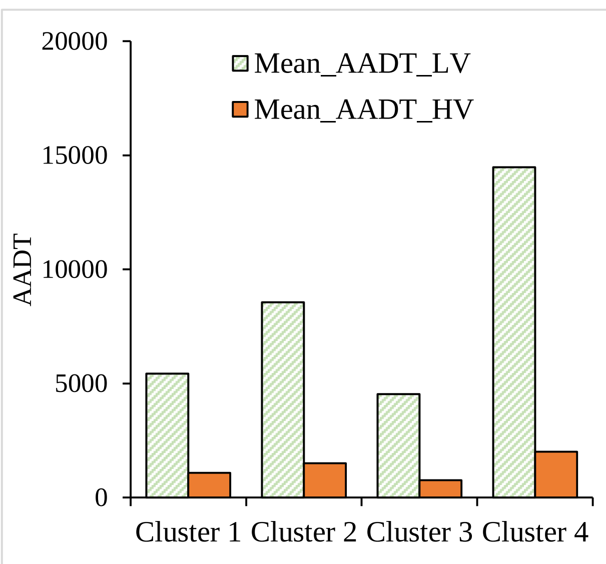
The GWP values estimated using the model described in the previous section are summarized in Figure 4. It can be seen on this figure that road sections of Cluster 2 had emitted the highest amount of GWP, which was due to the high IRI values of these roads. The GWP values for other clusters, for the analysis period per 1000 AADT light vehicles follows this pattern: Cluster 2 > Cluster 1 > Cluster 3 > Cluster 4, the same hierarchy of IRI increase rate.

When the GWP values are compared among the traffic loading type, it can be seen that, among the heavy and light vehicles, the heavy vehicle had significantly high GWP emission than light vehicle.

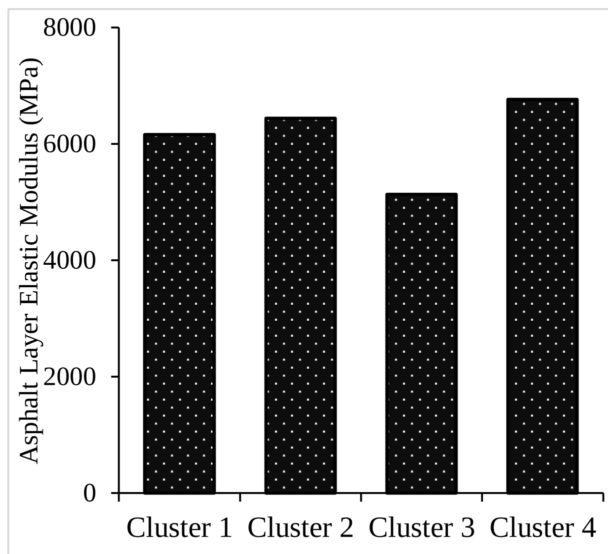
5.2 Pavement Deflection Effect

According to PVI Gen II model, vehicle load has a strong direct relationship with deflection. On the contrary, subgrade stiffness, surface layer elastic modulus and thickness of the asphalt layer have an inverse and comparatively less strong relationship. Among these inverse factors of deflection, the thickness of the asphalt surface layer has comparatively higher significance.

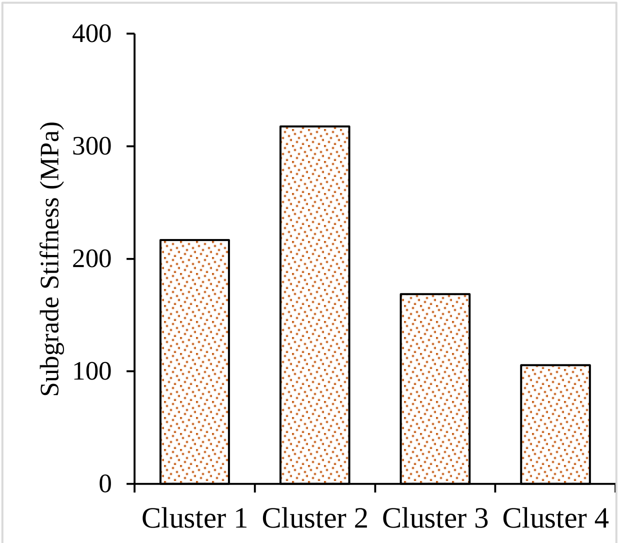
As shown in Figure 5-2, the GWP emission gradually increases from Cluster 1 to Cluster 4. Cluster 4 has the maximum vehicle load both for HV and LV, and the minimum subgrade stiffness. These properties, when combined, make the highest GWP emission for Cluster 4.



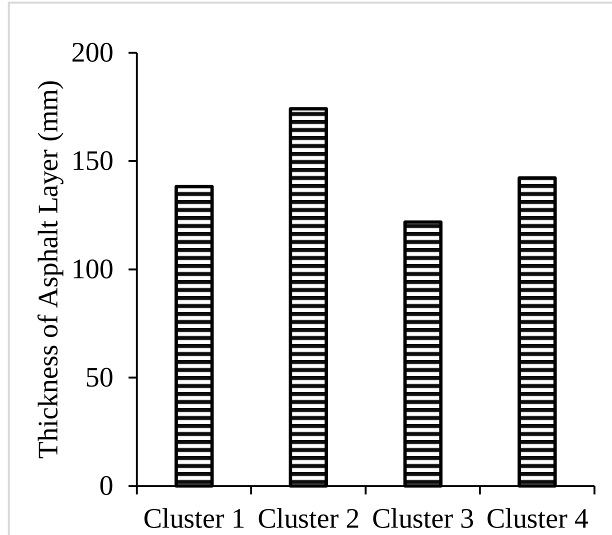
(a)



(b)



(c)



(d)

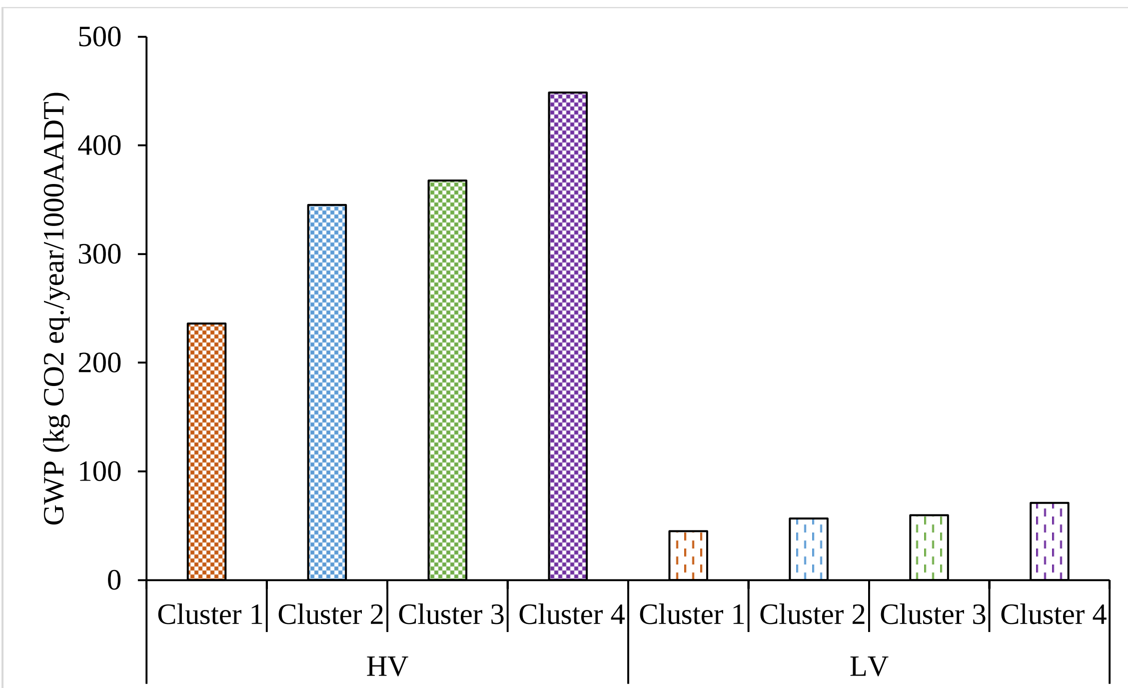


Figure 5-2: Deflection parameters in different clusters (a-d) and GWP due to deflection based PVI (e)

As for Cluster 3, though it has the lowest number of vehicle loads for both types, it has the minimum surface layer elastic modulus and asphalt layer thickness, which make Cluster 3 the second highest zone for the GWP emission.

As pavement material engineers do not have control over growing traffic loading, only material properties and pavement design can be improved to restrict pavement deflection based PVI. Therefore, subgrade stiffness can be increased in Cluster 4 to reduce high GWP emission. For Cluster 3, the elastic modulus of the asphalt layer and pavement design thickness can be upgraded to reduce further emission.

CHAPTER 6

LIFE CYCLE ASSESSMENT MODELS

This chapter describes the significance of model predictor variables, model parameter values and model accuracy for each phase of the pavement life cycle. In the last section of this chapter, a comparative analysis of CO_2 emission for different provinces using the models is also described.

6.1 Significance of Predictor Variables of Models

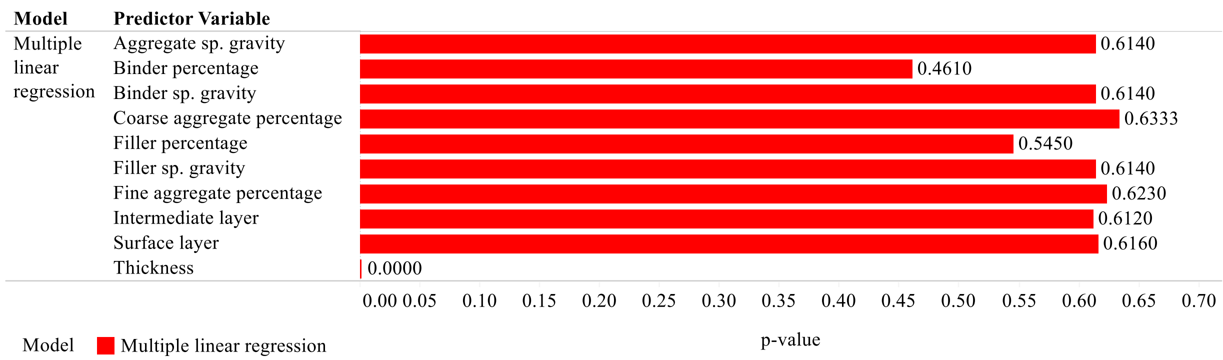
Four different types of models were developed for each pavement life cycle phase. *P*-values were considered for Conventional regression models (i.e., multiple linear and polynomial regression) in order to understand the significance of predictor variables. For SVR and decision tree regression, importance values were considered to understand the significance of predictor variables. In SVR, the squares of weights were used as importance values, whereas feature importance/ Gini importance values were used for decision tree regression.

6.1.1 Material Production and Initial Construction Phase

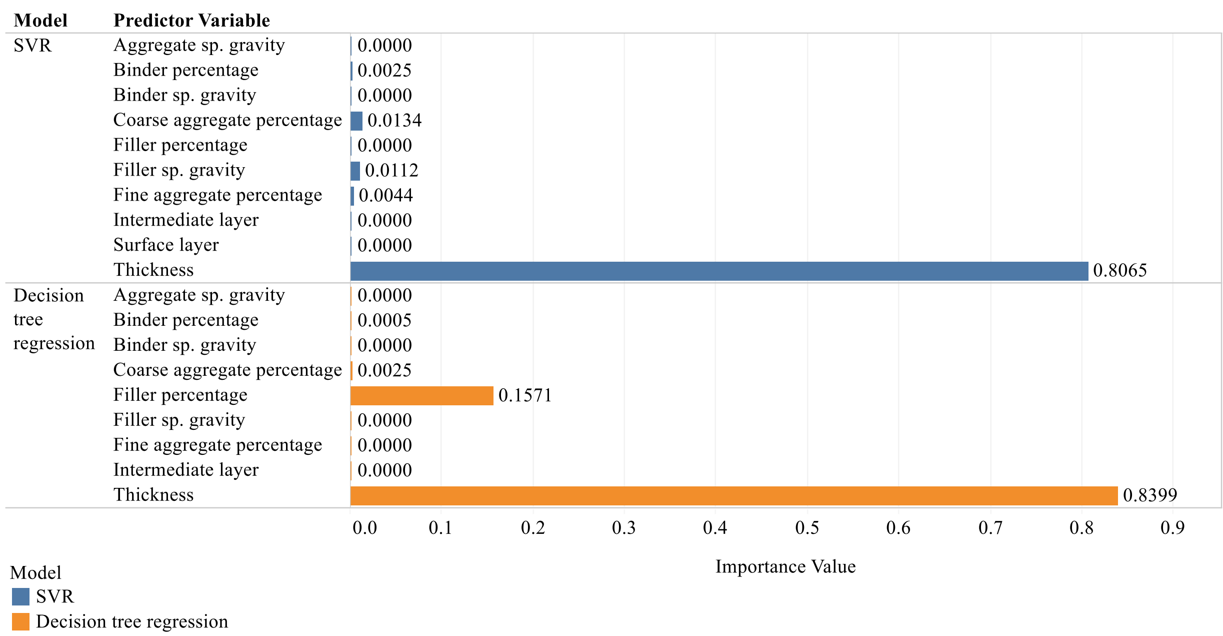
For material production and initial construction phase of asphalt layer in Canadian LTPP sections, three models are developed: multiple linear regression, SVR and decision tree regression. 10 predictor variables could generate 65 terms considering a second-degree polynomial regression. In order to avoid such complexity, polynomial regression is not considered in models for asphalt layer.

According to Figure 6-1(i) for multiple linear regression of the asphalt layer, only asphalt layer thickness shows significant (*p*-value less than 0.05) CO_2 emission. From the SVR model hyperplane, it was found the weight of thickness (0.898) is maximum compared with other predictor variables. Therefore, the square of weight of thickness (0.8065) will be the highest and as a result, thickness of asphalt layer has the largest relevance with the emission. From the decision tree regression, thickness and filler percentage have a large impact on emission according to importance feature values. The thickness layer has the largest impact (0.8399) among the predictor variables.

According to Figure 6-2(i) for multiple linear regression, all of the base predictor variables are significant (*p*-value less than 0.05). Except for base layer-coarse aggregate interaction, the second degree of thickness and subbase-coarse aggregate percentage interaction (p-values much higher than 0.05), the rest of the interactions show significance according to the polynomial regression model.

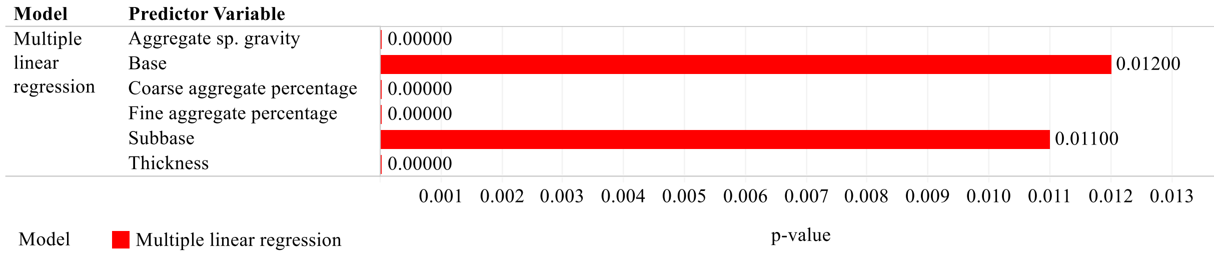


(i) *p*-value of predictor variables in multiple linear regression model

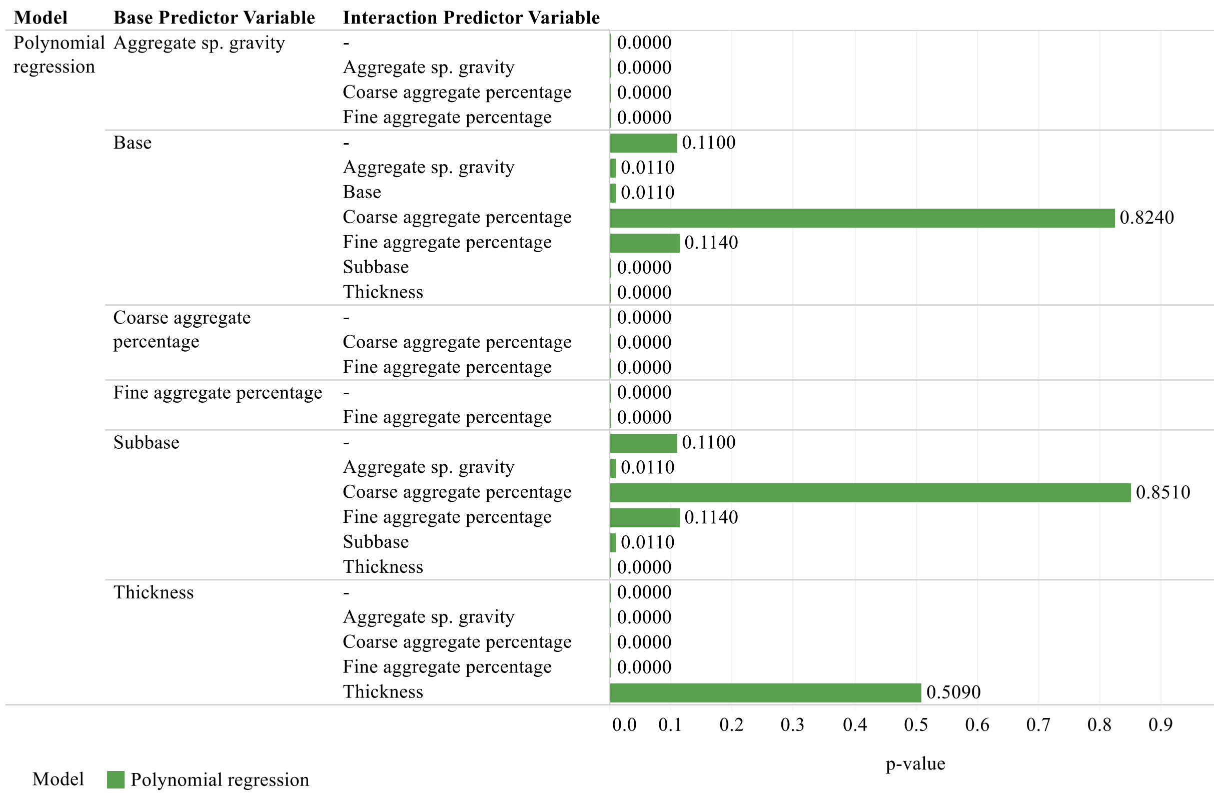


(ii) Importance value of predictor variables in SVR and decision tree regression model

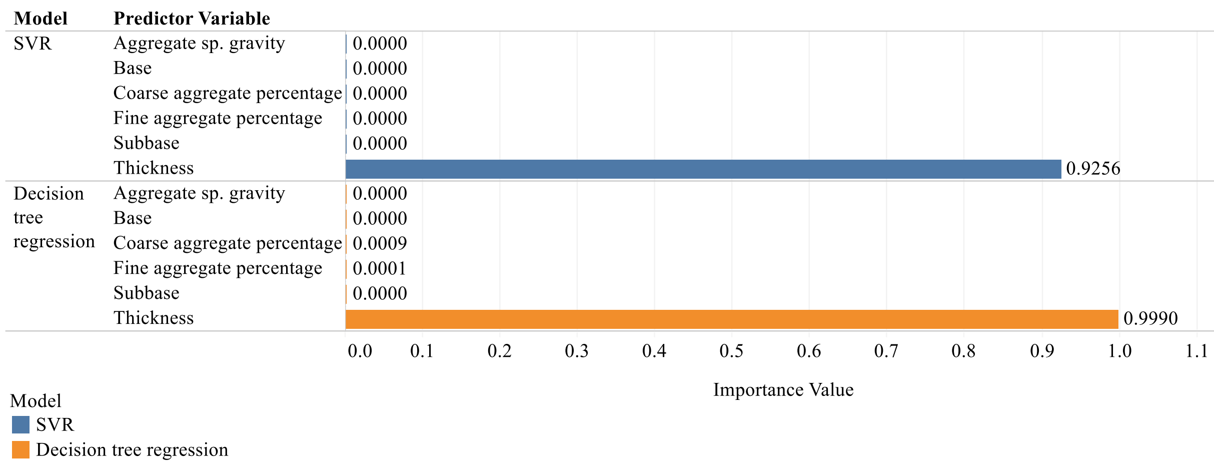
Figure 6-1: Visualization of importance value of models' predictor variables for asphalt layer in material production and initial construction of pavement



(i) *p*-value of predictor variables in multiple linear regression model



(ii) *p*-value of predictor variables in polynomial regression model



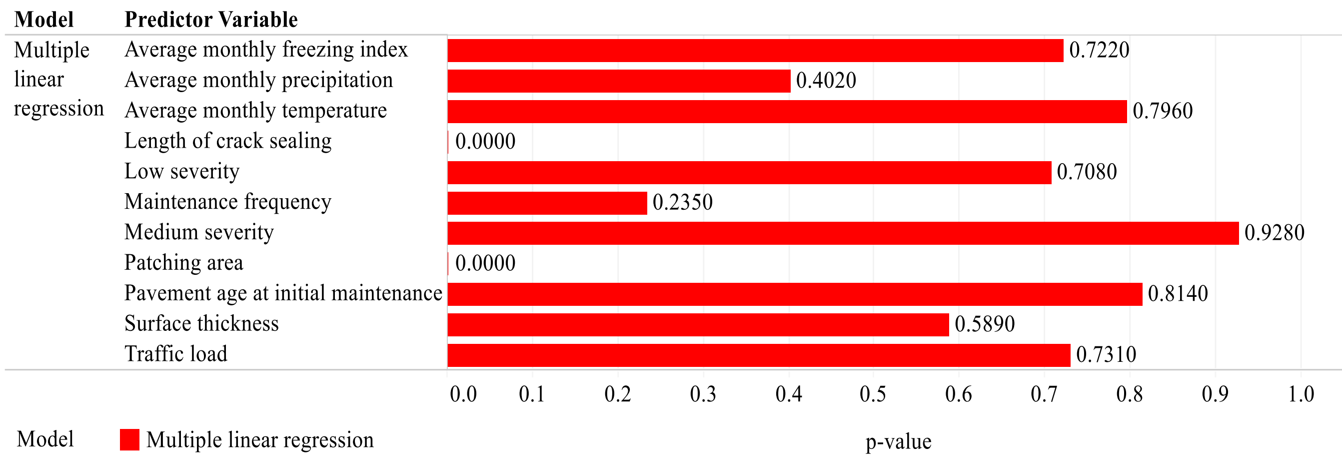
(iii) Importance value of predictor variables in SVR and decision tree regression model

Figure 6-2: Visualization of importance value of models' predictor variables for granular layer in material production and initial construction of pavement.

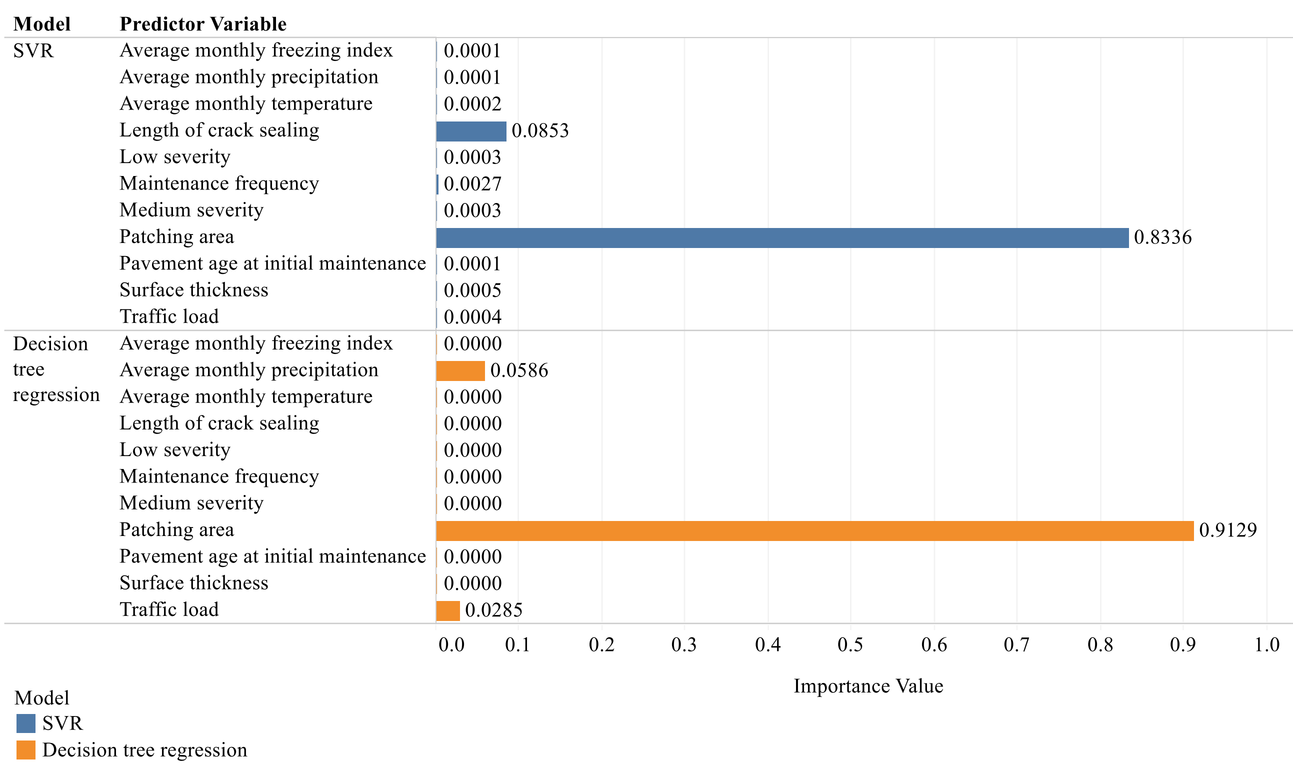
From the SVR model hyperplane, it is found the weight of thickness (0.962) has a large contribution to emission. The square of weight for thickness (0.925) becomes the highest and as a result, the thickness of the asphalt layer has the largest relevance with the emission. From decision tree regression, thickness has a large impact on emission according to importance feature values (0.999).

6.1.2 Maintenance Phase

The patching area and length of crack sealing (p-value less than 0.05) are significant. Unlike multiple linear models, SVR and decision tree regression model describe that the patching area has a much higher significance than the total length of crack sealing. The decision tree regression model shows that precipitation and traffic load has a larger impact on CO_2 emission next to patching area.



(i) *p*-value of predictor variables in multiple linear regression model

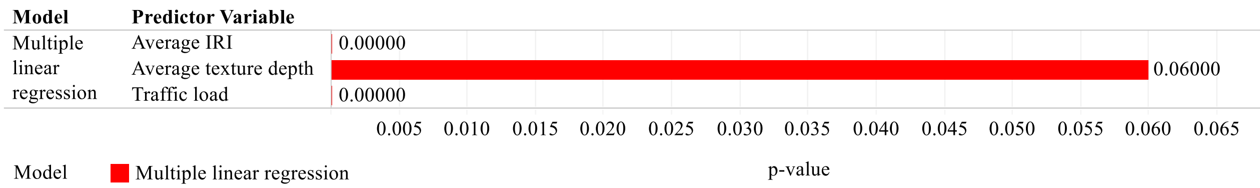


(ii) Importance value of predictor variables in SVR and decision tree regression model

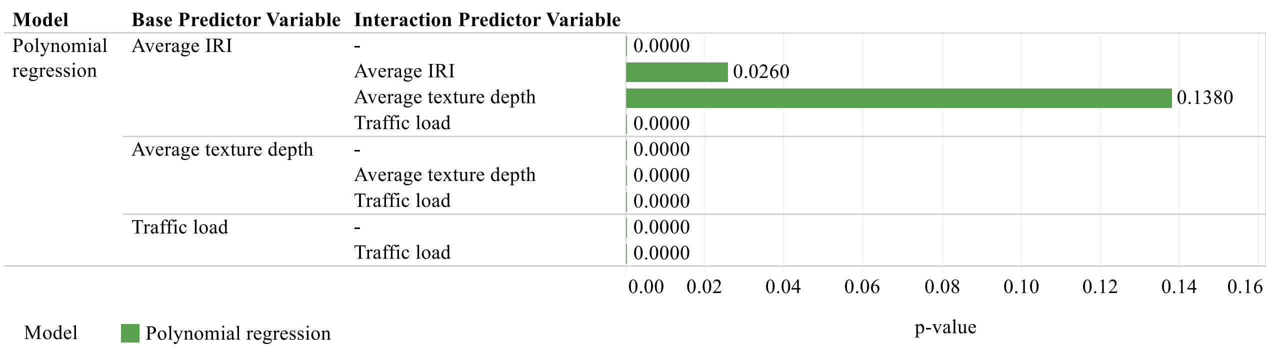
Figure 6-3: Visualization of significance of models' predictor variables in pavement maintenance phase

6.1.3 Use Phase

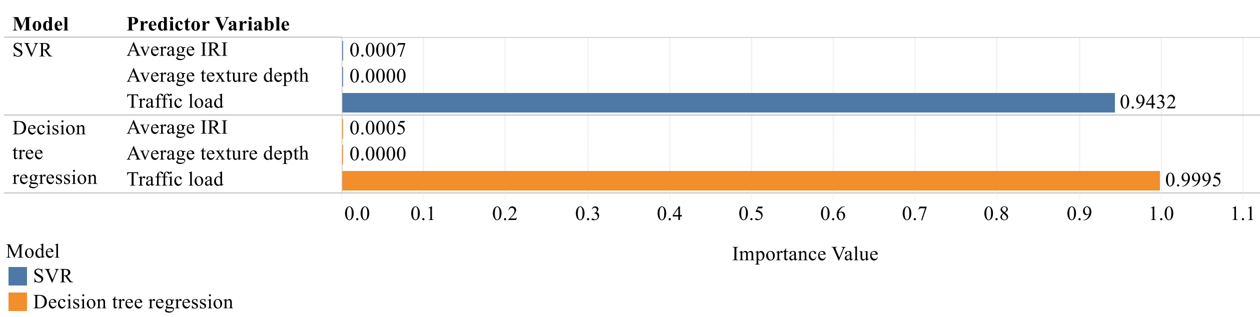
According to Figure 6-4(i) for multiple linear regression, IRI value and traffic load are significant (p-value less than 0.05). Except for IRI-texture depth, base predictor variable and rest of the interactions are significant according to the polynomial regression model.



(i) *p*-value of predictor variables in multiple linear regression model



(ii) *p*-value of predictor variables in polynomial regression model



(iii) Importance value of predictor variables in SVR and decision tree regression model

Figure 6-4: Visualization of significance of models' predictor variables in pavement use phase

6.2 Estimation of Model Parameters

6.2.1 Material Production and Initial Construction Phase

From the coefficient values in Table 6-1, the thickness of asphalt layer, percentage of aggregate, binder and filler have a direct relationship for CO_2 emission. Among these, thickness has the highest impact on increasing CO_2 emission. For one inch increase of asphalt layer, CO_2 emission increases by 4336000 gram. For one unit increase of asphalt binder percentage, CO_2 emission increases by 2453000 gram. On the other hand, the specific gravity of construction material shows an inverse relation with emission. In other words, if the density of aggregate, filler and binder increases the emission will decrease. When the specific gravity of coarse aggregate, binder and filler material is increased by one, CO_2 decreases by 52240000 (maximum reduction), 17410000 and 45280000 gram, respectively.

From the coefficient values of the multiple linear regression model, thickness of asphalt layer and coarse aggregate percentage have a direct relationship for CO_2 emission. For one inch increase of granular layer, CO_2 emission increases by 510200 gram. This emission rate is 8.5 times less than that of the asphalt layer. For one unit increase of coarse percentage, CO_2 emission increases by 734.941 gram. A similar relation is found in polynomial regression. In addition, very high interaction value of aggregate percentage (both for coarse and fine aggregate) with thickness value has been found in this model. The specific gravity of aggregate shows inverse relation with emission in multiple linear regression model. In other words, if the density of aggregate increases the emission will decrease. When the specific gravity of coarse aggregate is increased by one, CO_2 decreases by 27.383 gram. In polynomial regression, the coefficient values for specific gravity and its interaction terms are close to zero.

Table 6-1: Parameter values of LCA models for asphalt layers in material production and initial construction phase

Model Algorithm	Parameter Type	Constant	X1	X2	X3	X4	X5
Multiple linear regression	Coefficient	-17410000	-8661000	-8753000	4336000	-52240000	2453000
			X6	X7	X8	X9	X10
			-17410000	-45280000	3089000	3181000	3914000
SVR	Hyperplane coefficient		X1	X2	X3	X4	X5
			-0.003	0.003	0.898	0	0.05
			X6	X7	X8	X9	X10
			0	0.106	-0.116	-0.067	0
Decision tree regression	Number of leaves		25				

Material Production and Initial Construction (Asphalt Layer) Phase Predictor Variable	
X1	Surface layer
X2	Intermediate layer
X3	Thickness
X4	Aggregate sp. gravity
X5	Binder percentage
X6	Binder sp. gravity
X7	Filler sp. gravity
X8	Coarse aggregate percentage
X9	Fine aggregate percentage
X10	Filler percentage

Table 6-2: Parameter values of different models for granular layers in material production and initial construction phase of pavement

Model Algorithm	Model Parameter	Constant	X1	X2	X3	X4	X5
Multiple linear regression	Coefficient	-9.128	1682.936	-1692.064	510200	-27.383	734.941
			X6				
			-1565.331				
Polynomial Regression	Coefficient	Constant	X1	X2	X3	X4	X5
		0.000	0.014	-0.014	106.334	0.000	0.012
			X6	X7	X8	X9	X10
			-0.002	0.014	0.0000	48.870	0.042
			X11	X12	X13	X14	X15
			0.080	1.254	-0.014	57.464	-0.042
			X16	X17	X18	X19	X20
			-0.068	-1.256	0.014	319.003	5178.288
			X21	X22	X23	X24	X25
			5182.397	0.001	0.037	-0.005	-0.336
			X26	X27			
			1.481	-1.569			
SVR	Hyperplane coefficient		X1	X2	X3	X4	X5
			-0.001	0.001	0.962	0.000	0.001
			X6				
			-0.001				
Decision tree regression	Number of leaves		21				

Material Production and Initial Construction (Granular Layer) Phase Predictor Variables	
X1	Base
X2	Subbase
X3	Thickness
X4	Aggregate sp. gravity
X5	Coarse aggregate percentage
X6	Fine aggregate percentage

6.2.2 Maintenance Phase

For one unit (ft^2) increase of patching area, CO_2 emission increases by 201.661 gram. Moreover, for one unit (ft) increase of coarse percentage, CO_2 emission increases by 144.277 gram. The rest of the predictors show coefficient values almost zero.

Table 6-3: Parameter values of different models for maintenance phase of pavement

Model Algorithm	Parameter Type	Constant	X1	X2	X3	X4	X5
Multiple linear regression	Coefficient	-0.000011	-0.0000023	-0.000009	0.000001	0.0000001	-0.0000013
			X6	X7	X8	X9	X10
			0.0000001	-0.00000263	0.0000002	201.661	144.277
			X11				
			0.00000001				
SVR	Hyperplane coefficient		X1	X2	X3	X4	X5
			0.016	-0.016	-0.023	-0.008	0.014
			X6	X7	X8	X9	X10
			0.009	0.052	0.011	0.913	0.292
			X11				
			-0.021				
Decision tree regression	Number of leaves		6				

Maintenance Phase Predictor Variables	
X1	Medium severity
X2	Low severity
X3	Surface thickness
X4	Average monthly precipitation
X5	Average monthly temperature
X6	Average monthly freezing index
X7	Maintenance frequency
X8	Pavement age at initial maintenance
X9	Patching area
X10	Length of crack sealing
X11	Traffic load

6.2.3 Use Phase

From the coefficient values of the multiple linear regression model, for one m/km increase of mean IRI, CO_2 emission increases by 28790 gram. This emission rate is 3.4 times more than that of one kESAL/year increase of traffic load. Unlike multiple linear regression, traffic load has the highest contribution to CO_2 emission in polynomial regression, followed by the interaction IRI-traffic load. For one unit kESAL increase, CO_2 emission increases by 8208.7 gram and for the IRI-traffic load interaction, this emission value is 198.408 gram. According to SVR and decision tree regression model, traffic load has a large contribution for emission.

Table 6-4: Parameter values of different models for use phase of pavement

Model Algorithm	Parameter Type	Constant	X1	X2	X3		
Multiple linear regression	Coefficient	-42700	28790	670.855	8500.804		
Polynomial Regression	Coefficient	Constant	X1	X2	X3	X4	X5
		0.001	0.001	0.004	8208.698	-0.00005	0.00006
			X6	X7	X8	X9	
			198.408	0.004	0.000	0.000	
SVR	Hyperplane coefficient		X1	X2	X3		
			0.027	0.000	0.971		
Decision tree regression	Number of leaves		20				

Use Phase Predictor Variables	
X1	Average IRI
X2	Average texture depth
X3	Traffic load

6.3 Model Accuracy

In this component of the study, 20% of observations were selected as test data to estimate the performance of the model. Among the models, SVR shows the least root mean square error (RMSE) value which represents the highest prediction accuracy in the asphalt layer of first phase of this study. For the granular layer of the same phase, the polynomial regression model shows the least RMSE value (63.63) which represents the highest prediction accuracy, as shown in Table 6-5.

Among the models, multiple linear regression model shows the least RMSE value (zero) which represents the highest prediction accuracy for maintenance phase, whereas, polynomial and decision tree regression model shows the highest prediction accuracy as shown in Table 6-5.

Table 6-5: Summary of prediction accuracy of models

Pavement Structural Layer	Pavement Life Cycle Phase	RMSE			
		Multiple Linear Regression	Polynomial Regression	Decision Tree Regression	Support Vector Regression
Asphalt layer	Material production and initial construction	1612959.71	-	1754641.16	**877422**
Unbound granular layer		7907.20	**63.63**	355206.70	582648.96
Both layer	Maintenance	**0**	-	143006.82	13444.61
Both layer	Use	3857.41	**0**	**0**	35791.05

6.4 Comparative Analysis for Canadian Provinces

Before developing the machine learning based LCA model, CO_2 emission was estimated for all Canadian LTPP road sections. The emission report is summarized for different phases and provinces in a filled map as shown in Figure 6-5. More than 90% of emission in total pavement LCA goes to use phase from vehicle fuel emission including fuel loss from roughness and texture depth. Fuel loss from roughness and texture carries only 2% of CO_2 emission in total use phase emission. Because of data inadequacy, only Manitoba province was studied for the use phase. Besides use phase, material production and initial construction phase contributes a high quantity of emission of which material production carries the largest quantity as shown in Figure 6-5.

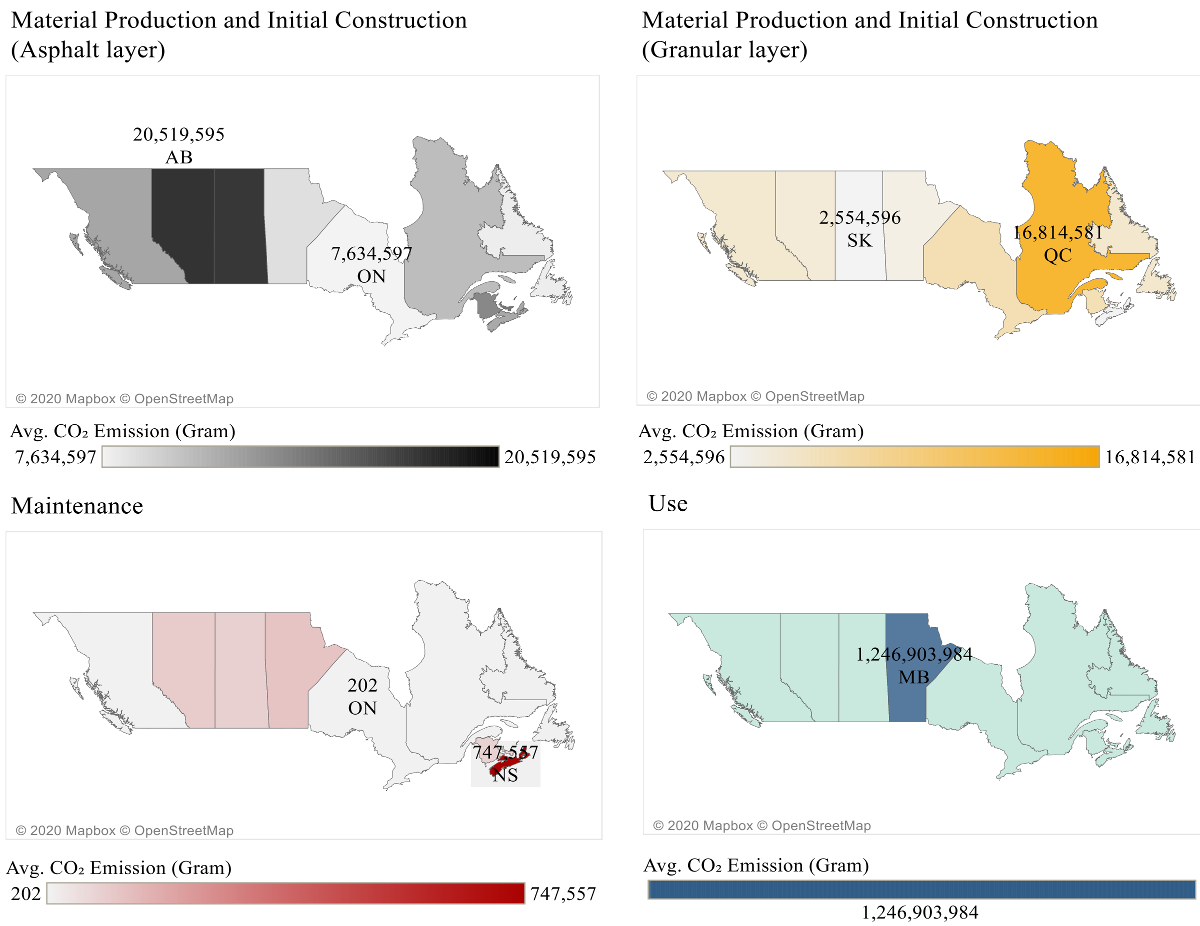


Figure 6-5: Geographical comparison of CO_2 emission (gram) phase-wise contribution from Canadian provinces

Material production and initial construction phase were also studied for asphalt and granular layer separately. Alberta emits the highest average CO_2 in asphalt layer whereas Quebec emits the highest one for granular layer. According to the SVR model (best fit model for material production and initial construction - asphalt layer), design asphalt layer thickness reduction, coarse aggregate percentage reduction and filler density increase can reduce the CO_2 emission in asphalt layer. According to the best fit model for the granular layer (polynomial regression model), both aggregate percentage and granular layer thickness combined reduction can reduce Quebec's maximum emission.

Ontario and Quebec both provinces emit less CO_2 emission for the maintenance phase, though they emit a high quantity of CO_2 in material production and initial construction. Proper pavement and material design and initial construction may reduce the requirement of frequent maintenance. However, sustainable material production and construction method is needed to be introduced to reduce the emission as much as possible.

Alberta, Manitoba and Saskatchewan consume a lot of fuel from material production to maintenance phase. Nova Scotia emits the maximum for the maintenance phase. The necessity to patching area and length of crack sealing reduction can reduce the emission in this province according to best fit multiple linear regression model. Actions need to be undertaken to reduce the generation of a higher amount of pavement distress area.

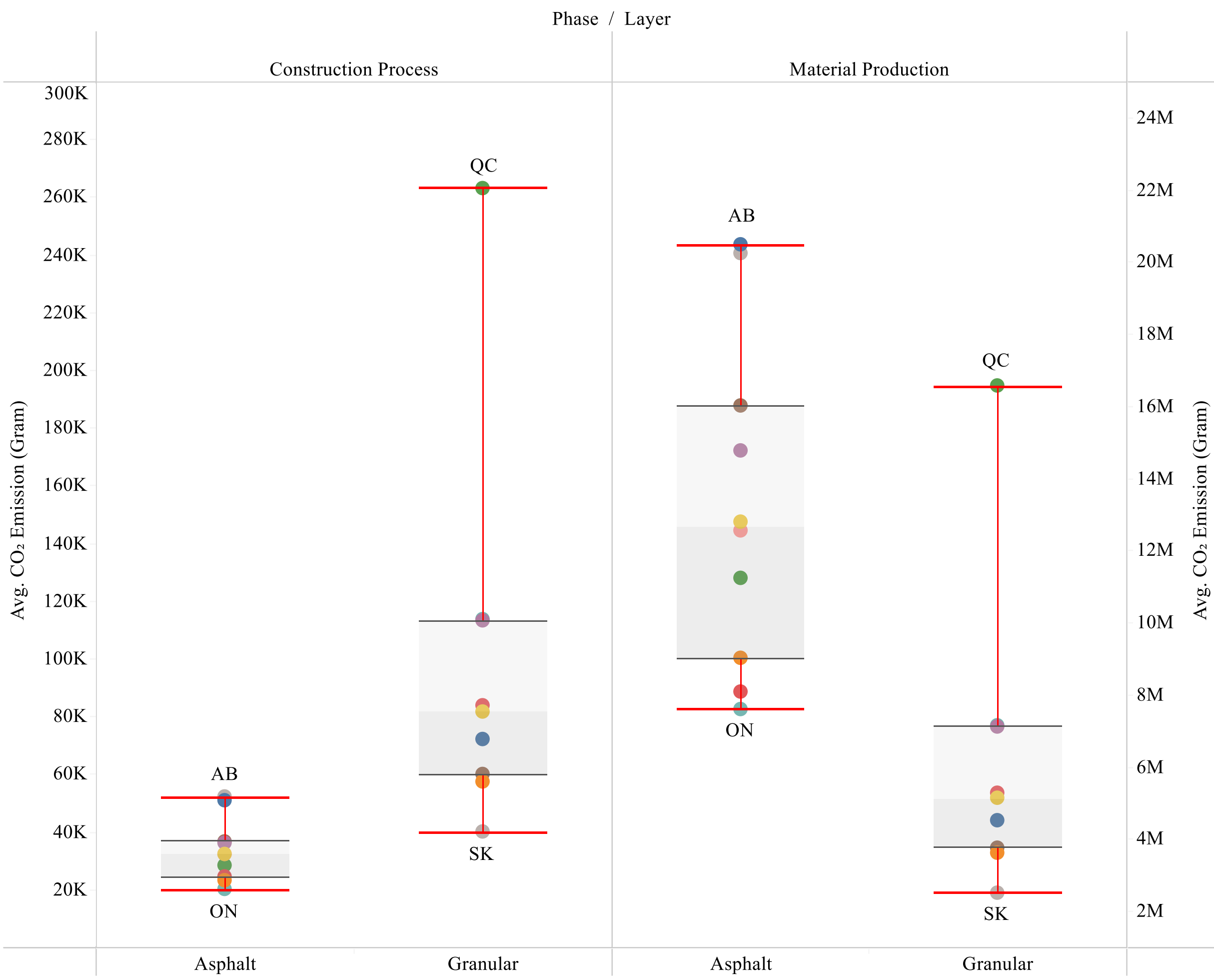


Figure 6-6: Comparison of asphalt and granular layer emission for different Canadian provinces in material production and initial construction phase

CHAPTER 7

CONCLUSIONS

7.1 Major Findings from Environmental Impact Assessment for M&R

- Among six compound-measured environmental impact categories, the global warming potential category, measured in emissions of CO_2-eq. in kg, held the highest values for all four M&R techniques including asphalt patching, rout and sealing, HIR and CIR.
- Based on GWP, the CIR technique produced the lowest percentage of CO_2-eq. (83.87%), and for asphalt patching, the CO_2 emission resulted in the highest percentage (92.22%) which is the least suitable option for M&R methods in light of GWP.
- CIR method which requires less machinery with no heating machinery leads to less diesel fuel for operation, and therefore, causes less reduction of CO_2 emissions.
- In terms of smog potential, asphalt patching (7.28%) appears as the most promising approach for pavement maintenance.
- Rout and sealing and CIR had the most significant impact on the percentage of acidification potential (0.96% each), whereas the contribution of HH particulate and eutrophication potential was minimal for each M&R technique.

7.2 Major Findings from Environmental Impact Assessment for PVI

- This component of the study adopted a statistical approach employing dendrogram and silhouette plotting techniques for clustering. Then, the GWP emission as a result of pavement vehicle interaction from pavement roughness and deflection was estimated for each cluster member. The main factors contributing to the GWP emission were also determined for each cluster.

- The combined impact of various climate factors including precipitation, temperature, and freezing index on PVI was estimated. IRI trend was significantly varying within and between the clusters. Overall, based on GWP emission from the IRI perspective, the clusters can be ranked as follows Cluster 2 > Cluster 1 > Cluster 3 > Cluster 4.

- In Cluster 2, the climate parameters included high annual precipitation, high annual freezing index, and medium annual temperature. It means that there was a high probability that the air voids in the soil layers in the pavement structure was filled and saturated with water because of high annual precipitation, which resulted in frost heave in winter. In the spring season, the frost melted and the trapped water weakened the base layer. This deterioration was accelerated by the dynamic loading of vehicles and resulted in increase of IRI, significantly.
- For light vehicles, the clusters can be ranked as follows: Cluster 2 > Cluster 1 > Cluster 3 > Cluster 4, the same hierarchy of IRI increase rate.
- For the heavy vehicles, GWP value follows: Cluster 2 > Cluster 1 > Cluster 4 > Cluster 3. For the same number of heavy vehicles, the GWP value was much higher. This result indicates that a relatively high impact from heavy vehicle traffic because of PVI for pavement roughness when compared with light vehicle traffic.
- For deflection based PVI effects, Cluster 4 had the maximum vehicle load both for HV and LV, and the minimum subgrade stiffness. These factors combined emitted the highest GWP in Cluster 4 among all the clusters
- As for Cluster 3, though it had the lowest number of vehicle load for both types, it had the minimum surface layer elastic modulus and asphalt layer thickness, which made Cluster 3 the second highest group in terms of GWP emission.
- For the same number of vehicles (1000 AADT), heavy vehicles are dominant rather than LV for GWP emission, considering both cases, roughness and deflection based PVI.
- In new pavement design analysis, attention can be given to increase subgrade stiffness (through soil stabilization) in Cluster 4 to reduce GWP emission due to deflection based PVI.
- For Cluster 3, the elastic modulus of the asphalt layer and pavement design thickness can be enhanced to reduce further emission.

7.3 Major Findings from Life Cycle Assessment Models

- Multiple linear regression, polynomial regression, SVR and decision tree regression algorithm were implemented to find out the best fit model for each LCA phase. SVR is the best fitted for material production and initial construction phase in asphalt layer. Thickness has the highest impact on increasing CO_2 emission. For one inch increase of asphalt layer, CO_2 emission increases by 4336000 gram. It was also found that if the density of aggregate, filler and binder increases the emission will decrease.
- For the granular layer in the same phase, the polynomial regression model is selected as the best one. A strong significant interaction of aggregate percentage (both for coarse and fine aggregate) with thickness value for emission has been found in this model.
- Patching area and length of crack sealing were significant factors for the maintenance phase. According to multiple linear regression (best fit) model, for one unit (ft^2) increase of patching area CO_2 emission increases by 201.661 gram. Moreover, for one unit (ft) increase of coarse percentage, CO_2 emission increases by 144.277 gram.
- Traffic load has the highest contribution to CO_2 emission according to polynomial regression, followed by the interaction IRI-traffic load. For one unit (kESAL) increase in traffic load, CO_2 emission increases by 8208.7 gram and for the IRI-traffic load interaction the emission becomes 198.408 gram.
- More than 90% of emission in total pavement LCA goes to use phase considering both vehicle fuel usage and extra fuel needed due to roughness and texture depth. The extra fuel is responsible for only 2% of CO_2 emission only in the use phase. Besides the use phase, material production and initial construction phase contributes a high quantity of emission of which material production accounts for the larger emission.
- From province-wise it is found that Ontario and Quebec both provinces emits less CO_2 emission for the maintenance phase, though they emit a high quantity of CO_2 in material production and initial construction. Moreover, provinces can take effective measures to reduce their emission.

7.4 Major Contributions

Based on our first and small-scale study of M&Rs in St. John's, it can be concluded that the LCA approach works effectively to comprehend the environmental impact of major maintenance and rehabilitation techniques of asphalt pavement. Furthermore, environmentally friendly road treatment was selected through the quantitative analysis of comparison among those M&R techniques. This methodology can be implemented to understand the environmental impact for rest of the M&Rs.

Canada is the second-largest country in the world and it has very large provinces and territories. Therefore, within a province, there are geometric regions that have completely different climates. As a result, in our second component of the study, a new systematic "climate-based clustering" approach is introduced rather than considering geometric boundaries for environmental impact analysis from the road system.

LCA of a particular pavement section needs lots of inventory data and lengthy calculation time, even using any LCA software. When there is a need to produce LCA results for different alternatives, it becomes more complex and time-consuming. As a result, in order to resolve this issue, a model is developed in the third component of the study. As the proposed LCA model can predict the emission for pavement projects and alternatives within a short time, this advantage allows the decision-makers to think better and eventually make the right decision.

7.5 Limitations and Recommendations for future research

7.5.1 LCA study of M&R

In the LCA study of M&R, the structural failure was not addressed. The reclamation depth for HIR and CIR also was kept constant at a 4inch depth to circumvent the complexity issue. The structural design for maintenance and rehabilitation was based upon the distress condition of the pavement infrastructure. Therefore, in future research, the proper remedy for structural failure and different reclamation depths can be considered.

The properties of asphalt mixture and its component materials are different depending on the provinces and their design guidelines. Based on the available data, an in-depth sensitivity analysis should be performed to find out the material properties which are significant for each type of environmental emission in future research.

The technology associated with the M&R used diesel fuel as an energy source. Athena LCA tool considered construction equipment from the default system. Instead of using default technology, the equipment and their specifications can be updated based on actual equipments in future LCA study.

The procedure of asphalt pavement M&R alternatives is emerging day by day. New technology is introduced for better performance and a more sustainable solution. Environmental emission study for different new M&R can be performed based on emerging technology in future research.

Pavement infrastructure susceptibility includes three essential elements: environmental protection, economic prosperity and social acceptability (Reza et al. 2014). Therefore, further LCA study should be multi-attributed which will evaluate the cost-effectiveness and performance of M&R along with environmental impact (Giustozzi et al. 2012; Yu et al. 2013). Notably, life-cycle cost analysis (LCCA) could be recommended to integrate with LCA researches. Considering both LCA and LCCA, a priority-based pavement management tool can be developed for decision-makers of the pavement management system.

In the case study-based LCA of pavement M&R, only four conventional techniques were considered. In future research, the consideration of all categories of pavement M&R techniques can be considered. An optimization study based on LCA and LCCA to find out the best preventive maintenance, minor and major rehabilitation might be another interesting future study.

7.5.2 LCA study for PVI

In the second component of the study, the annual climate data of only 22 road sections from LTPP had been used for clustering purposes. It can be recommended in a future study to consider monthly climate data, which could be more effective for clustering and understanding the PVI effect.

In this study, two built-in models (HDM IV and PVI Gen II) were used to find out emission due to PVI. For the Canadian climate condition, the calibration of these models can be done for future works. The calibration factor for the Canadian climate will be effective for increasing the accuracy of the research output.

The severity of climate factors are classified intuitively based on historical survey data. Probabilistic methodology for the estimation of climate severity threshold values may increase the acceptability of the severity classification and eventually increase the effectiveness of the study.

In our study, fixed linear relation of the increased rate of IRI was considered over the time period during LCA input. A new model can be developed in which the IRI increase rate from prior data can be added each year for a better understanding of PVI due to roughness. The probabilistic approach can be implemented to comprehend the pattern of IRI increase rate over the time.

The study reveals the effect of traffic load, material properties of subgrade and asphalt layer and design thickness of the asphalt layer based on the data analytics. However, the real mechanism of this significant factor needs to be discovered both for light vehicle and heavy vehicle.

7.5.3 LCA Models

A set of the machine learning model had been developed using calculated CO_2 emission as a response variable. Instead of using calculated CO_2 emission, field data using CO_2 meter can be used. As the proposed model is a prototype for a machine learning-based model, a new and calibrated model using field emission data could be a better research.

The proposed models from this thesis can be further developed by tuning model hyper-parameter (model hyper-parameter is a configuration that is external to the model and whose value cannot be estimated from data). Besides, k fold cross-validation can be used instead of a simple train-test split. K fold cross-validation will mitigate the overfitting of the model and therefore will increase the performance of the model.

Transportation distance from the plant to the site was not considered because of data inadequacy. This factor has a significant effect (concluded from the first component of the study) on the model in material production and the initial construction phase. In collaboration with the department of transportation of each province, the necessary data can be achieved for future research.

The LTPP data includes the arterial road section for the Canadian region. In city areas, collector roads have massive traffic and therefore road usage is higher. Besides, pavement maintenance and rehabilitation are more frequent in this type of roads in urban areas. Therefore, a new study considering collector roads can be performed over several years in future research.

The emission result from LCA models can be added as a geographical information system (GIS) over the road networks. An application of an automated GIS-based LCA model could be another research project for smart infrastructure management.

This LCA models have the potential to interact with the life cycle cost analysis (LCCA) models. Optimization of LCA and LCCA models could be extensively helpful for pavement management system decision-makers. Optimization research of LCA and LCCA can be another scope for future research.

REFERENCES

Ahammed, A. M., Sullivan, S., Finlayson, G., Goemans, C., & Meil, J. (2016). Life Cycle Environmental Assessment Using Athena Pavement LCA Tool: A Manitoba Case Study. *2016 Conference of the Transportation Association of Canada*, (613). Retrieved from http://www.tac-atc.ca/sites/default/files/conf_papers/ahammed_0.pdf

Akbarian, M., Louhghalam, A., & Ulm, F. J. (2014). Mapping pavement - vehicle interaction life cycle impacts on virginia's interstate system. In *LCA Symposium* (pp. 14–16).

Akbarian, M., Moeini-Ardakani, S. S., Ulm, F.-J., & Nazzal, M. (2012). Mechanistic approach to pavement–vehicle interaction and its impact on life-cycle assessment. *Transportation Research Record*, *2306*(1), 171–179.

Alam, M. R., Hossain, K., & Bazan, C. (2020). A Systematic Approach to Estimate Global Warming Potential from Pavement Vehicle Interaction Using Canadian Long-Term Pavement Performance Data. *Journal of Cleaner Production*. Retrieved from https://doi.org/10.1016/j.jclepro.2020.123106

Alam, M. R., Hossain, K., Butt, A. A., Caudle, T., Bazan, C., Caudle, T., & Bazan, C. (2019). Life Cycle Assessment of Asphalt Pavement Maintenance and Rehabilitation Techniques : A Study for the City of St . John ' s. *Canadian Journal of Civil Engineering*, 1–28.

Amos, D. (2006). *Pavement Smoothness and Fuel Efficiency: An Analysis of the Economic Dimensions of the Missouri Smooth Road Initiative.* Missouri. Dept. of Transportation.

Azhar, A., Toller, S., & Birgisson, B. (2015). Life cycle assessment for the green procurement of roads : a way forward. *Journal of Cleaner Production*, *90*, 163–170. https://doi.org/10.1016/j.jclepro.2014.11.068

Baker, J. W., & Lepech, M. D. (2009). Treatment of uncertainties in life cycle assessment. In *Intl. Congress on Structral Safety and Reliability*.

Benbow, E., Iaquinta, J., Lodge, R., & Wright, A. (2007). An Investigation of the effects of pavement stiffness on fuel consumption. *Transport Research Laboratory Limited: Berkshire, UK*.

Bester, C. J. (1984). Effect of pavement type and condition on the fuel consumption of vehicles. *Transportation Research Record*, *1000*, 28–32.

Butt, A. A., Mirzadeh, I., Toller, S., & Birgisson, B. (2014). Life cycle assessment framework for asphalt pavements: methods to calculate and allocate energy of binder and additives. *International Journal of Pavement Engineering*, *15*(4), 290–302. https://doi.org/10.1080/10298436.2012.718348

Cao, R., Leng, Z., Yu, H., & Hsu, S. (2019). Comparative life cycle assessment of warm mix technologies in asphalt rubber pavements with uncertainty analysis. *Resources, Conservation & Recycling*, *147*(December 2018), 137–144. https://doi.org/10.1016/j.resconrec.2019.04.031

Cenek, P. D. (1994). Rolling resistance characteristics of new zealand road surfaces. In *Vehicle-Road Interaction*. ASTM International.

Chatti, K., & Zaabar, I. (2012). Estimating the effects of pavement condition on vehicle operating costs. In *Transportation Research Board*.

Delanne, Y. (1994). The influence of pavement evenness and macrotexture on fuel consumption. In *Vehicle-road interaction*. ASTM International.

Department of Transportation and Works, H. D. D. (2011). *The Department of Transportation and Works Specification*. Retrieved from http://www.tw.gov.nl.ca/publications/specbook2011.pdf

Espinoza, M., Campos, N., Yang, R., Ozer, H., Aguiar-Moya, J. P., Baldi, A., … Al-Qadi, I. L. (2019). Carbon Footprint Estimation in Road Construction : La Abudancia-Florencia Case Study. *Sustainability*, 1–13.

Evangelista, L., & De Brito, J. (2007). Environmental life cycle assessment of concrete made with fine recycled concrete aggregates. *Portugal Sb07-Sustainable Construction, Materials and Practices: Challenge of the Industry for the New Millennium, Pts*, *1*, 789–794.

Farina, A., Zanetti, M. C., Santagata, E., & Blengini, G. A. (2017). Life cycle assessment applied to bituminous mixtures containing recycled materials: Crumb rubber and reclaimed asphalt pavement. *Resources, Conservation and Recycling*, *117*, 204–212.

https://doi.org/https://doi.org/10.1016/j.resconrec.2016.10.015

FHWA. (2018). *Long-Term Pavement Performance Program Brochure*.

Garber, N. J., & Hoel, L. A. (2009). *Traffic and Highway Engineering* (4th ed.).

Giustozzi, F., Crispino, M., & Flintsch, G. (2012). Multi-attribute life cycle assessment of preventive maintenance treatments on road pavements for achieving environmental sustainability. *The International Journal of Life Cycle Assessment*, *17*(4), 409–419. https://doi.org/10.1007/s11367-011-0375-6

Gregory, J. (2017). *Pavement-related research at the MIT Concrete Sustainability Hub*.

Häkkinen, T., & Mäkelä, K. (1996). *Environmental adaption of concrete*. *Asphalt*. Technical Research Centre of Finland.

Harvey, J. T., Kendall, A., Lee, I., Santero, N., Dam, T. Van, & Wang, T. (2010). Pavement Life Cycle Assessment Workshop. *International Journal of Life Cycle Assessment*, *16*(May), 944–946.

Horvath, A. (2004). *A Life-Cycle Analysis Model and Decision-Support Tool for Selecting Recycled Versus Virgin Materials for Highway Applications*.

Horvath, A., & Hendrickson, C. (1998). Comparison of Environmental Implications of Asphalt and Steel-Reinforced Concrete Pavements. *Transportation Research Record: Journal of the Transportation Research Board*, *1626*(September 1998), 105–113. https://doi.org/10.3141/1626-13

Huang, Y., Bird, R., & Heidrich, O. (2009). Development of a life cycle assessment tool for construction and maintenance of asphalt pavements. *Journal of Cleaner Production*, *17*(2), 283–296. https://doi.org/10.1016/j.jclepro.2008.06.005

Hultqvist, A., Christensen, M., Johansson, B., Richter, M., Nygren, J., Hult, J., & Aldén, M. (2002). The HCCI combustion process in a single cycle—high-speed fuel tracer LIF and chemiluminescence imaging. *SAE Transactions*, 913–927.

International Organization for Standardization. (2006). *Environmental management — Life cycle assessment — Requirements and guidelines* (First edit, Vol. 2006). International Standard

ISO 14044.

Kallas, B. F., Puzinauskas, V. P., & Krieger, H. C. (1962). Mineral Fillers in Asphalt Paving Mixtures. *Highway Research Board Bulletin*.

Kaloush, K. E., Carlson, J. D., Golden, J. S., & Phelan, P. E. (2008). *The thermal and radiative characteristics of concrete pavements in mitigating urban heat island effects*.

Kim, B., Lee, H., Park, H., & Kim, H. (2013). Framework for Estimating Greenhouse Gas Emissions Due to Asphalt Pavement Construction. *Journal of Construction Engineering and Management*, *138*(11), 1312–1321. https://doi.org/10.1061/(ASCE)CO.1943-7862.0000549

Kleynhans, T., Montanaro, M., Gerace, A., & Kanan, C. (2017). Predicting top-of-atmosphere thermal radiance using merra-2 atmospheric data with deep learning. *Remote Sensing*, *9*(11), 1133.

Lavin, P. (2003). *Asphalt Pavements : A Practical Guide to Design , Production and Maintenance for Engineers and Architects* (1st ed.). Taylor & Francis.

Lenngren, C. A., & Faldner, L. (2011). Fuel Cost Consideration Regarding Truck Rolling Resistance on Different Pavement Types. *Swedish Road and Transport Research Institute, VTI*.

Loijos, A. A. N. (2011). Life cycle assessment of concrete pavements: impacts and opportunities. Massachusetts Institute of Technology.

Mack, J. W., Akbarian, M., Ulm, F. J., & Louhghalam, A. (2018). Overview of Pavement-Vehicle Interaction Related Research at the MIT Concrete Sustainability Hub. In *Proceedings of the 13th International Symposium on Concrete Roads 2018*.

Meil, J. (2006). A Life Cycle Perspective on Concrete and Asphalt Roadways: Embodied Primary Energy and Global Warming Potential. *Athena Research Institute*.

Ministry of Transportation Ontario. (2017). *Provincial Standard Specification Construction Specification for Hot Mix Asphalt*.

Mroueh, U., Ruotoistenmäki, A., Eskola, P., Laine-Ylijoki, J., Wellman, K., Mäkelä, E., &

Juvankoski, M. (2000). Life cycle assessment of road construction. *Finnish National Road Administration*, (January 2000). Retrieved from http://www.tielaitos.fi/libr.htm

National Research Council (US). Transportation Research Board. Committee for the National Tire Efficiency Study. (2006). *Tires and passenger vehicle fuel economy: informing consumers, improving performance* (Vol. 286). Transportation Research Board.

Noshadravan, A., Wildnauer, M., Gregory, J., & Kirchain, R. (2013). Comparative pavement life cycle assessment with parameter uncertainty. *Transportation Research Part D: Transport and Environment, 25*, 131–138.

Pachauri, R. K., Allen, M. R., Barros, V. R., Broome, J., Cramer, W., Christ, R., … Dasgupta, P. (2014). *Climate change 2014: synthesis report. Contribution of Working Groups I, II and III to the fifth assessment report of the Intergovernmental Panel on Climate Change*. Ipcc.

Paulo, J., Araújo, C., Oliveira, J. R. M., & Silva, H. M. R. D. (2014). The importance of the use phase on the LCA of environmentally friendly solutions for asphalt road pavements. *Transportation Research Part D, 32*, 97–110. https://doi.org/10.1016/j.trd.2014.07.006

Plessis, H. W., Visser, A. T., & Curtayne, P. C. (1990). Fuel consumption of vehicles as affected by road-surface characteristics. In *Surface Characteristics of Roadways: International Research and Technologies* (1990th ed.). ASTM International.

Reza, B., Sadiq, R., & Hewage, K. (2014). Emergy-based life cycle assessment (Em-LCA) for sustainability appraisal of infrastructure systems : a case study on paved roads. *Clean Technologies and Environmental Policy*, *16*(2), 251–266. https://doi.org/10.1007/s10098-013-0615-5

Samieadel, A., Schimmel, K., & Fini, E. H. (2018). Comparative life cycle assessment (LCA) of bio-modified binder and conventional asphalt binder. *Clean Technologies and Environmental Policy*, *20*(1), 191–200. https://doi.org/10.1007/s10098-017-1467-1

Sandberg, U. S. I. (1990). Road macro-and megatexture influence on fuel consumption. In *Surface characteristics of roadways: International Research and Technologies*. ASTM International, 1990.

Santero, N., Masanet, E., & Horvath, A. (2010). Life Cycle Assessment of Pavements: A Critical

Review of Existing Literature and Research. *Lawrence Berkeley National Laboratory*, (April), 81. https://doi.org/10.1016/j.resconrec.2011.03.010

Santos, Joao, Bressi, S., Cerezo, V., Presti, D. Lo, & Dauvergne, M. (2018). Life cycle assessment of low temperature asphalt mixtures for road pavement surfaces: A comparative analysis. *Resources, Conservation and Recycling*, *138*, 283–297. https://doi.org/https://doi.org/10.1016/j.resconrec.2018.07.012

Santos, João, Ferreira, A., & Flintsch, G. (2015). A life cycle assessment model for pavement management : methodology and computational framework. *International Journal of Pavement Engineering*, *16*(3), 268–286.

Schlegel, T., Puiatti, D., Ritter, H.-J., Lesueur, D., Denayer, C., & Shtiza, A. (2016). The limits of partial life cycle assessment studies in road construction practices: A case study on the use of hydrated lime in Hot Mix Asphalt. *Transportation Research Part D: Transport and Environment*, *48*, 141–160. https://doi.org/https://doi.org/10.1016/j.trd.2016.08.005

Soliman, A. M. A. (2006). *Effect of road roughness on the vehicle ride comfort and rolling resistance*. SAE Technical Paper.

Soliman, H., Shalaby, A., & Kass, S. (2014). Overview of Specifications for Unbound Granular Base Materials in Selected Canadian Provinces and Neighbouring States. In *Annual Conference of the Transportation Association of Canada* (pp. 1–15). Montreal, Quebec.

Standard, I. (1997). Environmental mangement- Life cycle assessment - Principles and framework, *1997*.

Stripple, H. (2001). Life Cycle Assessment of Road: A Pilot Study for Inventory Analysis. *IVL RAPPORT*, (1210).

Sumitsawan, P., Ardenkani, S. A., & Romanoschi, S. (2009). Effect of pavement type on fuel consumption and emissions. In *Proceedings of the 2009 Mid-Continent Transportation Research Symposium*.

Taylor, G. W., Farrel, P., & Woodside, A. (2002). Additional analysis of the effect of pavement structures on truck fuel consumption. *National Research Council of Canada (NRC), Ottawa, Ontario*.

U.S. Department of Transportation. (2009). *Technical Support Document for the Lime Manufacturing Sector : Proposed Rule for Mandatory Reporting of Greenhouse Gases*.

Velinsky, S. A., & White, R. A. (1979). *Increased Vehicle Energy Dissipation Due to Changes in Road Roughness with Emphasis on Rolling Losses*. SAE Technical Paper.

Vidal, R., Moliner, E., Martínez, G., & Rubio, M. C. (2013). Life cycle assessment of hot mix asphalt and zeolite-based warm mix asphalt with reclaimed asphalt pavement. *Resources, Conservation and Recycling*, *74*, 101–114.

Yu, B., & Lu, Q. (2012). Life cycle assessment of pavement: Methodology and case study. *Transportation Research Part D: Transport and Environment*, *17*(5), 380–388.

Yu, B., Lu, Q., & Xu, J. (2013). An improved pavement maintenance optimization methodology : Integrating LCA and LCCA. *Transportation Research Part A*, *55*, 1–11.

Yu, B., Qiang, L., & Gu, X. (2016). Data quality and uncertainty assessment methodology for pavement LCA. *International Journal of Pavement Engineering*, (May), 1–7. https://doi.org/10.1080/10298436.2016.1176166

Zaabar, I., & Chatti, K. (2010). Calibration of HDM-4 models for estimating the effect of pavement roughness on fuel consumption for US conditions. *Transportation Research Record*, *2155*(1), 105–116.

Zaniewski, J. P., & Butler, B. C. (1985). Vehicle operating costs related to operating mode, road design, and pavement condition. In *Measuring road roughness and its effects on user cost and comfort*. ASTM Internationa.

Zapata, P., & Gambatese, J. A. (2005). Energy Consumption of Asphalt and Reinforced Concrete Pavement Materials and Construction. *Journal of Infrastructure Systems*, *11*(1), 9–20. https://doi.org/10.1061/(ASCE)1076-0342(2005)11:1(9)

Zhang, H., Keoleian, G. A., & Lepech, M. D. (2008). An integrated life cycle assessment and life cycle analysis model for pavement overlay systems. In *1st Int. Symp. on Life-Cycle Civil Engineering* (pp. 907–912). London: Taylor & Francis Group.

Zimmerman, K. A. (1997). *Guidelines for Using Economic Factors and Maintenance Costs in*

Life-Cycle Cost Analysis.

Ziyadi, M., & Al-qadi, I. L. (2019). Model uncertainty analysis using data analytics for life-cycle assessment (LCA) applications. *International Journal of Life Cycle Assessment*, 945–959.

Ziyadi, M., Ozer, H., & Al-Qadi, I. L. (2017). Functional Unit Choice for Comparative Pavement LCA Involving Use-Stage with Pavement Roughness Uncertainty Quantification (UQ). In *Proceedings of the Symposium on Life-Cycle Assessment of Pavements* (pp. 12–13).

Appendix A: R Coding

A.1 Hierarchical clustering

#read CSV file from working directory

```
roadsection<-read.csv(file="clustering by environment.csv",header=TRUE, sep=",")
```

#normalization

```
z1<- roadsection[,-c(1,1)]
m<-apply(z1,2,mean) #2 denotes column, 1 denotes row (not considered here)
s<-apply(z1,2,sd)
z2<-scale(z1,m,s)
```

#calculating Euclidian distance

```
distance<-dist(z2)
print(distance,digits=3)
```

#hierarchical clustering by cluster dendrogram by complete linkage

```
hc.c<-hclust(distance)
plot(hc.c, labels=roadsection$consid,hang=-1)
```

#cluster membership

```
member.c<-cutree(hc.c,3)
member.a<-cutree(hc.a,3)
table(member.c,member.a)
```

#cluster Means

```
aggregate(z2,list(member.c),mean) #z2=standardized value
aggregate(roadsection[,-c(1,1)],list(member.c),mean)#original values
```

#silhouette plot

```
library(cluster)
plot(silhouette(cutree(hc.c,4),distance))
```

Appendix B: PVI Input data for LCA

Province	SHRP	AC (mm)	Base (mm)	Subbase (mm)	Mean AADT LV	Mean AADT HV	Analysis Period (years)	Pavement Lift	Subgrade Stiffness, k	AC Modulus, E
AB	0501	165.1	73.66	294.64	4138	1477	16	1	129.33	3066.62
AB	2812	152.4	165.1	-	2154	385	16	1	155.7	9815
AB	0903	134.62	566.42	2565.4	11749	1709	6	2	469.07	5776.78
SK	0901	121.92	182.88	233.68	4096	1056	10	2	182.8	7784.74
SK	6410	116.84	132.08	106.68	4991	816	13	2	146.48	4337.08
NL	1801	81.28	251	284	10051	943	16	2	562.09	11299.6
NL	1803	81.28	157.5	381	2266	517	16	1	511.17	9039.68
ON	0960	129.5	228.6	50.8	7526	1225	16	2	191.95	4695.5
ON	2812	241.3	127	-	11790	3325	9	3	236.05	16480
QC	1021	287.02	386.08	1905	15169	2480	14	2	223.51	4233.91
QC	1127	188	416.56	594.36	11175	1523	9	2	254.93	4090.6
NB	6804	154.94	81.28	937.26	4406	1392	7	3	506.2	3348.82
NB	1684	127	83.82	543.56	9895	1479	10	2	149.31	2515.37
NB	1802	276.86	63.5	472.44	4824	681	13	2	223.17	2235.25
AB	A901	119.38	350.52	-	5200	990	14	1	112.86	3132
BC	1005	124.46	238.76	309.88	3867	532	12	2	224.06	7131
BC	6006	134.62	208.28	604.52	19653	1664	9	2	100.53	6782.52
BC	6007	149.86	314.96	-	9311	2356	13	2	110.65	6745.31

Appendix C: SQL Coding

C.1 Check the functional class of Canadian LTPP road sections

```
SELECT [STATE_CODE_EXP]
      ,[SHRP_ID]
      ,[FUNCTIONAL_CLASS]
FROM [Bucket_30922].[dbo].[PROJECT_ID_EXP]
```

C.2 Retrieve lane width and section length and find missing values as well

```
SELECT [STATE_CODE_EXP]
      ,[SHRP_ID]
      ,[MONITORED_LANE]
      ,ISNULL([LANE_WIDTH],0) AS Study_Lane_Width/*ft*/
      ,ISNULL([SECTION_LENGTH],0) AS Study_Section_Length/*ft*/
  FROM [Bucket_30922].[dbo].[SECTION_GENERAL_EXP]
```

C.3 Retrieve initial and final construction year with thickness of particular layer we considered in this study

```
SELECT [STATE_CODE_EXP]
      ,[SHRP_ID]
      ,datepart(yyyy,[START_DATE]) AS Initial_Year
      ,datepart(yyyy,[END_DATE]) AS Final_Year
      ,[DESCRIPTION] AS Layer_Code_No
      ,[REPR_THICKNESS]
      ,[MATL_CODE]
              INTO Layer_Thickness_Table/*Create new table based on query results*/
  FROM [Bucket_30922].[dbo].[TRF_ESAL_AC_THICK]
  where [DESCRIPTION] like '3'
                    OR [DESCRIPTION] LIKE '4'
                    OR [DESCRIPTION] LIKE '5'
```

```
            OR [DESCRIPTION] LIKE '6'
            OR [DESCRIPTION] LIKE '7'
ORDER BY SHRP_ID, CAST([DESCRIPTION] AS Varchar(1000));
```

C.4 Retrieve material properties

```
SELECT [STATE_CODE_EXP]
   ,[SHRP_ID]
   ,[LAYER_NO] AS Layer_Code_No
   ,ISNULL([BINDER_SPEC_GRAV],0) AS BINDER_SPEC_GRAV
   ,ISNULL([BINDER_PCT],0) AS BINDER_PCT
   ,ISNULL([AGG_COARSE_SPEC_GRAV],0) AS AGG_COARSE_SPEC_GRAV
   ,ISNULL([AGG_COARSE_PCT],0) AS AGG_COARSE_PCT
   ,ISNULL([AGG_FINE_SPEC_GRAV],0) AS AGG_FINE_SPEC_GRAV
   ,ISNULL([AGG_FINE_PCT],0) AS AGG_FINE_PCT
   ,ISNULL([AGG_FILLER_SPEC_GRAV],0) AS AGG_FILLER_SPEC_GRAV
   ,ISNULL([AGG_FILLER_PCT],0) AS AGG_FILLER_PCT
FROM [Bucket_30922].[dbo].[TST_SP02]
GROUP BY
[STATE_CODE_EXP],[SHRP_ID],[LAYER_NO],[BINDER_SPEC_GRAV],[BINDER_PCT]
    ,[AGG_COARSE_SPEC_GRAV],[AGG_COARSE_PCT],[AGG_FINE_SPEC_GRAV]
      ,[AGG_FINE_PCT],[AGG_FILLER_SPEC_GRAV],[AGG_FILLER_PCT]
ORDER BY [STATE_CODE_EXP],[SHRP_ID],[LAYER_NO]
```

C.5 Merge material properties and thickness table

```
SELECT *
FROM [LCA_thesis].[dbo].[Ini_Final_Year_Layer_Code_Thickness]
LEFT JOIN [LCA_thesis].[dbo].[Asphalt_Agg_Filler_Percent_Spgravity_Info_1]
ON [LCA_thesis].[dbo].[Ini_Final_Year_Layer_Code_Thickness].STATE_CODE_EXP
=[LCA_thesis].[dbo].[Asphalt_Agg_Filler_Percent_Spgravity_Info_1].STATE_CODE_EXP

AND [LCA_thesis].[dbo].[Ini_Final_Year_Layer_Code_Thickness].SHRP_ID
```

```
    =[LCA_thesis].[dbo].[Asphalt_Agg_Filler_Percent_Spgravity_Info_1].SHRP_ID

AND [LCA_thesis].[dbo].[Ini_Final_Year_Layer_Code_Thickness].Layer_Code_No
=[LCA_thesis].[dbo].[Asphalt_Agg_Filler_Percent_Spgravity_Info_1].Layer_Code_No

        UPDATE [LCA_thesis].[dbo].[Asphalt_Agg_Filler_Percent_Spgravity_Info_1]
    SET
[LCA_thesis].[dbo].[Asphalt_Agg_Filler_Percent_Spgravity_Info_1].[Layer_Code_No]='Origin
al Surface Layer'
        WHERE
[LCA_thesis].[dbo].[Asphalt_Agg_Filler_Percent_Spgravity_Info_1].[Layer_Code_No]='3'
        UPDATE [LCA_thesis].[dbo].[Asphalt_Agg_Filler_Percent_Spgravity_Info_1]
    SET
[LCA_thesis].[dbo].[Asphalt_Agg_Filler_Percent_Spgravity_Info_1].[Layer_Code_No]='AC
layer below surface (binder course)'
        WHERE
[LCA_thesis].[dbo].[Asphalt_Agg_Filler_Percent_Spgravity_Info_1].[Layer_Code_No]='4'
        UPDATE [LCA_thesis].[dbo].[Asphalt_Agg_Filler_Percent_Spgravity_Info_1]
    SET
[LCA_thesis].[dbo].[Asphalt_Agg_Filler_Percent_Spgravity_Info_1].[Layer_Code_No]='Base'
        WHERE
[LCA_thesis].[dbo].[Asphalt_Agg_Filler_Percent_Spgravity_Info_1].[Layer_Code_No]='5'
        UPDATE [LCA_thesis].[dbo].[Asphalt_Agg_Filler_Percent_Spgravity_Info_1]
    SET
[LCA_thesis].[dbo].[Asphalt_Agg_Filler_Percent_Spgravity_Info_1].[Layer_Code_No]='Subba
se'
        WHERE
[LCA_thesis].[dbo].[Asphalt_Agg_Filler_Percent_Spgravity_Info_1].[Layer_Code_No]='6'
        UPDATE [LCA_thesis].[dbo].[Asphalt_Agg_Filler_Percent_Spgravity_Info_1]
```

SET
[LCA_thesis].[dbo].[Asphalt_Agg_Filler_Percent_Spgravity_Info_1].[Layer_Code_No]='Subgrade'
WHERE
[LCA_thesis].[dbo].[Asphalt_Agg_Filler_Percent_Spgravity_Info_1].[Layer_Code_No]='7'
UPDATE [LCA_thesis].[dbo].[Ini_Final_Year_Layer_Code_Thickness]
SET
[LCA_thesis].[dbo].[Ini_Final_Year_Layer_Code_Thickness].[Layer_Code_No]='Original Surface Layer'
WHERE
[LCA_thesis].[dbo].[Ini_Final_Year_Layer_Code_Thickness].[Layer_Code_No]='3'
UPDATE [LCA_thesis].[dbo].[Ini_Final_Year_Layer_Code_Thickness]
SET
[LCA_thesis].[dbo].[Ini_Final_Year_Layer_Code_Thickness].[Layer_Code_No]='AC layer below surface (binder course)'
WHERE
[LCA_thesis].[dbo].[Ini_Final_Year_Layer_Code_Thickness].[Layer_Code_No]='4'
UPDATE [LCA_thesis].[dbo].[Ini_Final_Year_Layer_Code_Thickness]
SET
[LCA_thesis].[dbo].[Ini_Final_Year_Layer_Code_Thickness].[Layer_Code_No]='Base'
WHERE
[LCA_thesis].[dbo].[Ini_Final_Year_Layer_Code_Thickness].[Layer_Code_No]='5'
UPDATE [LCA_thesis].[dbo].[Ini_Final_Year_Layer_Code_Thickness]
SET
[LCA_thesis].[dbo].[Ini_Final_Year_Layer_Code_Thickness].[Layer_Code_No]='Subbase'
WHERE
[LCA_thesis].[dbo].[Ini_Final_Year_Layer_Code_Thickness].[Layer_Code_No]='6'
UPDATE [LCA_thesis].[dbo].[Ini_Final_Year_Layer_Code_Thickness]
SET
[LCA_thesis].[dbo].[Ini_Final_Year_Layer_Code_Thickness].[Layer_Code_No]='Subgrade'

```
WHERE
[LCA_thesis].[dbo].[Ini_Final_Year_Layer_Code_Thickness].[Layer_Code_No]='7'
```

C.6 Missing values are replaced with random value generated within real life range

```
UPDATE [LCA_thesis].[dbo].[Intial_Construction_2]
	SET
[LCA_thesis].[dbo].[Intial_Construction_2].[BINDER_PCT]=FLOOR(RAND(CHECKSUM(NEWID()))*(6-4+1)+4)
	WHERE [LCA_thesis].[dbo].[Intial_Construction_2].[Layer_Code_No] LIKE 'Original Surface Layer'
	AND [LCA_thesis].[dbo].[Intial_Construction_2].[BINDER_PCT] IS NULL;

UPDATE [LCA_thesis].[dbo].[Intial_Construction_2]
SET
[LCA_thesis].[dbo].[Intial_Construction_2].[BINDER_PCT]=FLOOR(RAND(CHECKSUM(NEWID()))*(6-4+1)+4)
	WHERE [LCA_thesis].[dbo].[Intial_Construction_2].[Layer_Code_No] LIKE 'Original Surface Layer'
	AND [LCA_thesis].[dbo].[Intial_Construction_2].[BINDER_PCT] = 0;

UPDATE [LCA_thesis].[dbo].[Intial_Construction_2]
	SET
[LCA_thesis].[dbo].[Intial_Construction_2].[BINDER_PCT]=FLOOR(RAND(CHECKSUM(NEWID()))*(6-4+1)+4)
	WHERE [LCA_thesis].[dbo].[Intial_Construction_2].[Layer_Code_No] LIKE 'AC layer below surface (binder course)'
	AND [LCA_thesis].[dbo].[Intial_Construction_2].[BINDER_PCT] IS NULL;

UPDATE [LCA_thesis].[dbo].[Intial_Construction_2]
```

```
	SET
[LCA_thesis].[dbo].[Intial_Construction_2].[AGG_COARSE_PCT]=FLOOR(RAND(CHECKSUM(NEWID()))*(55-48+1)+48)
	WHERE [LCA_thesis].[dbo].[Intial_Construction_2].[Layer_Code_No] LIKE 'Original Surface Layer'
	AND [LCA_thesis].[dbo].[Intial_Construction_2].[AGG_COARSE_PCT] IS NULL;

UPDATE [LCA_thesis].[dbo].[Intial_Construction_2]
	SET
[LCA_thesis].[dbo].[Intial_Construction_2].[AGG_COARSE_PCT]=FLOOR(RAND(CHECKSUM(NEWID()))*(55-48+1)+48)
	WHERE [LCA_thesis].[dbo].[Intial_Construction_2].[Layer_Code_No] LIKE 'Original Surface Layer'
	AND [LCA_thesis].[dbo].[Intial_Construction_2].[AGG_COARSE_PCT] LIKE '0';

UPDATE [LCA_thesis].[dbo].[Intial_Construction_2]
	SET
[LCA_thesis].[dbo].[Intial_Construction_2].[AGG_COARSE_PCT]=FLOOR(RAND(CHECKSUM(NEWID()))*(55-48+1)+48)
	WHERE [LCA_thesis].[dbo].[Intial_Construction_2].[Layer_Code_No] LIKE 'AC layer below surface (binder course)'
	AND [LCA_thesis].[dbo].[Intial_Construction_2].[AGG_COARSE_PCT] IS NULL;

UPDATE [LCA_thesis].[dbo].[Intial_Construction_2]
	SET
[LCA_thesis].[dbo].[Intial_Construction_2].[AGG_FINE_PCT]=FLOOR(RAND(CHECKSUM(NEWID()))*(40-35+1)+35)
	WHERE  [LCA_thesis].[dbo].[Intial_Construction_2].[Layer_Code_No] LIKE 'Original Surface Layer'
	AND [LCA_thesis].[dbo].[Intial_Construction_2].[AGG_FINE_PCT] IS NULL;
```

```
UPDATE [LCA_thesis].[dbo].[Intial_Construction_2]
	SET
[LCA_thesis].[dbo].[Intial_Construction_2].[AGG_FINE_PCT]=FLOOR(RAND(CHECKSUM(
NEWID()))*(40-35+1)+35)
	WHERE  [LCA_thesis].[dbo].[Intial_Construction_2].[Layer_Code_No] LIKE 'Original
Surface Layer'
	AND [LCA_thesis].[dbo].[Intial_Construction_2].[AGG_FINE_PCT] LIKE '0';

UPDATE [LCA_thesis].[dbo].[Intial_Construction_2]
	SET
[LCA_thesis].[dbo].[Intial_Construction_2].[AGG_FINE_PCT]=FLOOR(RAND(CHECKSUM(
NEWID()))*(40-35+1)+35)
	WHERE  [LCA_thesis].[dbo].[Intial_Construction_2].[Layer_Code_No] LIKE 'AC layer
below surface (binder course)'
	AND [LCA_thesis].[dbo].[Intial_Construction_2].[AGG_FINE_PCT] IS NULL;

UPDATE [LCA_thesis].[dbo].[Intial_Construction_2]
	SET [LCA_thesis].[dbo].[Intial_Construction_2].[AGG_COARSE_PCT]=61.5
	WHERE  [LCA_thesis].[dbo].[Intial_Construction_2].[Layer_Code_No] LIKE 'Base'
	AND [LCA_thesis].[dbo].[Intial_Construction_2].[AGG_COARSE_PCT] IS NULL
	AND [LCA_thesis].[dbo].[Intial_Construction_2].[STATE_CODE_EXP] LIKE
'Ontario';

UPDATE [LCA_thesis].[dbo].[Intial_Construction_2]
	SET [LCA_thesis].[dbo].[Intial_Construction_2].[AGG_COARSE_PCT]=67.5
	WHERE  [LCA_thesis].[dbo].[Intial_Construction_2].[Layer_Code_No] LIKE 'Base'
	AND [LCA_thesis].[dbo].[Intial_Construction_2].[AGG_COARSE_PCT] IS NULL
	AND [LCA_thesis].[dbo].[Intial_Construction_2].[STATE_CODE_EXP] LIKE 'British
Columbia';

UPDATE [LCA_thesis].[dbo].[Intial_Construction_2]
```

```
	SET [LCA_thesis].[dbo].[Intial_Construction_2].[AGG_COARSE_PCT]=58
	WHERE  [LCA_thesis].[dbo].[Intial_Construction_2].[Layer_Code_No] LIKE 'Base'
	AND [LCA_thesis].[dbo].[Intial_Construction_2].[AGG_COARSE_PCT] IS NULL
	AND [LCA_thesis].[dbo].[Intial_Construction_2].[STATE_CODE_EXP] LIKE
'Alberta';

UPDATE [LCA_thesis].[dbo].[Intial_Construction_2]
	SET [LCA_thesis].[dbo].[Intial_Construction_2].[AGG_COARSE_PCT]=58
	WHERE  [LCA_thesis].[dbo].[Intial_Construction_2].[Layer_Code_No] LIKE 'Base'
	AND [LCA_thesis].[dbo].[Intial_Construction_2].[AGG_COARSE_PCT] LIKE '0
	AND [LCA_thesis].[dbo].[Intial_Construction_2].[STATE_CODE_EXP] LIKE 'Alberta'

UPDATE [LCA_thesis].[dbo].[Intial_Construction_2]
	SET [LCA_thesis].[dbo].[Intial_Construction_2].[AGG_COARSE_PCT]=68.5
	WHERE  [LCA_thesis].[dbo].[Intial_Construction_2].[Layer_Code_No] LIKE 'Base'
	AND [LCA_thesis].[dbo].[Intial_Construction_2].[AGG_COARSE_PCT] IS NULL
	AND [LCA_thesis].[dbo].[Intial_Construction_2].[STATE_CODE_EXP] LIKE
'Saskatchewan';

UPDATE [LCA_thesis].[dbo].[Intial_Construction_2]
	SET [LCA_thesis].[dbo].[Intial_Construction_2].[AGG_COARSE_PCT]=75
	WHERE  [LCA_thesis].[dbo].[Intial_Construction_2].[Layer_Code_No] LIKE 'Base'
	AND [LCA_thesis].[dbo].[Intial_Construction_2].[AGG_COARSE_PCT] IS NULL
	AND [LCA_thesis].[dbo].[Intial_Construction_2].[STATE_CODE_EXP] LIKE
'Manitoba';

UPDATE [LCA_thesis].[dbo].[Intial_Construction_2]
	SET [LCA_thesis].[dbo].[Intial_Construction_2].[AGG_COARSE_PCT]=66.1
	WHERE  [LCA_thesis].[dbo].[Intial_Construction_2].[Layer_Code_No] LIKE 'Base'
	AND [LCA_thesis].[dbo].[Intial_Construction_2].[AGG_COARSE_PCT] IS NULL
```

AND [LCA_thesis].[dbo].[Intial_Construction_2].[STATE_CODE_EXP] LIKE 'New Brunswick';

UPDATE [LCA_thesis].[dbo].[Intial_Construction_2]
SET [LCA_thesis].[dbo].[Intial_Construction_2].[AGG_COARSE_PCT]=66.1
WHERE [LCA_thesis].[dbo].[Intial_Construction_2].[Layer_Code_No] LIKE 'Base'
AND [LCA_thesis].[dbo].[Intial_Construction_2].[AGG_COARSE_PCT] IS NULL
AND [LCA_thesis].[dbo].[Intial_Construction_2].[STATE_CODE_EXP] LIKE 'Newfoundland';

UPDATE [LCA_thesis].[dbo].[Intial_Construction_2]
SET [LCA_thesis].[dbo].[Intial_Construction_2].[AGG_COARSE_PCT]=66.1
WHERE [LCA_thesis].[dbo].[Intial_Construction_2].[Layer_Code_No] LIKE 'Base'
AND [LCA_thesis].[dbo].[Intial_Construction_2].[AGG_COARSE_PCT] IS NULL
AND [LCA_thesis].[dbo].[Intial_Construction_2].[STATE_CODE_EXP] LIKE 'NovaScotia';

UPDATE [LCA_thesis].[dbo].[Intial_Construction_2]
SET [LCA_thesis].[dbo].[Intial_Construction_2].[AGG_COARSE_PCT]=66.1
WHERE [LCA_thesis].[dbo].[Intial_Construction_2].[Layer_Code_No] LIKE 'Base'
AND [LCA_thesis].[dbo].[Intial_Construction_2].[AGG_COARSE_PCT] IS NULL
AND [LCA_thesis].[dbo].[Intial_Construction_2].[STATE_CODE_EXP] LIKE 'Prince Edward Island';

UPDATE [LCA_thesis].[dbo].[Intial_Construction_2]
SET [LCA_thesis].[dbo].[Intial_Construction_2].[AGG_COARSE_PCT]=66.1
WHERE [LCA_thesis].[dbo].[Intial_Construction_2].[Layer_Code_No] LIKE 'Base'
AND [LCA_thesis].[dbo].[Intial_Construction_2].[AGG_COARSE_PCT] IS NULL
AND [LCA_thesis].[dbo].[Intial_Construction_2].[STATE_CODE_EXP] LIKE 'Quebec';

```
UPDATE [LCA_thesis].[dbo].[Intial_Construction_2]
	SET [LCA_thesis].[dbo].[Intial_Construction_2].[AGG_COARSE_PCT]=61.5
	WHERE  [LCA_thesis].[dbo].[Intial_Construction_2].[Layer_Code_No] LIKE 'Subbase'
	AND [LCA_thesis].[dbo].[Intial_Construction_2].[AGG_COARSE_PCT] IS NULL
	AND [LCA_thesis].[dbo].[Intial_Construction_2].[STATE_CODE_EXP] LIKE
'Ontario';

UPDATE [LCA_thesis].[dbo].[Intial_Construction_2]
	SET [LCA_thesis].[dbo].[Intial_Construction_2].[AGG_COARSE_PCT]=67.5
	WHERE  [LCA_thesis].[dbo].[Intial_Construction_2].[Layer_Code_No] LIKE 'Subbase'
	AND [LCA_thesis].[dbo].[Intial_Construction_2].[AGG_COARSE_PCT] IS NULL
	AND [LCA_thesis].[dbo].[Intial_Construction_2].[STATE_CODE_EXP] LIKE 'British
Columbia';

UPDATE [LCA_thesis].[dbo].[Intial_Construction_2]
	SET [LCA_thesis].[dbo].[Intial_Construction_2].[AGG_COARSE_PCT]=58
	WHERE  [LCA_thesis].[dbo].[Intial_Construction_2].[Layer_Code_No] LIKE 'Subbase'
	AND [LCA_thesis].[dbo].[Intial_Construction_2].[AGG_COARSE_PCT] IS NULL
	AND [LCA_thesis].[dbo].[Intial_Construction_2].[STATE_CODE_EXP] LIKE
'Alberta';

UPDATE [LCA_thesis].[dbo].[Intial_Construction_2]
	SET [LCA_thesis].[dbo].[Intial_Construction_2].[AGG_COARSE_PCT]=58
	WHERE  [LCA_thesis].[dbo].[Intial_Construction_2].[Layer_Code_No] LIKE 'Subbase'
	AND [LCA_thesis].[dbo].[Intial_Construction_2].[AGG_COARSE_PCT] LIKE '0'
	AND [LCA_thesis].[dbo].[Intial_Construction_2].[STATE_CODE_EXP] LIKE
'Alberta';

UPDATE [LCA_thesis].[dbo].[Intial_Construction_2]
	SET [LCA_thesis].[dbo].[Intial_Construction_2].[AGG_COARSE_PCT]=68.5
	WHERE  [LCA_thesis].[dbo].[Intial_Construction_2].[Layer_Code_No] LIKE 'Subbase'
```

```
AND [LCA_thesis].[dbo].[Intial_Construction_2].[AGG_COARSE_PCT] IS NULL
AND [LCA_thesis].[dbo].[Intial_Construction_2].[STATE_CODE_EXP] LIKE
'Saskatchewan';

UPDATE [LCA_thesis].[dbo].[Intial_Construction_2]
SET [LCA_thesis].[dbo].[Intial_Construction_2].[AGG_COARSE_PCT]=75
WHERE  [LCA_thesis].[dbo].[Intial_Construction_2].[Layer_Code_No] LIKE 'Subbase'
AND [LCA_thesis].[dbo].[Intial_Construction_2].[AGG_COARSE_PCT] IS NULL
AND [LCA_thesis].[dbo].[Intial_Construction_2].[STATE_CODE_EXP] LIKE
'Manitoba';

UPDATE [LCA_thesis].[dbo].[Intial_Construction_2]
SET [LCA_thesis].[dbo].[Intial_Construction_2].[AGG_COARSE_PCT]=66.1
WHERE  [LCA_thesis].[dbo].[Intial_Construction_2].[Layer_Code_No] LIKE 'Subbase'
AND [LCA_thesis].[dbo].[Intial_Construction_2].[AGG_COARSE_PCT] IS NULL
AND [LCA_thesis].[dbo].[Intial_Construction_2].[STATE_CODE_EXP] LIKE 'New
Brunswick';

UPDATE [LCA_thesis].[dbo].[Intial_Construction_2]
SET [LCA_thesis].[dbo].[Intial_Construction_2].[AGG_COARSE_PCT]=66.1
WHERE  [LCA_thesis].[dbo].[Intial_Construction_2].[Layer_Code_No] LIKE 'Subbase'
AND [LCA_thesis].[dbo].[Intial_Construction_2].[AGG_COARSE_PCT] IS NULL
AND [LCA_thesis].[dbo].[Intial_Construction_2].[STATE_CODE_EXP] LIKE
'Newfoundland';

UPDATE [LCA_thesis].[dbo].[Intial_Construction_2]
SET [LCA_thesis].[dbo].[Intial_Construction_2].[AGG_COARSE_PCT]=66.1
WHERE  [LCA_thesis].[dbo].[Intial_Construction_2].[Layer_Code_No] LIKE 'Subbase'
AND [LCA_thesis].[dbo].[Intial_Construction_2].[AGG_COARSE_PCT] IS NULL
AND [LCA_thesis].[dbo].[Intial_Construction_2].[STATE_CODE_EXP] LIKE
'NovaScotia';
```

```
UPDATE [LCA_thesis].[dbo].[Intial_Construction_2]
	SET [LCA_thesis].[dbo].[Intial_Construction_2].[AGG_COARSE_PCT]=66.1
	WHERE  [LCA_thesis].[dbo].[Intial_Construction_2].[Layer_Code_No] LIKE 'Subbase'
	AND [LCA_thesis].[dbo].[Intial_Construction_2].[AGG_COARSE_PCT] IS NULL
	AND [LCA_thesis].[dbo].[Intial_Construction_2].[STATE_CODE_EXP] LIKE 'Prince
Edward Island';

UPDATE [LCA_thesis].[dbo].[Intial_Construction_2]
	SET [LCA_thesis].[dbo].[Intial_Construction_2].[AGG_COARSE_PCT]=66.1
	WHERE  [LCA_thesis].[dbo].[Intial_Construction_2].[Layer_Code_No] LIKE 'Subbase'
	AND [LCA_thesis].[dbo].[Intial_Construction_2].[AGG_COARSE_PCT] IS NULL
	AND [LCA_thesis].[dbo].[Intial_Construction_2].[STATE_CODE_EXP] LIKE
'Quebec';

UPDATE [LCA_thesis].[dbo].[Intial_Construction_2]
	SET [LCA_thesis].[dbo].[Intial_Construction_2].[BINDER_PCT]=0
	WHERE [LCA_thesis].[dbo].[Intial_Construction_2].[Layer_Code_No] LIKE 'Subbase'
	OR [LCA_thesis].[dbo].[Intial_Construction_2].[Layer_Code_No] LIKE 'Base'


UPDATE [LCA_thesis].[dbo].[Intial_Construction_2]
		SET [LCA_thesis].[dbo].[Intial_Construction_2].[AGG_FILLER_PCT]
		=100-CAST([LCA_thesis].[dbo].[Intial_Construction_2].[AGG_COARSE_PCT]
AS decimal(4,2))
			CAST([LCA_thesis].[dbo].[Intial_Construction_2].[AGG_FINE_PCT]
AS decimal(4,2))


UPDATE [LCA_thesis].[dbo].[Intial_Construction_2]
```

```
SET
[LCA_thesis].[dbo].[Intial_Construction_2].[AGG_COARSE_SPEC_GRAV]=3;

UPDATE [LCA_thesis].[dbo].[Intial_Construction_2]
        SET [LCA_thesis].[dbo].[Intial_Construction_2].[AGG_FINE_SPEC_GRAV]=3;

UPDATE [LCA_thesis].[dbo].[Intial_Construction_2]
        SET
[LCA_thesis].[dbo].[Intial_Construction_2].[AGG_FILLER_SPEC_GRAV]=2.6;

UPDATE [LCA_thesis].[dbo].[Intial_Construction_2]
        SET [LCA_thesis].[dbo].[Intial_Construction_2].[BINDER_SPEC_GRAV]=1;
```

C.7 Remove the column with similar value

```
ALTER TABLE [LCA_thesis].[dbo].[INITIAL CONSTRUCTION_2]
DROP COLUMN
[MATL_CODE],[,[AGG_COARSE_PCT],[AGG_FINE_PCT],[AGG_FILLER_PCT];

 ALTER TABLE [LCA_thesis].[dbo].[INITIAL CONSTRUCTION_2]
 DROP COLUMN [AGG_COARSE_PCT],[AGG_FINE_PCT],[AGG_FILLER_PCT];

SELECT [STATE_CODE_EXP]
   ,[SHRP_ID]
   ,[Initial_Yr]
   ,[Final_Yr]
   ,[Layer_Code_No]
   ,[REPR_THICKNESS]
   ,[AGG_COARSE_SPEC_GRAV]
   ,[AGG_FINE_SPEC_GRAV]
   ,[BINDER_PCT]
   ,[BINDER_SPEC_GRAV]
```

```
      ,[AGG_FILLER_SPEC_GRAV]
      ,[Aggregate_Pct]
      ,[CA_Pct]
      ,[FA_Pct]
      ,[Filler_Pct]
  INTO Initial_Construction_3
  FROM [LCA_thesis].[dbo].[INITIAL CONSTRUCTION_2];

ALTER TABLE [LCA_thesis].[dbo].[Initial_Construction_3]
  ADD Volume decimal(10,2);
```

C.8 Measure volume of each material

```
UPDATE [LCA_thesis].[dbo].[Initial_Construction_3]
SET Volume = 12*499.90* [REPR_THICKNESS]/12;

SELECT [STATE_CODE_EXP]
      ,[SHRP_ID]
      ,[Initial_Yr]
      ,[Final_Yr]
      ,[Layer_Code_No]
      ,[REPR_THICKNESS]
      ,[AGG_COARSE_SPEC_GRAV]
      ,[AGG_FINE_SPEC_GRAV]
      ,[BINDER_PCT]
      ,[BINDER_SPEC_GRAV]
      ,[AGG_FILLER_SPEC_GRAV]
      ,[Aggregate_Pct]
      ,[CA_Pct]
      ,[FA_Pct]
      ,[Filler_Pct]
      ,[Volume]
```

```
    ,([Volume]*.01* (TRY_CAST([CA_Pct] as decimal(5,2)))) as Coarse_Agg_Volume
    ,([Volume]*.01* (TRY_CAST([FA_Pct] as decimal(5,2)))) as Fine_Agg_Volume
    ,([Volume]*.01* (TRY_CAST([Filler_Pct] as decimal(5,2)))) as Filler_Volume
Into Initial_Construction_4
FROM [LCA_thesis].[dbo].[Initial_Construction_3]
```

C.9 Find emission for each material

```
ALTER TABLE [LCA_thesis].[dbo].[Initial_Construction_4]
ADD CO2EMISSION_CA DECIMAL(10,2);

  UPDATE [LCA_thesis].[dbo].[Initial_Construction_4]
  set CO2EMISSION_CA=[Coarse_Agg_Volume]*62.4*5.46*[AGG_COARSE_SPEC_GRAV]

ALTER TABLE [LCA_thesis].[dbo].[Initial_Construction_4]
  ADD CO2EMISSION_FA DECIMAL(10,2);


  UPDATE [LCA_thesis].[dbo].[Initial_Construction_4]
  set CO2EMISSION_FA=[Fine_Agg_Volume]*62.4*5.46*[AGG_FINE_SPEC_GRAV];

  ALTER TABLE [LCA_thesis].[dbo].[Initial_Construction_4]
  ADD CO2EMISSION_Filler DECIMAL(10,2);


  UPDATE [LCA_thesis].[dbo].[Initial_Construction_4]
  set CO2EMISSION_Filler=[Filler_Volume]*62.4*340*[AGG_FILLER_SPEC_GRAV];

 ALTER TABLE [LCA_thesis].[dbo].[Initial_Construction_4]
  ADD Binder_Volume DECIMAL(10,2);

 UPDATE [LCA_thesis].[dbo].[Initial_Construction_4]
```

```
 set Binder_Volume=[Volume]*.01*(TRY_CAST([BINDER_PCT] AS DECIMAL(5,2)));

ALTER TABLE [LCA_thesis].[dbo].[Initial_Construction_4]
 ADD CO2EMISSION_Binder DECIMAL(10,2);

 UPDATE [LCA_thesis].[dbo].[Initial_Construction_4]
 set CO2EMISSION_Binder=[Binder_Volume]*62.4*560.98*[BINDER_SPEC_GRAV];

 ALTER TABLE [LCA_thesis].[dbo].[Initial_Construction_4]
 ADD TOTAL_CO2_EMISSION DECIMAL(10,2);

UPDATE [LCA_thesis].[dbo].[Initial_Construction_4]
Set
TOTAL_CO2_EMISSION=[CO2EMISSION_CA]+[CO2EMISSION_FA]+[CO2EMISSION_FILLER]+[CO2EMISSION_BINDER]
```

C.10 Find emission for construction process

```
SELECT *
Into Initial_Construction_6
 FROM [LCA_thesis].[dbo].[Initial_Construction_5];

 ALTER TABLE [LCA_thesis].[dbo].[Initial_Construction_5]
 ADD CO2_EMISSION_ASPHALT_PAVER decimal(10,2);

 UPDATE [LCA_thesis].[dbo].[Initial_Construction_5]
 SET CO2_EMISSION_ASPHALT_PAVER/*gram*/ = [Volume]*0.046*49.1*852*3.16/2400
 WHERE [Layer_Code_No] LIKE 'Original Surface Layer';

 ALTER TABLE [LCA_thesis].[dbo].[Initial_Construction_5]
 ADD CO2_EMISSION_Pneumatic_Roller decimal(10,2);
```

```
UPDATE [LCA_thesis].[dbo].[Initial_Construction_5]
SET CO2_EMISSION_Pneumatic_Roller/*gram*/ = [Volume]*0.046*26.1*852*3.16/668
WHERE [Layer_Code_No] LIKE 'Original Surface Layer';

ALTER TABLE [LCA_thesis].[dbo].[Initial_Construction_5]
ADD CO2_EMISSION_Tandem_Roller decimal(10,2);

UPDATE [LCA_thesis].[dbo].[Initial_Construction_5]
SET CO2_EMISSION_Tandem_Roller/*gram*/ = [Volume]*0.046*32.7*852*3.16/285
WHERE [Layer_Code_No] LIKE 'Original Surface Layer';

/* below codes for AC layer below surface (binder course)*/
UPDATE [LCA_thesis].[dbo].[Initial_Construction_5]
SET CO2_EMISSION_ASPHALT_PAVER/*gram*/ = [Volume]*0.046*49.1*852*3.16/2400
WHERE [Layer_Code_No] LIKE 'AC layer below surface (binder course)';

UPDATE [LCA_thesis].[dbo].[Initial_Construction_5]
SET CO2_EMISSION_Pneumatic_Roller/*gram*/ = [Volume]*0.046*26.1*852*3.16/668
WHERE [Layer_Code_No] LIKE 'AC layer below surface (binder course)';

UPDATE [LCA_thesis].[dbo].[Initial_Construction_5]
SET CO2_EMISSION_Tandem_Roller/*gram*/ = [Volume]*0.046*32.7*852*3.16/285
WHERE [Layer_Code_No] LIKE 'AC layer below surface (binder course)';

/* below codes for granular unbound layer*/
ALTER TABLE [LCA_thesis].[dbo].[Initial_Construction_5]
ADD CO2_EMISSION_Excavator decimal(10,2);

UPDATE [LCA_thesis].[dbo].[Initial_Construction_5]
SET CO2_EMISSION_Excavator/*gram*/ = [Volume]*0.048*34.2*852*3.16/315
WHERE [Layer_Code_No] LIKE 'Base'
```

```
        OR [Layer_Code_No] LIKE 'Subbase'

  ALTER TABLE [LCA_thesis].[dbo].[Initial_Construction_5]
  ADD CO2_EMISSION_Vibratory_Compactor decimal(10,2);

  UPDATE [LCA_thesis].[dbo].[Initial_Construction_5]
  SET CO2_EMISSION_Vibratory_Compactor/*gram*/ =
[Volume]*0.048*27.6*852*3.16/1832
  WHERE [Layer_Code_No] LIKE 'Base'
        OR [Layer_Code_No] LIKE 'Subbase'
```

C.11 Climate data

```
SELECT ISNULL([TOTAL_MON_PRECIP],0) AS TOTAL_MON_PRECIP
            ISNULL([TOTAL_SNOWFALL_MONTH],0) AS
[TOTAL_SNOWFALL_MONTHs]
     INTO PrecipitationTable1
FROM [Bucket_33198].[dbo].[CLM_VWS_PRECIP_MONTH]

ALTER TABLE [Bucket_33198].[dbo].[PrecipitationTable1]
ADD Total_Month_Precipitation Decimal (10,5);

UPDATE [Bucket_33198].[dbo].[PrecipitationTable1]
SET Total_Month_Precipitation =
[TOTAL_MON_PRECIPITATION]+[TOTAL_SNOWFALL_MONTHS]

SELECT *
        INTO LCA_thesis.dbo.Total_Monthly_Precipitation
  FROM [Bucket_33198].[dbo].[PrecipitationTable1]

SELECT [STATE_CODE_EXP]
                  ,[SHRP_ID]
```

```
                    ,AVG(CAST([Total_Month_Precipitation] AS DECIMAL (10,5))) AS
AVG_Month_Precipitation
                    INTO Total_Monthly_Precipitation2
  FROM [LCA_thesis].[dbo].[Total_Monthly_Precipitation]
  GROUP BY [STATE_CODE_EXP]
                    ,[SHRP_ID]
  ORDER BY [STATE_CODE_EXP]
                    ,[SHRP_ID]

/*Tempereature*/

SELECT [SHRP_ID]
      ,[STATE_CODE_EXP]
      ,AVG (CAST ([MEAN_MON_TEMP_AVG] AS DECIMAL (10,5))) AS
AVG_MON_TEMP
      ,AVG (CAST ([FREEZE_INDEX_MONTH] AS DECIMAL (10,5))) AS
AVG_MON_FREEZINDEX
  INTO  LCA_thesis.dbo.Monthly_Temperture
  FROM [Bucket_33215].[dbo].[CLM_VWS_TEMP_MONTH]

  GROUP BY [SHRP_ID]
      ,[STATE_CODE_EXP]
  ORDER BY [SHRP_ID]
      ,[STATE_CODE_EXP]
```

C.12 Join three column of surface layer thickness, precipitation and temperature

```
SELECT *
FROM [LCA_thesis].[dbo].[Original_Surface_Thickness2]
                    surfthickness inner join [LCA_thesis].[dbo].[Total_Monthly_Precipitation2]
                    preci on surfthickness.SHRP_ID=preci.SHRP_ID
```

```
                    AND
surfthickness.STATE_CODE_EXP=preci.STATE_CODE_EXP
                    inner join [LCA_thesis].[dbo].[Monthly_Temperture]
                    temp on temp.SHRP_ID=preci.SHRP_ID
                         AND
temp.STATE_CODE_EXP=preci.STATE_CODE_EXP
```

C.13 Calculate mean texture and combine tables of IRI and texture depth

```
SELECT  datepart(yyyy,[VISIT_DATE]) AS Year
   ,[STATE_CODE_EXP]
   ,[SHRP_ID]
   ,[MRI]
       INTO Roughness_Table
 FROM [Bucket_33775].[dbo].[MON_HSS_PROFILE_SECTION]

SELECT [STATE_CODE_EXP]
   ,[SHRP_ID]
   ,datepart(yyyy,[VISIT_DATE]) AS Year
   ,[Mean_MTD]
       INTO Texture_Depth_Table
 FROM [Bucket_33775].[dbo].[MON_HSS_TEXTURE_SECTION]

SELECT [STATE_CODE_EXP]
   ,[SHRP_ID]
   ,[YEAR]
   ,[KESAL_YEAR]
       INTO ESAL_Table
 FROM [Bucket_33775].[dbo].[TRF_ESAL_COMPUTED]

SELECT *
FROM [Bucket_33775].[dbo].[ESAL_Table] ET
```

```
INNER JOIN [Bucket_33775].[dbo].[Roughness_Table] RT
ON ET.STATE_CODE_EXP = RT.STATE_CODE_EXP
AND ET.SHRP_ID = RT.SHRP_ID
AND ET.[YEAR] = RT.[YEAR]

INNER JOIN [Bucket_33775].[dbo].[Texture_Depth_Table] TT
ON TT.STATE_CODE_EXP = RT.STATE_CODE_EXP
AND TT.SHRP_ID = RT.SHRP_ID
AND TT.[YEAR] = RT.[YEAR]
```

Appendix D: Python Coding

Material Production and Initial construction

D.1 Multiple linear regression (asphalt layer)

```
#importing libraries
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd

#importing the datatset
dataset=pd.read_csv('Initial_construction_asphalt_layer_final_version.csv')
X = dataset.iloc[:,:-1].values
y = dataset.iloc[:,9].values

#encoding categorical data
from sklearn.preprocessing import LabelEncoder, OneHotEncoder
labelencoder_X = LabelEncoder()
X[:,0]=labelencoder_X.fit_transform(X[:,0]) #different label assigned
onehotencoder = OneHotEncoder(categorical_features = [0])
X = onehotencoder.fit_transform(X).toarray()  #different column for diff. label

#splitting the dataset into the Training set and Test set
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size = 0.2, random_state = 0)

#fitting multiple linear regression with train data
from sklearn.linear_model import LinearRegression
regressor = LinearRegression()
regressor.fit(X_train, y_train)
```

#predict values

y_predict = regressor.predict(X_test)

D.2 Decision tree regression (asphalt layer)

#importing libraries

import numpy as np

import matplotlib.pyplot as plt

import pandas as pd

from IPython.display import Image

from sklearn import tree

import pydotplus

#importing the datatset

dataset=pd.read_csv('Initial_construction_asphalt_layer_final_version.csv')

X = dataset.iloc[:,:-1].values

y = dataset.iloc[:,9].values

#encoding categorical data

from sklearn.preprocessing import LabelEncoder, OneHotEncoder

labelencoder_X = LabelEncoder()

X[:,0]=labelencoder_X.fit_transform(X[:,0]) #different label assigned

onehotencoder = OneHotEncoder(categorical_features = [0])

X = onehotencoder.fit_transform(X).toarray() #different column for diff. label

#splitting the dataset into the Training set and Test set

from sklearn.model_selection import train_test_split

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size = 0.2, random_state = 0)

#fitting decision tree regression to the dataset

from sklearn.tree import DecisionTreeRegressor

regressor = DecisionTreeRegressor(min_samples_leaf=5,random_state = 0) regressor.fit(X_train,y_train)

regressor.feature_importances_
regressor.get_n_leaves()

 # Create DOT data
dot_data = tree.export_graphviz(regressor, out_file=None)

 # Draw graph
graph = pydotplus.graph_from_dot_data(dot_data)

 # Create PNG
graph.write_png("iris.png")

#predict
predictedval=regressor.predict(X_test)

D.3 SVR (asphalt layer)

#importing libraries
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd

#importing the dataset
dataset=pd.read_csv('Initial_construction_asphalt_layer_final_version.csv')
X = dataset.iloc[:,:-1].values
y = dataset.iloc[:,9].values

#encoding categorical data
from sklearn.preprocessing import LabelEncoder, OneHotEncoder
labelencoder_X = LabelEncoder()

```
X[:,0]=labelencoder_X.fit_transform(X[:,0]) #different label assigned
onehotencoder = OneHotEncoder(categorical_features = [0])
X = onehotencoder.fit_transform(X).toarray()  #different column for diff. label
```

#splitting the dataset into the Training set and Test set

```
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size = 0.2, random_state = 0)
```

#feature scaling

```
from sklearn.preprocessing import StandardScaler
sc_X = StandardScaler()
sc_Y = StandardScaler()
X_train_scaled=sc_X.fit_transform(X_train)
y_train_scaled=sc_Y.fit_transform(y_train.reshape(-1,1))
```

#fitting SVR

```
from sklearn import svm
regressor = svm.SVR(kernel = 'linear')
regressor.fit(X_train_scaled,y_train_scaled)
```

#predicting a new result

```
y_pred = sc_Y.inverse_transform(regressor.predict(sc_X.transform(X_test)))
```

#weight vectors

```
regressor.coef_
```

D.4 Multiple linear regression (granular layer)

#importing libraries

```
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
```

#importing the dataset

dataset=pd.read_csv('INIT_CONSTR_6_Granular_layer_prepared.csv')

X = dataset.iloc[:,:-1].values

y = dataset.iloc[:,6].values

#encoding categorical data

from sklearn.preprocessing import LabelEncoder, OneHotEncoder

labelencoder_X = LabelEncoder()

X[:,0]=labelencoder_X.fit_transform(X[:,0]) #different label assigned

onehotencoder = OneHotEncoder(categorical_features = [0])

X = onehotencoder.fit_transform(X).toarray() #different column for diff. label

#splitting the dataset into the Training set and Test set

from sklearn.model_selection import train_test_split

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size = 0.2, random_state = 0)

#fitting multiple linear regression with train data

from sklearn.linear_model import LinearRegression

regressor = LinearRegression()

regressor.fit(X_train, y_train)

#predict values

y_predict = regressor.predict(X_test)

D.5 Polynomial regression (granular layer)

#importing the datatset

dataset=pd.read_csv('INIT_CONSTR_6_Granular_layer_prepared.csv')

X = dataset.iloc[:,:-1].values

y = dataset.iloc[:,6].values

#encoding categorical data

```
from sklearn.preprocessing import LabelEncoder, OneHotEncoder
labelencoder_X = LabelEncoder()
X[:,0]=labelencoder_X.fit_transform(X[:,0]) #different label assigned
onehotencoder = OneHotEncoder(categorical_features = [0])
X = onehotencoder.fit_transform(X).toarray()  #different column for diff. label
```

#splitting the dataset into the Training set and Test set

```
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size = 0.2, random_state = 0)
```

#fitting polynomial regression with train data

```
from sklearn.preprocessing import PolynomialFeatures
poly_reg= PolynomialFeatures(degree=2)
X_poly = poly_reg.fit_transform(X_train)

from sklearn.linear_model import LinearRegression
lin_reg2 = LinearRegression()
lin_reg2.fit(X_poly,y_train)
```

#predict values

```
y_predict = lin_reg2.predict(poly_reg.fit_transform(X_test))
```

D.6 Decision tree regression (granular layer)

#importing libraries

```
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
```

#importing the datatset

```
dataset=pd.read_csv('INIT_CONSTR_6_Granular_layer_prepared.csv')
X = dataset.iloc[:,:-1].values
```

y = dataset.iloc[:,6].values

#encoding categorical data

from sklearn.preprocessing import LabelEncoder, OneHotEncoder

labelencoder_X = LabelEncoder()

X[:,0]=labelencoder_X.fit_transform(X[:,0]) #different label assigned

onehotencoder = OneHotEncoder(categorical_features = [0])

X = onehotencoder.fit_transform(X).toarray() #different column for diff. label

#splitting the dataset into the Training set and Test set

from sklearn.model_selection import train_test_split

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size = 0.2, random_state = 0)

#fitting decision tree regression to the dataset

from sklearn.tree import DecisionTreeRegressor

regressor = DecisionTreeRegressor(random_state = 0)

regressor.fit(X_train,y_train)

#predict

predictedval=regressor.predict(X_test)

D.7 SVR (granular layer)

#importing libraries

import numpy as np

import matplotlib.pyplot as plt

import pandas as pd

#importing the datatset

dataset=pd.read_csv('INIT_CONSTR_6_Granular_layer_prepared.csv')

X = dataset.iloc[:,:-1].values

y = dataset.iloc[:,6].values

#encoding categorical data

```
from sklearn.preprocessing import LabelEncoder, OneHotEncoder
labelencoder_X = LabelEncoder()
X[:,0]=labelencoder_X.fit_transform(X[:,0]) #different label assigned
onehotencoder = OneHotEncoder(categorical_features = [0])
X = onehotencoder.fit_transform(X).toarray()  #different column for diff. label
```

#splitting the dataset into the Training set and Test set

```
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size = 0.2, random_state = 0)
```

#feature scaling

```
from sklearn.preprocessing import StandardScaler
sc_X = StandardScaler()
sc_Y = StandardScaler()
X_train_scaled=sc_X.fit_transform(X_train)
y_train_scaled=sc_Y.fit_transform(y_train.reshape(-1,1))
```

#fitting SVR to the dataset

```
from sklearn import svm
regressor = svm.SVR(kernel = 'linear')
regressor.fit(X_train_scaled,y_train_scaled)
```

#predicting a new result

```
y_pred = sc_Y.inverse_transform(regressor.predict(sc_X.transform(X_test)))
```

#weight vectors

```
regressor.coef_
```

Maintenance Phase

D.8 Multiple linear regression

#importing libraries

```
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
```

#importing the datatset

```
dataset=pd.read_csv('MAINTENANCE_PHASE.csv')
X = dataset.iloc[:,:-1].values
y = dataset.iloc[:,10].values
```

#encoding categorical data

```
from sklearn.preprocessing import LabelEncoder, OneHotEncoder
labelencoder_X = LabelEncoder()
X[:,8]=labelencoder_X.fit_transform(X[:,8]) #different label assigned
onehotencoder = OneHotEncoder(categorical_features = [8])
X = onehotencoder.fit_transform(X).toarray()  #different column for diff. label
```

#splitting the dataset into the Training set and Test set

```
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size = 0.2, random_state = 0)
```

#fitting multiple linear regression with train data

```
from sklearn.linear_model import LinearRegression
regressor = LinearRegression()
regressor.fit(X_train, y_train)
```

#predict values

```
y_predict = regressor.predict(X_test)
```

#OLS

```
import statsmodels.regression.linear_model as sm
X = np.append(arr= np.ones((47,1)).astype(int), values = X, axis =1)
X_opt= X[:,:]
regressor_OLS = sm.OLS(endog= y, exog = X_opt).fit()
regressor_OLS.summary()
```

D.9 Decision tree regression

#importing libraries

```
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
from IPython.display import Image
from sklearn import tree
import pydotplus
```

#importing the datatset

```
dataset=pd.read_csv('MAINTENANCE_PHASE.csv')
X = dataset.iloc[:,:-1].values
y = dataset.iloc[:,10].values
```

#encoding categorical data

```
from sklearn.preprocessing import LabelEncoder, OneHotEncoder
labelencoder_X = LabelEncoder()
X[:,8]=labelencoder_X.fit_transform(X[:,8]) #different label assigned
onehotencoder = OneHotEncoder(categorical_features = [8])
X = onehotencoder.fit_transform(X).toarray()  #different column for diff. label
```

#splittting the dataset into the Training set and Test set

```
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size = 0.2, random_state = 0)
```

#fitting decisiontree regression to the dataset

```
from sklearn.tree import DecisionTreeRegressor
regressor = DecisionTreeRegressor(min_samples_leaf=5,
                                  random_state = 0)
regressor.fit(X_train,y_train)

regressor.feature_importances_
regressor.get_n_leaves()
```

#create DOT data

```
dot_data = tree.export_graphviz(regressor, out_file=None)
```

#draw graph

```
graph = pydotplus.graph_from_dot_data(dot_data)
```

#show graph

```
Image(graph.create_png())
```

#create PDF

```
graph.write_pdf("MAINTENACEPHASE.pdf")
```

#create PNG

```
graph.write_png("MAINTENACEPHASE.png")
```

#predict

```
predictedval=regressor.predict(X_test)
```

D.10 SVR

#importing libraries

```
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
```

#importing the datatset

```
dataset=pd.read_csv('MAINTENANCE_PHASE.csv')
X = dataset.iloc[:,:-1].values
y = dataset.iloc[:,10].values
```

#encoding categorical data

```
from sklearn.preprocessing import LabelEncoder, OneHotEncoder
labelencoder_X = LabelEncoder()
X[:,8]=labelencoder_X.fit_transform(X[:,8]) #different label assigned
onehotencoder = OneHotEncoder(categorical_features = [8])
X = onehotencoder.fit_transform(X).toarray()  #different column for diff. label
```

#splittting the dataset into the Training set and Test set

```
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size = 0.2, random_state = 0)
```

#feature scaling

```
from sklearn.preprocessing import StandardScaler
sc_X = StandardScaler()
sc_Y = StandardScaler()
X_train_scaled=sc_X.fit_transform(X_train)
y_train_scaled=sc_Y.fit_transform(y_train.reshape(-1,1))
```

#fitting SVR ro the dataset

```
from sklearn import svm
regressor = svm.SVR(kernel = 'linear')
regressor.fit(X_train_scaled,y_train_scaled)
```

#predicting a new result

```
y_pred = sc_Y.inverse_transform(regressor.predict(sc_X.transform(X_test)))
```

Use Phase

D.11 Multiple linear regression

```
#importing libraries
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd

#Importing the datatset
dataset=pd.read_csv('USE_PHASE_READY_FOR_MODEL.csv')
X = dataset.iloc[:,:-1].values
y = dataset.iloc[:,3].values

#splitting the dataset into the Training set and Test set
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size = 0.2, random_state = 0)

#fitting multiple linear regression with train data
from sklearn.linear_model import LinearRegression
regressor = LinearRegression()
regressor.fit(X_train, y_train)

#predict values
y_predict = regressor.predict(X_test)
```

D.12 Polynomial regression

```
#importing libraries
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
```

#Importing the datatset

```
dataset=pd.read_csv('USE_PHASE.csv')
X = dataset.iloc[:,:-1].values
y = dataset.iloc[:,3].values
```

#splitting the dataset into the Training set and Test set

```
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size = 0.2, random_state = 0)
```

#fitting polynomial regression with train data

```
from sklearn.preprocessing import PolynomialFeatures
poly_reg= PolynomialFeatures(degree=2)
X_poly = poly_reg.fit_transform(X_train)

from sklearn.linear_model import LinearRegression
lin_reg2 = LinearRegression()
lin_reg2.fit(X_poly,y_train)
```

#predict values

```
y_predict = lin_reg2.predict(poly_reg.fit_transform(X_test))
```

#OLS

```
import statsmodels.regression.linear_model as sm
X_opt= X_poly[:,:]
regressor_OLS = sm.OLS(endog= y_train, exog = X_opt).fit()
regressor_OLS.summary()
```

D.13 Decision tree regression

#importing libraries

```
import numpy as np
import matplotlib.pyplot as plt
```

```
import pandas as pd
from IPython.display import Image
from sklearn import tree
import pydotplus

#importing the datatset
dataset=pd.read_csv('USE_PHASE.csv')
X = dataset.iloc[:,:-1].values
y = dataset.iloc[:,3].values

#splitting the dataset into the Training set and Test set
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size = 0.2, random_state = 0)

#fitting decision tree regression to the dataset
from sklearn.tree import DecisionTreeRegressor
regressor = DecisionTreeRegressor(min_samples_leaf=5, random_state = 0)
regressor.fit(X_train,y_train)

regressor.feature_importances_
regressor.get_n_leaves()

# Create DOT data
dot_data = tree.export_graphviz(regressor, out_file=None)

# Draw graph
graph = pydotplus.graph_from_dot_data(dot_data)

# Show graph
Image(graph.create_png())
```

 # Create PDF

```
graph.write_pdf("MAINTENACEPHASE.pdf")
```

 # Create PNG

```
graph.write_png("MAINTENACEPHASE.png")
```

#predict

```
predictedval=regressor.predict(X_test)
```

D.14 SVR

#importing libraries

```
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
```

#importing the datatset

```
dataset=pd.read_csv('USE_PHASE.csv')
X = dataset.iloc[:,:-1].values
y = dataset.iloc[:,3].values
```

#splitting the dataset into the Training set and Test set

```
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size = 0.2, random_state = 0)
```

#feature scaling

```
from sklearn.preprocessing import StandardScaler
sc_X = StandardScaler()
sc_Y = StandardScaler()
X_train_scaled=sc_X.fit_transform(X_train)
y_train_scaled=sc_Y.fit_transform(y_train.reshape(-1,1))
```

#fitting SVR to the dataset

```
from sklearn import svm
regressor = svm.SVR(kernel = 'linear')
regressor.fit(X_train_scaled,y_train_scaled)
```

#predicting a new result

```
y_pred = sc_Y.inverse_transform(regressor.predict(sc_X.transform(X_test)))
```

#weight vectors

```
regressor.coef_
```