Master Thesis

Service-life Pavement Performance Simulation: Capturing the Influence of Traffic Flows by Big Data Analysis

Zili Wang

CIE5060-09 MSc Thesis
Faculty of Civil Engineering and Geosciences
February 15, 2020

Master Thesis

Service-life Pavement Performance Simulation: Capturing the Influence of Traffic Flows by Big Data Analysis

by

Zili Wang

as the assignment of CIE5060-09 MSc Thesis

at the Delft University of Technology,

in cooperation with Rijkswaterstaat.

Student number:	4698320	
Project duration:	April 21, 2019 – January 31, 2020	
Graduation committee:	Dr. Kumar Anupam	Daily supervisor, TU Delft
	Prof. dr. ir. JWC (Hans) van Lint	Chair of board, TU Delft
	Dr.ir. H. Farah	TU Delft
Company mentors:	Frank Bouman	Rijkswaterstaat
	Thijs Bennis	Rijkswaterstaat

An electronic version of this thesis is available at `http://repository.tudelft.nl/`.

Preface

Welcome to my master thesis! Here I will present the outcomes of my research for the traffic influence on pavement performance, and based on the knowledge, a simulation software to forecast whether the performance indicators meet the maintenance requirements timely will be developed. Before presenting the findings from the research I would like to share the journey I have gone through in the last few months.

When I was looking for a graduation project, Building Information Modeling (BIM) was my first interest. After a preliminary investigation, I found out the real challenge of BIM application was not a technology problem, but about the low willingness of the engineers to learn the new software. Commonly, the selection of technology is inertial. Only if the new technology saves much time and costs, people will have the preference to use it. Thus, it seems the solution to improve the BIM application is making a more efficient software. But there is a dilemma. The more advanced the computer program is, the bigger gap between the old and new systems there is, and the more learning difficulties the technicians have. Based on the surveys and my consideration of self professional backgrounds, the topic was abandoned.

Inspired by the core idea of BIM, I believe the data aggregation platform can avoid the needless reduplicative work of the later phase in a system development life cycle, and it is beneficial for the stakeholders to make the optimal decisions with the data transparency. Accordingly, I was coming up with the other question: Can the information integration of the design, the construction and the operation achieve a better maintenance? Because I studied structural engineering and mechanics in 4-year bachelor, structural design and construction are the familiar fields to me. With the 2-year master education of transport planning in Delft University of Technology (TU Delft), I believe that I am capable to do the case.

The first and most important step is to find the data. I contacted Prof. Hans van Lint, who is the chair of Delft Integrated Traffic and Travel Laboratory (DiTTlab). After hearing my topic, he was very willing to support and permitted me to access all the data DiTTlab had. Not only that, but Hans also introduced me to Prof. Sandra Erkens, the chair of pavement engineering department in TU Delft. Step by step, the idea landed on the multidisciplinary research, which is relevant to pavement performance, traffic data, and maintenance management.

After the determination of which direction my thesis would go to, I showed the pilot proposal to my daily supervisor, Dr. Kumar Anupam. Via him, the idea was presented to Rijkswaterstaat, the executive agency of Ministry of Infrastructure and Water Management in the Netherlands (RWS) and gained their data support of the annual measurement of roughness, rutting, and skid index in the long term, as well as the intensities of various vehicle types. At that point, all the preparation had completed and the project was ready to carry out.

Before starting the thesis, I had the basic knowledge of statistics and programming. I got the government policy of road management in the Netherlands and had the training to analyze the big data set. So I am so thankful for this opportunity. This project is by far, the most challenge I have ever done. I have tried to meet the practical expectations from RWS, and the academic requirements from TU Delft. During this project I have learned so much about data analysis, pavement performance, and road management. Besides, I also learned so much about myself, I have been struggling a lot with understanding the survival model, and the code bugs in Matlab and Python, and got to knew that facing your own strengths, weaknesses and uncertainties is also an important part of growing.

I wish you a pleasant reading journey here!

Zili Wang

February 15, 2020

Acknowledgements

I can not complete the project on my own, and in this page, I want to express my gratitude to all the people who have contributed to this project.

First of all, I would like to thank my daily supervisor Kumar Anupam for all the help, guidance, and encouragement during the last few months. Anupam not only arranged an office and a desktop for me, but also treated me as a member of the pavement engineering group. I struggled with many difficulties of the project, and every time his praise gave me confidence in overcoming them.

Secondly, I would like to thank Hans van Lint, the chair of my graduation committee, for opening the traffic data access to me, the great help with organising the project, and the effective direction. I am grateful to Haneen Farah, my graduation committee member, for the advice on the case feasibility in the preparation phase and the support during the whole project.

Besides, I would like to thank my company mentors, Frank Bouman and Thijs Bennis, from RWS. Without your open minds of the multidisciplinary idea and the willingness of helping a student, I can neither achieve to pioneer the application of real-time traffic data to the road performance simulation, nor complete my graduation as scheduled. Also, thanks for the internship opportunity you provided. My professional ideal is to be an engineer who creates value for the society, and the internship made me really close to the dream.

Further I would like to thank Prof. Sandra Erkens for her efforts to set up the connection between my master thesis and RWS and all the supports. I am thankful to Sander van Nederveen, for sharing the knowledge of BIM and construction management engineering, all the help with my preparation, and the attendance to my mid-term meeting.

Thank Saurabh Sharma, Ghazal Masah, Grace Priskila Kastanya for the great help with the report and the programming part of my project.

A special thanks for my parents, Hongqiong Shu and Tao Wang. Your generous investment in my education made all the things happen. The life lessons from you of being independent, having the moral uprightness, and viewing things dialectically make the base of my way from the past to the future. The freedom you set enables me to arrive wherever I can and back home whenever I want.

The life is long, and it is my honor to meet you guys all in the journey.

Executive summary

This report presents the outcomes of a graduation project, completed with the intent to obtain a master degree in Civil Engineering with Integral Design and Management (IDM) annotation issued by TU Delft.

Problem as given
The graduation project is finished in RWS. The main objective is to figure out the transportation influence on the pavement performance and simulate the road surface performance progression in the service life.

Problem as taken
The importance of taking the challenges is underpinned by the following facts. A good prediction of pavement performance progression has been expected for a long term. In addition, an ideal road management system requires the real-time assessment of pavement condition. There are many root causes of the status quo that the expectation and the system have not been achieved yet. One of them is about the monitoring system. An all-time monitoring provides more timely and precise evaluation of pavement performance. However, in the reality, the measurement is carried out once a year because of the budget and also because it is time-costly. Accordingly, this thesis presents the novelty that using the real-time traffic data as the indirect way of monitoring the road. The data is collected 24/7 automatically, and the flows can be perfectly forecast by the transport flow modeling. The idea is based on the assumption that the transportation has the essential influence on the road performance. Although it is in line with the common sense, the correlation needs to be supported by data. Therefore, the main research question is defined as: *What are the effects of traffic flow characteristics on pavement response and performance in a quantitative way?*

Approach
The research focuses on the collaboration between traffic engineering, pavement engineering and road management, in the road operation and maintenance phase. A design-based research approach is used to develop the simulation software for road performance applied to Pavement Maintenance Information System (IVON). It works out the assessment of whether the road condition meets the maintenance requirement at any determined time according to the data of the influencing factors. The core code of the software is the performance model. Multiple models are used during the research project to figure out the effects of the influencing factors, especially the traffic flows, on the roadway performance progression. There are several data sources. Performance data is obtained by the annual measurement of IVON, as the same database as the modal split. Minute-traffic data is accessed from DiTTlab. Four steps are defined to find the best applicable performance model to the case. Firstly, the raw data is processed and aggregated to the yearly matrices in Matlab. Secondly, the correlation analysis is done for reducing the dimensions of the matrices. Thirdly, the regression models, the survival model and the decision tree classifier model derived from the literature all apply. Lastly, according to the goodness of fitting or the prediction accuracy, the best applicable model is chosen and it is used in the simulation software. The establishment of the software has four phases: building an interface, data preparation, coding the performance model; and the output visualisation.

Findings
The findings of this research are used to evaluate the existing situation and to improve the road management. All the findings are based on the observation of A15 from 2015 to 2018. Of most sections, the road roughness and the rutting were under fluctuation. The findings of the phases for modelling the performance model are elaborated as follows. The correlation analysis shows the weak linear correspondence between the test performance indicators and traffic intensities. Different forms of the performance models explain the effects of the influencing factors in the different ways. The linear regression models indicate that the important factors that affected the test pavement performance indicators were their initial values, traffic volumes and climate, and various vehicle types had different magnitudes of the effects on the road performance. The survival model shows traffic flow significantly decreased the expected survival time and the survival probability of pavement concerning roughness or rutting. Decision tree model predicts whether a section met the maintenance requirements by its characteristics and the traffic situation, and the prediction results are comparable with the actual situation. The three performance models provide the feasible ways of quantifying the influence of traffic factors. But since the results are all

data-sensitive, the quantities of the effects of traffic factors on road performance computed in the study describe the scenario of A15, but not address a general conclusion. The findings verify the assumption that the transportation has the essential influence on the road performance as the problem as taken elaborates. Thus, it is feasible to simulate the pavement performance by the properties of the road and the traffic situations.

Application

The practical application of the findings is to set up the simulation software as a plug-in of IVON system. It is to real-time emulate whether a roadway meets the maintenance requirements and sends the alarm to the road management organisation when it does. Decision Tree Classifier (DTC) is the core algorithm. The data input includes two parts. One from the IVON dataset contains the last measurements of roughness, rutting, the construction times, the surface material types, the locations, the directions, and the modal splits. The other provided by DiTTlab includes the minute-traffic flows on per 100-meter section. The output of the software is the simulation result of whether the section needs maintenance. The test result of using the historical data from 2015 to 2017 to predict the situation in 2018 is more than 88% identical to the reality.

Contents

List of Figures

List of Tables

List of Acronyms

AASHO American Association of State Highway Officials

AASHTO American Association of State Highway and Transportation Officials

ANN Artificial Neural Networks

ARAN Automatic Road Analyzer

AS Actual Skid Resistance Score

BIM Building Information Modeling

CD Condition Points

COMBID Special Pavement Surface Layer

DAB Close Asphalt Concrete

DBFM Design, Build, Finance and Maintain Contract

DiTTlab Delft Integrated Traffic and Travel Laboratory

DTC Decision Tree Classifier

DWW Civil Engineering Department in Rijkswaterstaat

EAB Emulsion Asphalt Concrete

ESAL Average Equivalent Single Axle Loads

GPO Major Projects and Maintenance Department in Rijkswaterstaat

HDM Highway Design and Maintenance Standard Study

HRI Half-car Roughness Index

IDM Integral Design and Management

INWEVA Database of Traffic Intensities of National Road Network in the Netherlands

IRI International Roughness Index

IVON Pavement Maintenance Information System

LTPP Long-Term Pavement Performance Program published by the Federal Highway Administration of the United States

MJPV Multi-year Pavement Maintenance Planning System

OAB Open Asphalt Concrete

PSI Present Serviceability Index

RD Coordinates National Triangle Coordinates in the Netherlands

RWS Rijkswaterstaat, the executive agency of Ministry of Infrastructure and Water Management in the Netherlands

SCW Dutch Study Center for Road Construction

SI Skid Index

SMA Special Asphalt Concrete

TRRL Transport and Road Research Laboratory

TU Delft Delft University of Technology

VTM Voids in Total Mix

WINFRABASE Center Database of DWW in RWS

ZOAB Zeer Open Asphalt Concrete

ZOAB+ New Zeer Open Asphalt Concrete

ZOABTW Double-layer Zeer Open Asphalt Concrete

ZOEAB Zeer Open Emulsion Asphalt Concrete

List of Symbols

α	Coefficients of the roughness increment to be estimated in the regression model
β	Coefficients of the rut depth increment to be estimated in the regression model
d	Difference of the rankings of X_a and Y_b
ΔCR	Incremental values of area of indexed cracking, percent
ΔIRI	Increment of IRI
ΔNE_t	Incremental values of cumulative traffic loading during the time period t, million ESAL
p_1	The share of the passenger cars of the mix traffic
p_2	The share of the light trucks of the mix traffic
p_3	The share of the heavy trucks of the mix traffic
ΔPAT	Incremental values of area of patching, percent
ΔPOT	Incremental values of total volume of potholing, with the unit of m3/lane/km
ΔRUT	Increment of the rut depth, mm
t	Time period
x	Exponent of the variable of T
y	Exponent of the variable of the cumulative traffic flows
f	Influencing factor of pavement performance
IRI_0	Initial IRI at time T = 0, constant for given range of modified structural number
λ	Coefficient of the influence factor in the survival model
n	Coefficient of age
NE_t	Cumulative traffic loading at time t in the unit of million ESAL
$\overline{X_a}$	Average of X_a
$\overline{Y_b}$	Average of Y_b
p_i	Probabilities of each class i in the given branch
rho^K	Kendall's tau coefficient
rho^P	Pearson's linear correlation coeffcient
rho^S	Spearman's rho
RUT	The rut depth, mm
S	Number of the training samples that falls in the branch
T	Age of the pavement since rehabilitation or construction year
T_s	Survival time
T_{s0}	Baseline survival time
T_{str}	The start time of the time interval of the survival time
T_{stp}	The end time of the time interval of the survival time
u	Constant value standing for the non-traffic associated influence
w	Constant value standing for the non-traffic associated influence
X	Matrix of the variables
X_a	Column of one variable
Y	Matrix of the variables
Y_b	Column of one variable

I

Discover

Discover phase starts the research project. The phase is a “phase of divergent though” to see problems from a wide perspective. The phase aims to identify the research problem that occurred during the road operation and maintenance phase and estimate the chance to improve road management by introducing the traffic data to the present system.

There are two chapters in Discover phase. Chapter 1 gives the introduction of the problem, the goal, and the questions, the scope, and the scientific relevance of the research, as well as the report structure. Apart from these, IVON system, as the context of the project, is included in the chapter. Chapter 2 explores the problem by reviewing the previous literature.

1

Introduction

This research is a graduation project for the master degree of Civil Engineering in TU Delft, performed for RWS in Rotterdam, the Netherlands. The content of this research project is described in this chapter. Firstly, the problem that will be researched is introduced in Section 1.1, followed by the goal and the questions of the research including the main research question and the sub questions in Section 1.2. Section 1.3 describes the scope of the project. The scientific relevance and the report structure are elaborated on in Section 1.4 and Section 1.5 respectively. The context of the project is the present road management system RWS is using, which described in Section 1.6. The chapter ends with the short summary of the chapter and the answers to the sub question 1 to 3.

1.1. Problem introduction

Many studies have found that pavement conditions significantly influence traffic safety. 16% of traffic crashes are due to roadway environmental factors (NHTSA, 2008). Pavement roughness or present serviceability rating has a large impact on crash rates (Al-Masaeid, 1997). The pavement factors of International Roughness Index (IRI), pavement rutting, and pavement condition rating significantly influence the frequency of crashes (Abdel-Aty, Devarasetty, & Pande, 2009). Therefore, it is significant for a good pavement management system to have the accurate and timely assessment of pavement state. In reality, most pavement distress is detected by inspection, and only a few important roads or bridges are monitored by sensors. In the Netherlands, the detection of pavement poor performance is dependent on the field measurement (Rijkswaterstaat, 2017b). The pavement performance assessment on the basis of observations is very labour-intensive and it is difficult to observe the damages in time together, especially internal failures of the materials. The recent developing sensor-based structural health monitoring system with the characteristics of sensitivity, timeliness, and accuracy is being applied to more and more infrastructure projects (Farrar & Worden, 2006). However, the technical challenges like the limited life spans of sensors or the large expense of the operation of the whole system, impede the wider application of the technology. With respect to the difficulties together with the availability of the current technologies for a road health evaluation, this thesis proposes a new method to simulate in-service pavement performance based on the real-time traffic data. It is an innovation that the real-time traffic data is taken as the predictor variables for roadway damage simulation.

The current pavement performance assessment of RWS relies on the measurement of the materials properties and the defects (Rijkswaterstaat, 2017b). But the pavement health simulation of the project is driven by the real-time traffic data and looks into the external factors and their potential causes. As the output of the 24-hour vehicle loop detectors, traffic data captures the live transportation situation. The research proposal is to further employ the traffic data for simulating road performance during its service life. Compared with the traditional manual inspection, the pavement health simulation based on traffic data is timely and efficient, because the data collection is real-time, 24/7 and automated. In addition, the new way of roadway health monitoring can have better predictability for road health development compared with the existing system. Road damages are observable only after the deterioration has occurred in the present, but by estimating the future transit intensity based on the current data, the new method is expected to predict the pavement condition development by the prediction of the traffic flow. With the advantage over the recent road performance assessments, the traffic data-driven simulation has a bright prospect for the timely, automated, and predictive pavement health evaluation system. Stakeholders, like the Ministry of Transport and Infrastructure, RWS and other maintenance organisations, can all benefit from the technology application if it is able to reduce a great cost of manual patrols for damage detection, provide potential distress alarm for maintenance, and simulate the pavement performance for asset management.

1.2. Research goal and questions

With respect to academics, the thesis aims at contributing to the existing body of knowledge on the performance models. In the industry level, the research objective is to improve the road management system by setting up the real-time assessment for monitoring the pavement condition. The real-time assessment is based on the data-driven model, and expected to be easily utilized by engineers of road operation and maintenance field.

To be able to set up the real-time assessment of pavement condition, the real-time traffic data is considered as the indirect way of monitoring the road. The idea will be feasible only if the transportation has the nonnegligible influence on the road performance progression. Although it is in line with common sense, the correlation needs to be supported by data. Therefore, the main research question is defined as follows.

Main question

What are the effects of traffic flow characteristics on pavement response and performance in a quantitative way?

In order to answer the main research question, seven sub research questions are formulated:

Sub questions

Discover:
SQ 1. What are the roles of pavement performance and traffic in the current road management system? What kinds of pavement performance indices and traffic characteristics are important in this system?
SQ 2. What can cause pavement deterioration? Which causes are related to or determined by the traffic characteristics?
Define:
SQ 3. What performance models are feasible to perform in this case and can achieve the research goal?
SQ 4. Which pavement deterioration, its potential causes and traffic characteristics can be captured in the current data collection system? Which can not?
Develop:
SQ 5a. How to apply the regression models?
SQ 5b. How to apply the survival model?
SQ 5c. How to apply the decision tree classifier model?
SQ 6a. What are the results of the regression models by the test data?
SQ 6b. What are the results of the survival model by the test data?
SQ 6c. What are the results of the decision tree classifier models by the test data?
Deliver:
SQ 7. How to apply the findings of the quantitative effects of the traffic characteristics to improve the road management field?

1.3. Research scope

The thesis project is two-folded:

- From an academic perspective, on the one hand, this research aims at contributing to the existing body of knowledge on the performance models, that quantify the effects of the influencing factors on road deterioration. On the other hand, the application of the models to simulating the pavement performance progression will be evaluated.
- In the industry level, this thesis aims at setting up a practical tool for real-time evaluating the pavement conditions, which helps to solve the problem that the relatively long measurement intervals of pavement performance may cause.

The influence of the factors on road performance
The research aim is to study the effects of the influencing factors on the pavement performance by the data analysis only. The statistical methods are applied to see the closeness of the correlation between the variables. The pavement performance is considered as the response variable, and the influencing factors are the independent variables. The main focus of the factors is traffic-associated.

The technical feasibility of the service-life pavement performance simulation
The project examines the technical feasibility of simulating road performance by three principles. The first is, there should be a strong correlation between the influencing factors with the road surface performance analyzed by data. The second is the performance model should be perfectly applied to the case, which is assessed by the goodness of fitting or prediction accuracy in the test. The third is that it should be possible to link the traffic data to the measurement results on the same road section.

The establishment of the simulation software
Developed in Python, the simulation software is to provide the engineers in the road operation and maintenance field with the useful and reliable simulation results of the Dutch highway performance in the real time. It has been designed by the report with three functions. The first is that it can read the after-processing data originating the historical performance data in IVON and real-time traffic data from DiTTlab. The second is that it gives the visualisation of the decision tree, which is the core algorithm of the simulation. The third is that it shows the simulation results which makes the statement as the investigated road section has (not) met the maintenance requirements concerning the performance indicator yet.

1.4. Scientific relevance

Not much research has been conducted on the interface between traffic engineering and pavement engineering. This thesis project with the multidisciplinary topic explores the correlation between the two tracks and finds the cooperation valuable in the road operation and maintenance field.

It should be noticed that the road maintenance strategy is relevant to the case. RWS has made the great achievement of the road management system for all classes of the roadways in the Netherlands (Rijkswaterstaat, 2006). The system includes the criteria of road construction and repair, the service standards of the urban and regional infrastructure, the code of conduct of the inspection and the measurement, the damage assessments, the maintenance requirements in the Design, Build, Finance and Maintain Contract (DBFM), the standard process of the road maintenance decision, and the other subsystems. Accordingly, the thesis applies the current system and the predefined criteria when it requires the knowledge of road maintenance.

As stated before, the core of the report is data analysis. With the development of machine learning, data analysis is filled up with tools. Some complex relations between the variables are hard to define clearly by the classic statistics, but that may be discovered by the advanced data analysis tool. This case is expected to benefit for the development of the machine learning. The project explores the practical application of the algorithm in the field of the road maintenance decision and it develops the technology by adding one practical case of the service-life pavement performance simulation.

1.5. Report structure

This research is based on the design-based research approach. The design approach is used to get a full understanding of the complex problems and to come up with solutions that tackle these problems. “Design is a creative approach to problem solving with the power to tackle complex and pressing social issues. It is people-centred, getting straight to the heart of an issue to encourage new perspectives and generate powerful ideas." (Design Council, 2003). There are many design models, such as the service design thinking model, the product innovation process model and so on. This report chooses the double diamond model developed by the UK Design Council because it defines the process clearly and it can be easily and flexibly applied (Design Council, 2005). The double diamond is divided into 4 phases: 1) Discover the context, (2) Define the current situation, (3) Develop the desired situation, and (4) Deliver the outcome of this research. The report is organised based on the process of the design approach. In Fig. 1.1, an overview is shown as well as the connection between the chapters. The elaboration on each phase is given as follows.

Discover phase
In this phase, at first a problem is formulated clearly, and then the context in which the research takes place is elaborated on. There are two key points when formulating the problem. One is that it should be valuable, and the other is that it should be feasible. On the foundation of the literature overview of the field, a valuable problem is formulated. With the application of the design-thinking method, the feasible research approach is proposed. Then the investigation of IVON, the currently-used road management system in the Netherlands, is carried out to understand the context in which the research takes place.

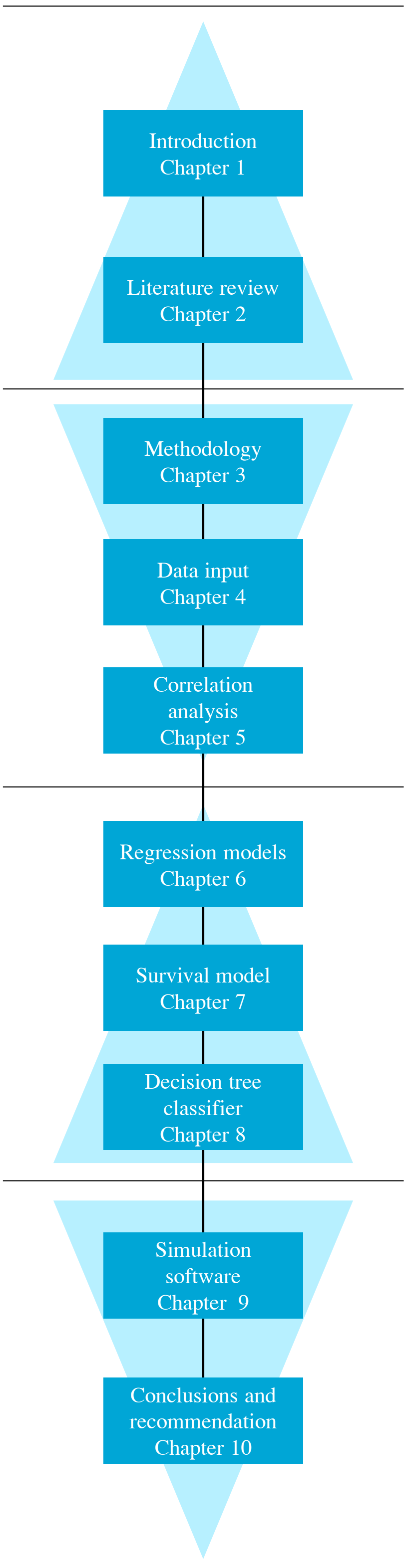
Define phase
Define phase for determining the focus of the study area is a convergent process as the design-thinking method indicates (Design Council, 2005). The phase is based on the investigation in Discover phase, that presents the pavement performance is evaluated in various aspects in IVON, and there are many factors considered in the road operation and maintenance. The phase starts with the formulation of the model framework of the pavement performance, the performance models, and the potential influencing factors. But actually there are many variables in the framework does not included in the current database. It means that only the data is available, the analysis of the effects in the quantitative way can be achieved further. Therefore, the focus is narrowed to study the variables that can be gained from the current data collect set. Correlation analysis, the classic statistical method to quantify the strength of the relation between two variables is carried out. After that, this phase will end with the focus on the rather strong correlation between the pavement performance indicators and the influencing factors as the correlation analysis shows.

Develop phase
In the phase, the quantitative approach is applied to provide the answer to the main research question, and the different models are used to come up with ideas about the pavement performance progression simulation. According to the literature reviews of the performance models in Discover phase and the methodology definition in Define phase, the regression models, the probabilistic model, and the classification models are chosen. This phase inputs the test data in the multiple forms of the regression models, the survival model, and the decision tree classifier. At last, the model results are presented for the selected case study.

Deliver phase
The final phase is the practical application of the findings of the research. As Develop phase proposes, several performance models can be used to set up the simulation software. Thus Deliver phase begins with the selection of the best applicable model for the simulation of this case. The selection result is the synthetical consideration with the assessment by the goodness of fitting or the prediction accuracy of the test results and RWS requirement of user-friendliness. After then, the simulation software is designed and built. The phase ends with the overall of the thesis, the conclusions, and some recommendations.

1.6. Practical context: IVON system

The thesis addresses the problem in the road management field. So it is necessary to understand the current used road management system, as the practical context of the problem. One of the main tasks of Major Projects and Maintenance Department in Rijkswaterstaat (GPO) is consultancy in the field of construction, management and maintenance of road structures. The department prefers the fast, simple and cheap solution to implement one of the primary tasks of RWS, the management and maintenance of the national road network. The task requires the quality level of the roads (the freeways, the distributor roads, and the access roads), the time estimate of when the maintenance is needed, the maintenance implementation plan and the costs. To take on it, the quality of the national roads network is evaluated annually through Multi-year Pavement Maintenance Planning System (MJPV). The whole system is elaborated in the document by Rijk-



I
Discover

The aim of the discover phase is to address the research problem and explore the context of in which problems occur. In this phase IVON system is described and the present knowledge of the field. It provides an overview of the involved factors during the road operation and maintenance process. At the end of the discover phase the valuable and feasible problem is shown that occurs in the road operation and maintenance phase.

II
Define

The define phase aims to determine the focus of the study area and the research approach. In this phase it is analyzed what the factors and pavement performance are applicable to study under the current conditions and in what way the collaboration between the factors and pavement performance can be analyzed. At the end of the phase, the selection results of the variables and the performance models are defined that will be used in the next phase.

III
Develop

The objective of the develop phase is to test the performance models for the simulation in the next stage. In the phase, various widely-used performance models are applied to the case study. The different ways of describing the influence of the factors on the road as the model results indicate are the critical answers to the main research question.

IV
Deliver

In the deliver phase, the evaluation of the technical feasibility of the simulation is given as well as the selection results of the best applicable performance model to the case. The implementation phase of the simulation software is designed. The deliver phase ends with the conclusion of the overall project and some recommendations.

Figure 1.1: Report structure

swaterstaat (2006). This planning is also used in the context of the road management cycle in Fig. 1.2 in the long term, and the performance indicators recorded in Center Database of DWW in RWS (WINFRABASE) determine the service of the national road network as the primary service evaluation reports annually.

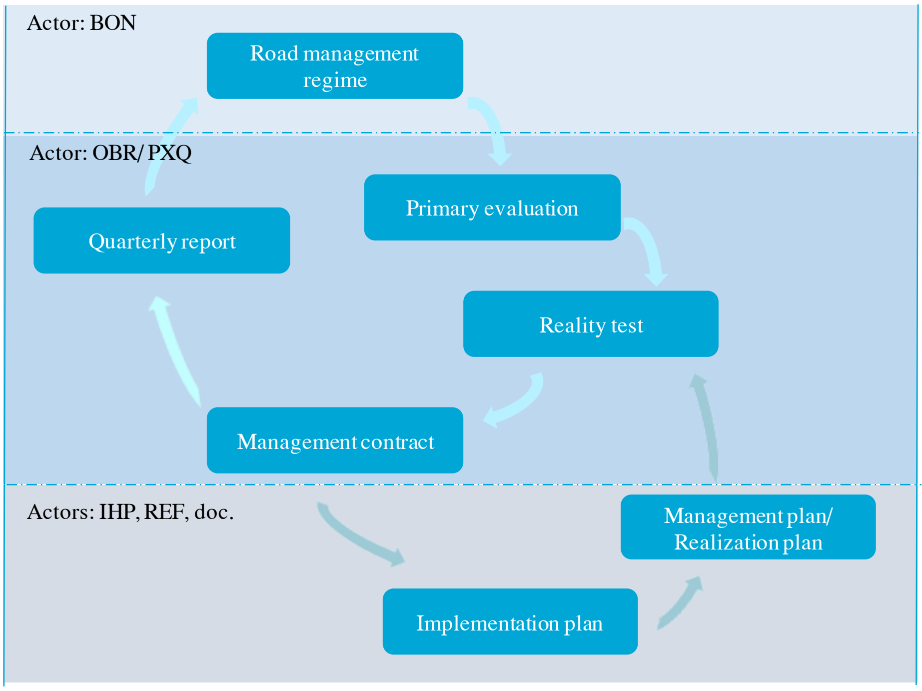


Figure 1.2: Road management cycle

MJPV sets up the strategic plan for the task, and IVON as framed in Fig. 1.3 provides the tactical plan. The framework starts with the data records in WINFRABASE and the preconditions that are translated into knowledge tables. WINFRABASE contains the following data supplied to IVON:

- Road structure: details of the paving (construction year and material type). The data is collected via the road data inventory procedure. It is updated sometimes by the video images captured by Automatic Road Analyzer (ARAN) or the visual inspections of the counselors.
- Road network: the locations in National Triangle Coordinates in the Netherlands (RD Coordinates), the directions, the lengths and the widths of the roads, and the districts the roads are situated at.
- Quality: the measurement data for the damage characteristics, including the rut depth and the longitudinal flatness, collected by the ARAN and the skid index captured by Civil Engineering Department in Rijkswaterstaat (DWW).

In addition to the input from WINFRABASE, as described above, IVON also uses:

- Data from the visual inspections by the consultants. These are initially copied from the previous MJPV.
- Knowledge tables: the maintenance guidelines, damage assessment and the interventions levels for the pavement maintenance, the assumption of the average damage development (with the consideration of climate, traffic, etc.), and the advice on the major maintenance on asphalt pavement.

IVON is a software package developed by DWW, which is a part of the production line ARAN-WINFRABASE-IVON. The consultants support for the whole process of the long-term plan. As Fig. 1.3 shows, IVON has two key elements: paving technical planning and implementation planning. In Fig. 1.4, paving technical planning consists of eight modules and implementation planning has two modules. The detailed introduction of the two elements is given as follows.

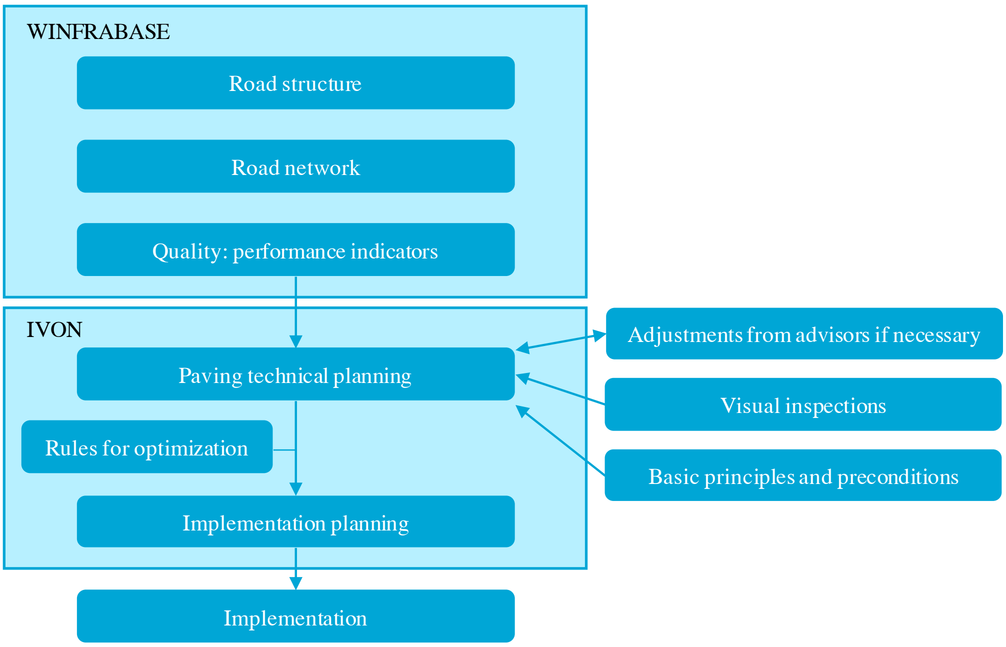


Figure 1.3: IVON system framework

Paving technical planning
With the input data and the basic principles and preconditions described above, IVON program creates the paving technical planning. 8 modules in total has been developed to make the plan in Fig. 1.4. The technical planning is divided into two phases. One covers the lifetime from the construction to the fifth year, and the other is made for the period from the sixth year until the tenth. In the first five-year plan, it is according to all kinds of relevant information, such as the damage characteristics, the characteristics of materials, traffic (in particular trucks), climate, and so on. The paving technical planning is finally determined by the advisors from DWW based on their professional knowledge. As for the road sections operated longer than five years, the technical plannings are based on the measured damage characteristics and the road quality as recorded. As a result, the paving technical planning provides where and when the maintenance is required on the road network.

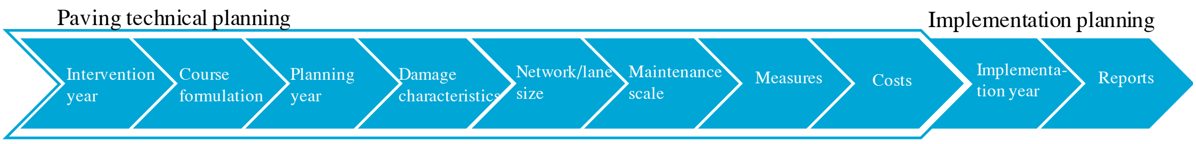


Figure 1.4: IVON program

Implementation planning
Based on the paving technical planning, the implementation planning is made. IVON uses the rules for optimization, which implies that the account has been taken into: the current states of the pavements, the intervention levels, the project costs, and the probability of the pavement performance progression to the worse result in the following years. The planning finalizes the maintenance plan of year one and two, and it does not allow any adjustments once the maintenance decision has been made up. Besides, it indicates the required compartment lengths of the new paving materials, like Zeer Open Asphalt Concrete (ZOAB), and calculates the costs.

Damage assessment
Damage assessment concerning rutting and roughness in the article by Rijkswaterstaat (2017b) is set up by

the comprehensive considerations of the influence of the damage on serviceability and the maintenance costs. The intervention levels of the skid resistance is presented in another article by Rijkswaterstaat (2017a). The severity of pavement damage is classified into four classes in Tab. 1.1. Class I is the good performance. Class II should be paid attention but still acceptable. Class III is the intervention level, and Class IV is so inferior that the road section can not provide the safe service and the maintenance action should be taken.

Table 1.1: Damage assessment

Performance index	Class			
	I	II	III	IV
Rut depth (mm)	<10	10 ~17	18 ~23	>23
IRI (m/km)	<2.6	2.6 ~3.4	3.5 ~4.0	>4.0
SI	>0.04	0.03 ~0.04	-0.03 ~0.02	<-0.03
AS	>0.04	0.00 ~0.04	-0.06 ~-0.01	<-0.06

1.7. Conclusion

The chapter states the research problem, the goal, the research questions, the scope and the scientific relevance and the report structure of the project. It gives an overview of the road management of IVON and the roles of the pavement performance and the influencing factors in the road operation and maintenance phase. Below the sub research will be answered.

Sub question 1

What are the roles of pavement performance and traffic in the current road management system? What kinds of pavement performance indices and traffic characteristics are important in this system?

Pavement performance is the core evaluation of the road service quality, and it is the elementary input to make up the paving technical planning. The performance indices are the decisive factors of whether the road is in need of repair or not.

Rut depth, IRI or Half-car Roughness Index (HRI), and Skid Index (SI) or Actual Skid Resistance Score (AS) are taken into account as three kinds of pavement performance indices in the current system. They describe the road performance, an abstract concept, in terms of three aspects: the tracks, the longitudinal flatness, and the skid resistance respectively. The road maintenance decision is made up when any index develops up to the intervention level. The intervention level of the rut depth is above 17 mm. As for the roughness, when IRI is larger than 2.6 m/km, it is the warning level, and if it develops above 3.4 m/km, the repair should be planned. About the skid resistance, the intervention level of SI and AS are below 0.03 and 0 respectively.

As for the traffic, it plays the role as a factor in the paving technical planing in the road management system. The essential goal of road management is to provide good service to transportation. Any restraint of free traffic flow caused by the road management (in most cases, the road construction) is computed as the vehicle loss hours, transferred in the form of money, and counted in the cost of the planning. As for the influence of the traffic on the road maintenance decision, traffic, more specifically, the number of trucks, is considered as the consultancy information, while the core indicators in the current road management system are the measurements of the distress.

The next chapter will introduce the relevant literature. Based on the existing knowledge, the model framework will be established in the next phase.

2

Literature Review

The chapter gives a theoretical review of existing knowledge about pavement performance and its causes during the service life from an academic perspective. Theories described in this chapter will be used as a lens to look at the effects of the traffic and other factors on the road performance in the operation and maintenance phase. Section 2.1 provides an overview of the pavement performance. Section 2.2 is to study what can cause the pavement deterioration. Section 2.3 focuses on the traffic causes. Section 2.4 introduces the performance models that are most used in the researches. Section 2.5 concludes what has been found in the literature.

2.1. Pavement performance

Pavement performance is a widely-used technical item, but previous articles defined it in different ways. Cary (1960) first presented the pavement performance concept and defined it as “serviceability history”. After that, the concept of performance was enriched. The discussion about the definitions of pavement performance is not the focus of the study, but helps to build up the technical context. The thesis refers the work of Huang (1993) to get an overview of the item. Pavement performance can be evaluated with the respective of:

- Distresses. Various defects develop in the pavement under the combined impacts of traffic loading and environmental conditions, which affect the functionality negatively. According to the distress identification manual of Long-Term Pavement Performance Program published by the Federal Highway Administration of the United States (LTPP) (J. S. Miller, Bellinger, et al., 2014), the deterioration for pavements with asphalt concrete surfaces can be categorized into 5 categories: cracking, patching and potholes, surface deformation, surface defects, and miscellaneous distresses.
- Serviceability. It is the ability of a specific section to serve traffic in its conditions. Present Serviceability Index (PSI) and IRI are used to determine the serviceability. PSI developed at the American Association of State Highway and Transportation Officials (AASHTO) Road Test is based on pavement roughness as well as distress conditions, and IRI is determined only by the road profile.
- Skid resistance. Adequate skid resistance highly dependent on the surface friction guarantees no loss of control and skid accident in normally expected situations when the pavement is wet.
- Structural capacity. As an important index of design and structural condition evaluation, structural capacity is commonly assessed by deflection measurements during the service life.

2.2. Causes of pavement deterioration

The cause of the defects is not simplex. Many reasons or combination of reasons that contribute to failures of bituminous pavements has been recognized as traffic, environmental factors, structural failures, design, and construction quality. The environment factors mainly contain moisture and frost, and the structural failures are the deficiencies of subgrades, joints, road shoulders, or layers (Adlinge & Gupta, 2013). Based on the desk research, the distress and the causes are elaborated respectively and the results are simplified as shown in Tab. 2.1.

- Fatigue cracking is primarily caused by the accumulation of damage imparted by the traffic load. Observed in LTPP fatigue cracking data, it has the feature that no significant cracks appear for several years, but then occur and propagate to a significant level in a short time. To analyze the effects, Average Equivalent Single Axle Loads (ESAL) and traffic levels categorized by Asphalt Institute (1996) are used. Observations of LTPP fatigue cracking data revealed that most cracking in the wheel path followed this pattern: no significant cracks appear for several years; then cracks appear and soon propagate to a significant level. Some typical fatigue cracking development patterns is the cracking area with the unit of the square meter.

- Longitudinal cracking is caused from the perspective of mechanics by the lack of restraints at the mid-slab edge or at the slab corner. In reality, it can occur when a truck is driven along with the slab (Heath, Roesler, & Harvey, 2003).

- Transverse cracking was found that transverse cracking could occur under either environmental loading or combined environmental and traffic loading. The findings illustrated here are based on the test that the slab thickness is 200 mm (Heath et al., 2003). For the environmental loading situation, transverse cracking would only occur if the slab is fairly long and the shrinkage gradient is high: bottom-up cracking under edge loading with no shrinkage gradient and a positive temperature gradient; top-down cracking of a long slab with a medium or high shrinkage gradient (0.35 me/mm in the research) and a negative temperature gradient.

- Slippage cracking, theoretically, is the consequence of ineffective adhesion due to poor quality viscous coatings and/or inadequate rates of application. Other factors include the low asphalt content and the high aging rate, which reduce the effectiveness of the binder (Chen, 2009).

- Rutting, can increase under the traffic load. Some studies have shown typical truck tire pressures to be approximately 120 psi. The amount of Voids in Total Mix (VTM) is likely the most important physical property of asphalt mixtures that relates to rutting. Once rutting starts, VTM may lower underneath and actually gets worse with the additional traffic. The various layers in the pavement will also have variations in the amount of VTM. Low VTM near the surface of the pavement can result in serious rutting problems (Brown & Cross, 1989). Besides, some study predicts rutting by repeating load trial compression test.

- Corrugations & Shoving, the forms of plastic movement typified by ripples or an abrupt wave across the pavement surface, usually occur at places where vehicles accelerate or decelerate (Adlinge & Gupta, 2013).

- Potholing & Patching, are related to the load-associated foundation breakdown (Adlinge & Gupta, 2013).

- Ravelling, is categorized into three levels: low, moderate and high. The correlation factors of ravelling with material or construction properties respectively are 0.986, 0.926, and 0.976, which has been studied by Artificial Neural Networks (ANN) (Miradi, 2004).

- Polishing, caused by inadequate resistance, aggregates particularly in areas of heavy traffic movements or where high stresses are developed between surface and tyres (Sorum, Guite, & Martina, 2014).

Additionally, skid resistance, as the important aspect of the pavement performance, is closely related to traffic safety. It is a common fact that the lower the skid resistance value, the higher the percentage of the traffic accidents. Skid resistance is generally quantified by the friction measurement such as the coefficient of friction or skid number. There are four main factors (Huang, 1993):

- The vehicle: friction demands vary greatly, dependent on the speed of a vehicle as well as its safety system.

- The weather: especially wet pavements and the thickness of water film affects the available friction.

- The driver: the skill of the operators affects the potential for loss of control or skidding.

- The roadway: some distresses on the pavements have an influence on the skid resistance, like bleeding, polishing, smooth macrotexture, rutting, and inadequate cross slope.

Table 2.1: Pavement distress and its causes

Distress type	Classification	Causes	Traffic	Moisture	Frost	Structure	Design	Construction
Cracking	Fatigue cracking	Repeated traffic load	√					
	Longitudinal cracking	Load-induced, frost heaving or joint failures	√		√	√		√
	Transverse cracking	Load-induced, frost heaving or joint failures	√		√	√		√
	Block cracking	Lack of compaction during construction						√
	Slippage cracking	Horizontal force from traffic, poor bonding between the surface and the layer below or lack of the tack coat	√				√	√
	Reflective cracking	Pavement is overlaid with hot mix asphalt concrete and cracks reflect up through the new surface						√
	Edge cracking	Lack of support of the shoulder due to weak material or excess moisture		√		√		
Surface deformation	Rutting	A very narrow rut is usually a surface failure, while a wide one is indicative of a subgrade failure. Inadequate compaction can lead to rutting				√		√
	Corrugation	Vehicles accelerate or decelerate too much. Too much fine aggregate, or rounded/smooth textured coarse aggregate.	√				√	
	Shoving	Vehicles accelerate or decelerate too much. Too much fine aggregate, or rounded/smooth textured coarse aggregate.	√				√	
	Depression	Localized consolidation or movement of the supporting layers beneath the surface course due to instability				√	√	√
	Swell	Frost heaving (subgrades with highly plastic clays can swell in a manner similar to frost heaves but usually in warmer months) or by moisture (an expansion of the supporting layers beneath the surface course or the subgrade)		√	√			
Disintegration	Potholes	The pavement disintegrates under traffic loading, due to inadequate strength in one or more layers, usually accompanied by the presence of water	√	√		√		
	Patches	The pavement disintegrates under traffic loading, due to inadequate strength in one or more layers, usually accompanied by the presence of water	√	√		√		
Surface defect	Ravelling	Ravelling can be accelerated by traffic and freezing weather. Some ravelling in chip seals is due to improper construction technique	√	√				√
	Bleeding	Improperly-applied seal coat. Excessively high asphalt cement content in the mix with too low viscosity (too flowable), too heavy a prime or tack coat					√	√
	Polishing	Traffic	√					

2.3. Transport factors

Pavement performance evaluation is a complex question that includes distress, serviceability, skid resistance, and structural capacity (Huang, 1993). Theoretically, these features are the reflections of both internal and external factors. The internal factors relate to the pavement quality resulting from the material production and overlay construction, while the external factors mainly contain the temperature, rainfalls, and traffic flow in the service life (Zhang, Lepech, Keoleian, Qian, & Li, 2009).

Traffic, as one of the main exterior influence on roadways during service life, especially affect pavements by heavy vehicles (Adlinge & Gupta, 2013). A number of researchers have focused on the failure mechanism of road surfaces generated by repeating load, trucks and overloaded vehicles. Most empirical distress models use several parameters to represent the mixed traffic, such as ESAL and truck percentage. However, the thesis takes into account a variety of behaviors of all kinds of vehicles captured by real-time traffic data, and aims to figure out the pavement deterioration model that predicts future pavement condition on the basis of the transport impacts.

Accordingly, the study divides the relative damaging effects into traffic- and non-traffic associated attributions. Both significantly affect the deterioration initiation and progression, but they are influential variously on the different distress modes. It is complicated to identify the individual effects of the factors or their own attribution in a quantitative way because of the interactions and the combined effects. The traditional approach combines the real world factors of mixed traffic and long-term environmental impacts and use a fourth-power law to compare (Paterson, 1987).

Traffic, as the heterogeneous mixture of many factors, varies from road to road, over time with the flow growth and changes, and across countries. It can cause the pavement fatigue and failures, and has the trifling influence on the scale of roadway profile. Accordingly, the most relevant road performance indicators to transport factors are distress, serviceability, and skid resistance.

Essentially, the traffic-associated cause of pavement deterioration is the dynamic load which generates the road failure mechanisms. It can be assessed by three aspects: the vehicle composition, the loading configuration and the whole intensities (Gillespie, 1993). There are many pieces of researches about the load effects on the road structure, especially the heavy and overloaded vehicles. Different forms of pavement distress, fatigue, and permanent deformation are contributed significantly by heavy vehicles (Gillespie, 1993). The axle load above the maximum legal limit causes significant damage on a pavement, increasing the construction and rehabilitation cost (Pais, Amorim, & Minhoto, 2013). In Section 2.2, the defects are elaborated about how they are caused by transportation characteristics respectively according to the literature. The simple correlations between pavement performance and transport are listed in Tab. 2.2.

2.4. Performance models

The approach to predict the pavement performance has developed among ages. The regression model was in the first stage. The damage function, as the first performance model on the basis of AASHTO road test data used simple linear regression to predict a specific distress occurrence in the function of loading. The input was the equivalent standard 18-kip axle load applied up to the time and damage at the same time. Estimated by the regression analysis, two parameters were calculated, which varied from the deterioration types. The one was the equivalent standard axle load which required to produce the damage levels defined as failure, and the other represented the increasing rate of damage. The function had the limitation that it was not precise for all distress modes and only for the specific climate condition and one subgrade. Thus, later on the predictive models based on the regression analysis for all kinds of deterioration being the functions of a variety of factors were presented to a wider range of conditions.

Generally, there are two forms of the regression models. One form deriving from Transport and Road Research Laboratory (TRRL) road costs study in Kenya believed the relationship between the influencing factors and the deterioration was linear fitting (Parsley & Robinson, 1982). The other supposed the regression should be non-linear (Paterson, 1987). The model assumes that the increment of unevenness was related to the value of the original roughness and the external effects on the road during that period of time. According to the Brazilian database, the model was formulated and gave the quantitative relation between the roughness variance and the structural and time-related environmental mechanisms, as well as the effects of surface

Table 2.2: Pavement deterioration and its traffic-relevant causes

Distress	Classification	Causes	Repeated traffic	Edge traffic	Lane change	Heavy vehicle	Speed change
Cracking	Fatigue cracking	Repeated traffic load	√				
	Longitudinal cracking	Edge load		√			
	Transverse cracking	Edge load		√			
	Slippage cracking	Horizontal force from traffic			√		
Surface deformation	Rutting	Increasing truck tire pressure, axle load and volume of traffic	√				
	Corrugations, shoving	Vehicles accelerate or decelerate	√				√
Disintegration	Potholes, patches	Overloaded traffic				√	
Surface defects	Ravelling	Traffic and freezing weather	√				
	Polishing	Heavy traffic				√	

distress such as cracking, patching, and potholing.

The second stage of developing the performance model introduced the statistical procedure in order to overcome the common difficulties in the development of distress models from the empirical data, and proposed the duration modeling (Paterson, 1987). The widely-cited paper used failure-time theory and maximum likelihood estimation methods to include censored data and the prevent biased parameter estimates. Failure time, as the response variable, was modeled via a Weibull distribution. An example application to Brazilian pavement condition gave the quantitative result of the effects of weather and its variability on pavement deterioration. Later on, one research improved the probabilistic duration modeling techniques with the adoption of the Weibull distribution to rate hazard function (Prozzi & Madanat, 2000). The method is successfully applied by undertaking the European database of construction, traffic and climate data in oeder to predict the duration of the crack initiation (Loizos & Karlaftis, 2005). In addition, the survival model, a method of statistics for analyzing the expected duration of time until one or more events happen, have been applied to many pieces of researches. It was firstly used in pavement performance modeling about survival curves of highways in the 1930s (Winfrey, 1969), although they relied more on empirical methods than statistical procedures. Highway Design and Maintenance Standard Study (HDM) initiated by the World Bank, employed the methodology to predict the initiation of fatigue cracking in the HDM-III model. American Association of State Highway Officials (AASHO) road test data were reanalyzed using survival analysis and the result indicated that the survival model is more appealing than the original AASHO formulations (Prozzi & Madanat, 2000). Most recently, researchers also attempted to employ survival models to predict in the pavement fatigue performance from laboratory fatigue test results (Tsai, Harvey, & Monismith, 2003), and fatigue cracking of flexible pavements based on long-term pavement performance data (Wang, Mahboub, & Hancher, 2005).

With the development of machine learning, the classification model is the new approach to establish the corresponding relations between the damaging factors and the road performance. Decision tree, a good solver of decision-making cases, can give the maintenance decision with regard to any performance indicators and, the importance ranking of the attributions as well as their critical values as a result. In addition, ANN, another machine learning classification model, can work out the interconnection between various factors and have been applied to predict non-linear interactions between various variables in complex concrete performance (Miradi, 2004).

2.5. Conclusion

The chapter finds the existing knowledge of the pavement performance, the performance indices, the potential causing factors, and their correlations, and the performance models. In this review the sub research question 2 will be answered as follows.

Sub question 2

What can cause the pavement deterioration? Which causes are related to or determined by the traffic characteristics?

Many factors can cause the pavement deterioration. Design and construction determine the original pavement surface characteristics. Traffic and climate mainly affect its performance in the service time. The structure defines the base, and the damage in it can be transmitted to the surface and observed. Time is a special influencing factor, and the process of aging is essentially through oxidation. It is hard to clarify that the dual and multiple factors result in the additive effects, or positive cycling, or negative cycling. Because in the reality, they always exist in the meantime.
Traffic, itself is an important factor of the pavement performance. The effect of traffic on the pavement performance essentially is because of dynamic load. Mechanically, there are five forms of the dynamic load that can have an influence (particularly, negative) on the road performance progression : the repeated load, the load on the edges of the slabs, the load in the transverse direction, the overload, and the varying load. Macroscopically, the five forms are related to the total intensities, the tracks of the vehicles, lane changing, the truck percentage, and the driving behavior of acceleration or deceleration. Generally, traffic is one of the influencing factors during the service life, but it is hard to clarify how important the traffic factor during the service years because the pavement performance is always an integrated consequence by all kinds of the factors. In fact, due to the slight uncontrollable variables in the construction process, every road section is unique. Therefore, even if there is a ranking of all factors according to the impact on the road performance progression, the sorting result may vary because each road segment has the different features.

The next phase will dive deeper into the methodology, the data availability and the correlation between pavement performance and influencing factors.

II

Define

In the previous phase, the research problem as well as the practical and academic context of the study is addressed. Road performance can be indicated by multiple pavement performance indices during the in-service process. There are many influencing factors of pavement performance theoretically.

Define phase is a convergent process, which starts with the methodology definition, including the model framework, the data correlation analysis method and the pre-defined performance models. In the model framework on the basis of the existing knowledge, not all the pavement performance indicators and the potential influencing factors are accessible in the present data collection system. Thus the entire field is concentrated on the study area by two following steps. Firstly filtering out the factors that the database does not have, and secondly focusing on the strong correspondence between the influencing factors and the pavement performance indices found by the correlation analysis. The findings will be used to set up the performance models in the next phase.

3 Methodology

This chapter elaborates on the methods used in Define phase and Develop phase. Define phase is to address the methodology, the input data, and the correlation of the variables. It starts with the establishment of the model framework of all the relevant variables, and focuses on the variables of which the data is accessible. The model framework is proposed in Section 3.1. The study variables in the research require either the values that can be accessed directly from the database provided by RWS or are able to be computed by the raw measurement data in Chapter 4. After then, the correlation analysis is applied in Chapter 5 to figure out the strength of the relation in a quantitative way between the variables, including the correlation between the performance indices, the correlation between the performance indices and the factors, as well as the correlation between the factors. It is important to do the analysis because of some principles of setting up the mathematics model. One principle is that the independent variables should be independent enough, that means the factors which has the strong relation found by the correlation analysis should not be considered at the same time in one model. The other principle is that in order to achieve the good fitting results, the regression models take the factors which are data dependent on the road performance into account. The functions for the correlation analysis are elaborated on Section 3.2.

As for the methods used in Develop phase, according to the previous study introduced in Section 2.4, three kinds of models are chosen because they are well applied, which are the regression models, the probabilistic model, and the classification model. However, because not all the variables in the pre-defined models are accessible in the study, some adjustments are made when applying the models. The elaboration on the pre-defined regression models, the survival model, and the decision tree model respectively is in Section 3.3, Section 3.4 and Section 3.5 respectively. The model adjustments to the case study are introduced in Chapter 6, Chapter 7 and Chapter 8 of Develop phase. The chapter answers the sub question in the end, which is what performance models are feasible to perform in this case and can achieve the research goal.

3.1. Model framework

The findings of Discover phase presented the pavement performance, the causes of the road damage, especially the traffic-relevant factors, and the performance models. Accordingly, the model framework contains all the aspects of the pavement performance, the performance indices, the potential causing factors, and their correlations in the various performance models, as Fig. 3.1 illustrates. It should be noticed that not all the data is available, for example, the road structure stiffness is not in the database, but it may be relevant to the pavement performance. For the further quantitative analysis, only the factors which were obtained from the database will be studied.

3.2. Correlation analysis

The research defines an independent variable in the way that a change in any other quantities does not cause a change in the independent variable. The definition of the dependent variable is that the variable changes

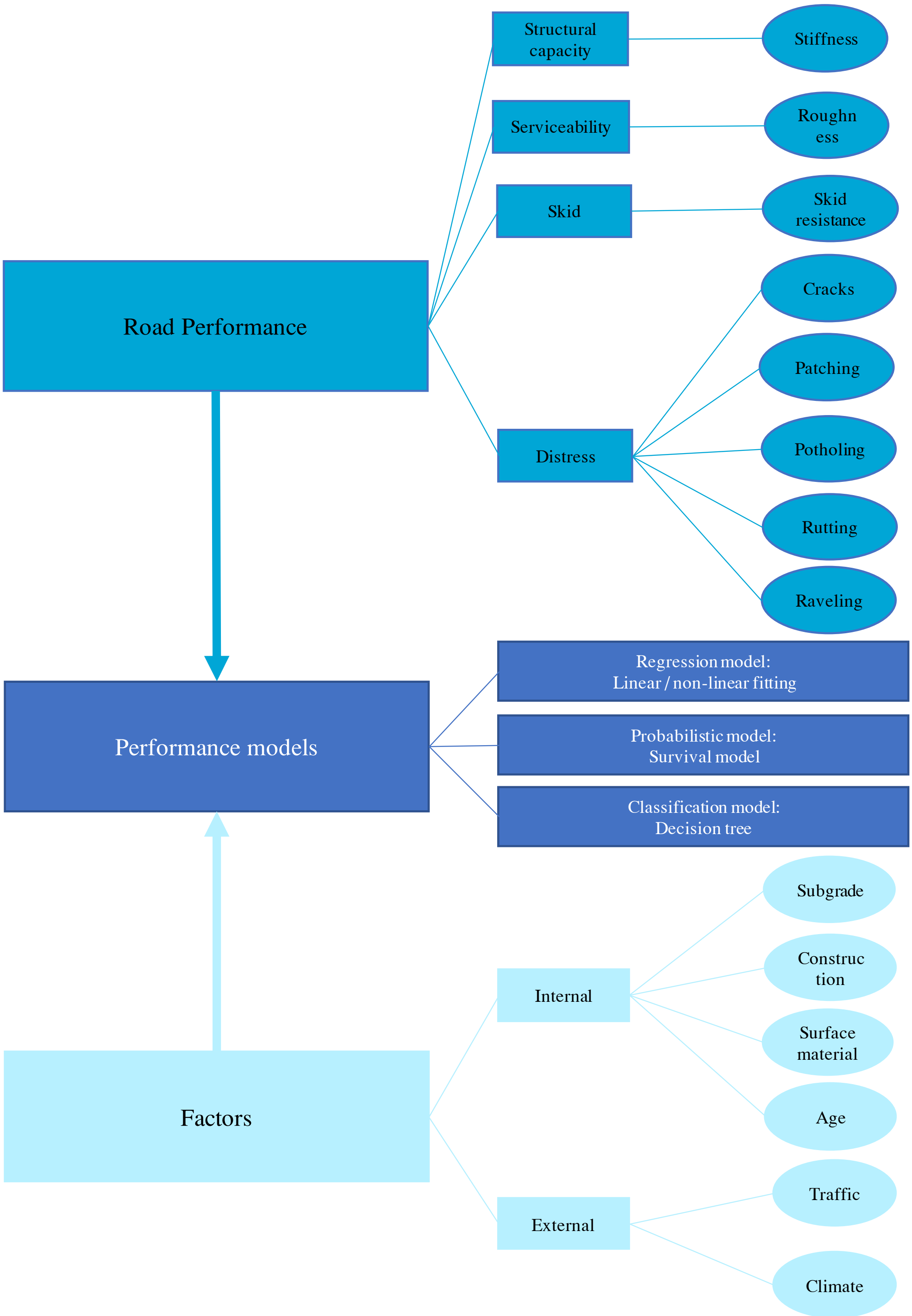


Figure 3.1: The model framework

as any independent variable changes. Taking the dependent variables as the independent variables is a incredible mistake when determining the relationship between the physical items, and vice versa. To avoid it, the study uses the three correlation analysis methods to analyze the linear correlations as well as the rank correlations between the variables, deriving the dependence and independence between the variables, of which the correlation coefficients are Pearson's linear correlation coefficient (rho^P), Kendall's tau coefficient (rho^K), and Spearman's rho (rho^S). The three are most-applied correlation coefficients, aiming to find the correlation of the variables in the different ways (Lee Rodgers & Nicewander, 1988).

Pearson's linear correlation coefficient is to study if there is a linear correlation between two data sets. It is formulated by the column X_a and the average $\overline{X_a}$ in matrix X, and the other column Y_b and its average $\overline{Y_b}$ in matrix Y as equation 3.1 by Mathworks (2006b). rho^P ranges from -1 to 1. A value of -1 indicates a perfect negative correlation, while a value of 1 indicates a perfect positive correlation. If rho^P is 0, it means no linear correlation between X_a and Y_b.

$$rho^P = \frac{\sum_{i=1}^{n}(X_{a,i} - \overline{X_a})(Y_{b,i} - \overline{Y_b})}{\sqrt{\sum_{i=1}^{n}(X_{a,i}-\overline{X_a})^2 \sum_{i=1}^{n}(Y_{b,i}-\overline{Y_b})^2}} \tag{3.1}$$

Kendall's tau coefficient (rho^K) is the coefficient of rank correlation. It is based on counting the number of the pairs of X_a and Y_b that are concordant. The concordance is defined as the difference value of $X_{a,i}$ and $X_{a,j}$ has the same sign as the difference value of $Y_{a,i}$ and $Y_{a,j}$. The formulation of rho^K by Mathworks (2006b) is equation 3.2. It ranges from -1 to 1. If rho^K is -1, it indicates the ranking of X_a is the reverse of the ranking of Y_b. If it is 1, it means the rankings of two variables are the same. The value of 0 means no correlation between them.

$$\begin{aligned} rho^K &= \frac{2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\xi^*(X_{a,i}, X_{a,j}, Y_{b,i}, Y_{b,j})}{n(n-1)} \\ \xi^*(X_{a,i}, X_{a,j}, Y_{b,i}, Y_{b,j}) &= \begin{cases} 1 & if (X_{a,i} - X_{a,j})(Y_{b,i} - Y_{b,j}) > 0 \\ 0 & if (X_{a,i} - X_{a,j})(Y_{b,i} - Y_{b,j}) = 0 \\ -1 & if (X_{a,i} - X_{a,j})(Y_{b,i} - Y_{b,j}) < 0 \end{cases} \end{aligned} \tag{3.2}$$

Spearman's rho (rho^S) is the coefficient of rank correlation, which is equivalent to rho^P applied to the ranking of two variables, X_a and Y_b (Mathworks, 2006b). When the rankings are distinct, rho^S is formulated as equation 3.3. In the equation, d is the difference in the rankings of two variables. The range of rho^S is between -1 and 1. The larger the absolute value is, the more relevant the two variables are.

$$rho^S = 1 - \frac{6\sum d^2}{n(n^2-1)} \tag{3.3}$$

3.3. Regression models

The pre-defined regression models have two forms: linear regression and non-linear regression. The most application form varies in the different case studies. Derived from TRRL road cost study in Kenya, the linear regression models of roughness and rutting are formulated in Eq. 3.4 and Eq. 3.5 (Hodges, Rolt, and Jones, 1975; Parsley and Robinson, 1982). The functions assume the association between the increment of roughness or the rut depth (ΔIRI or ΔRUT) and the cumulative traffic load (NE_t) is linear. α and β are the coefficients of the roughness increment and the rut depth increment respectively to be estimated. u and w are constants standing for the other influence. Besides, the literature also provides the models of other performance indicators, such as cracking, ravelling, and potholing. The study only has the performance data of roughness and rutting, and the relevant models are focused.

$$\Delta IRI = \alpha NE_t + u \tag{3.4}$$

$$\Delta RUT = \beta NE_t + w \tag{3.5}$$

The non-linear regression models of the roughness have the cumulative model as Eq. 3.6 and the increment model which function is Eq. 3.7 (Paterson, 1987). The cumulative model assumes the exponential relation between the age of the road section (T) and IRI, and the coefficient is attributed to the initial roughness after the construction or the repaving and traffic load (IRI_0 and NE_t). The coefficient of T and NE_t (n and α) should be estimated by the regression. The increment model supposes the different variables contribute to the roughness increment in various ways. The increments of rutting, cracking, patching, potholing and time (ΔRUT, ΔCR, ΔPAT, ΔPOT, and t) are related to the roughness increment (ΔIRI) as the linear fitting. The age of the road section (T) has the natural exponential relation with ΔIRI, and the increment of traffic load (ΔNE_t) is taken in the coefficient.

$$IRI = (IRI_0 + \alpha NE_t)e^{nT} \tag{3.6}$$

$$\Delta IRI = \alpha \Delta NE_t e^{nT} + \omega_1 \Delta RUT + \omega_2 \Delta CR + \omega_3 \Delta PAT + \omega_4 \Delta POT + \gamma IRIt \tag{3.7}$$

The non-linear regression model of the rut depth has the functions as Eq. 3.8 (Paterson, 1987). The model assumes the rut depth (RUT) is the function of the cumulative traffic load (NE_t) and the age (T) with the mathematical form in Eq. 3.8. β, x, x and w indicate the coefficient, the exponent of the variable of age and the error term respectively. The form of the cumulative model is exponential, and in Paterson's research it is has the shape in which x and y have the values between 0 and 1.

$$RUT = \beta T^x NE_t^y + w \tag{3.8}$$

To compare the applicability of different models, all the pre-defined regression models elaborated on above will be applied, expect the linear functions in Eq. 3.4 and Eq. 3.5. Because according to the correlation analysis in Chapter 5, the linear relation between the pavement performance variables of ΔIRI and ΔRUT (in the dataset named as IRI_VAR and RUT_VAR) and the traffic variables (NE_t, in the dataset, named as I_AL) is not obvious. rho^P analysed by the data of ΔIRI and NE_t in the case is 0.01, and it is -0.02 by the data of ΔRUT and NE_t. The coefficient indicates how well if the correlation is described as linear, and the closer the absolute value is to 1, the closer two variables are linearly correlated. Therefore, the pre-defined linear regression model will not be achieved. In addition, to get a more accurate result, some changes will be devised when applying the pre-defined regression models. The traffic variable in the model is ESAL (or the increment of ESAL), which is a common solution to represent the mixed traffic. With the data support in the study, the traffic load is categorised into various vehicle types. Rather than using one variable to represent the traffic influence, the study regards several variables meaning the different kinds of vehicles as the traffic factors. The model application will be in Chapter 6.

3.4. Survival model

Survival model, is also named as accelerated life model or accelerated failure time model. Its function assumes that all predictor variables can affect the lifetime of the pavement. It is a widely-applied time-related model. It is able to predict the expected survival time (T_s) of the roadways with a certain material paving layer on the surface under the conditions which are determined by traffic-associated and non-traffic associated factors (f). The function is as Eq. 3.9. The model assumes that a survival-time (T_{s0}) probability of the pavement covered by the specific material on the surface layer is a certain distribution. The survival time of a particular road section is effected by the specific situation in which it is located. The gap between the actual survival time and the baseline survival-time probability distribution is caused by the influence factors (f), and the quantitative influence of external factors evaluated in the coefficients (λ) can be described in the exponential model (Wang et al., 2005; Kartsonaki, 2016; Ebrahimi, Wallbaum, Svensson, and Gryteselv, 2019).

$$ln \frac{T_s}{T_{s0}} = \sum f_i \lambda_i \tag{3.9}$$

The process of applying the survival analysis shown in Fig. 3.2 contains 5 steps: (1) selecting the performance indicators; (2) defining the threshold values of the performance indicators; (3) modelling the survivor function; (4) estimating how the expected survival time depends on traffic characteristics; (5) calibration In Fig. 3.2, the green boxes represent the input, and the dark-blue ones are the models. Both the light blue and grey modules are the results, but the grey cubes give the values and the other shows other calculations. The first step for modeling the survival analysis replies on the pavement conditions of A15 that are evaluated by the

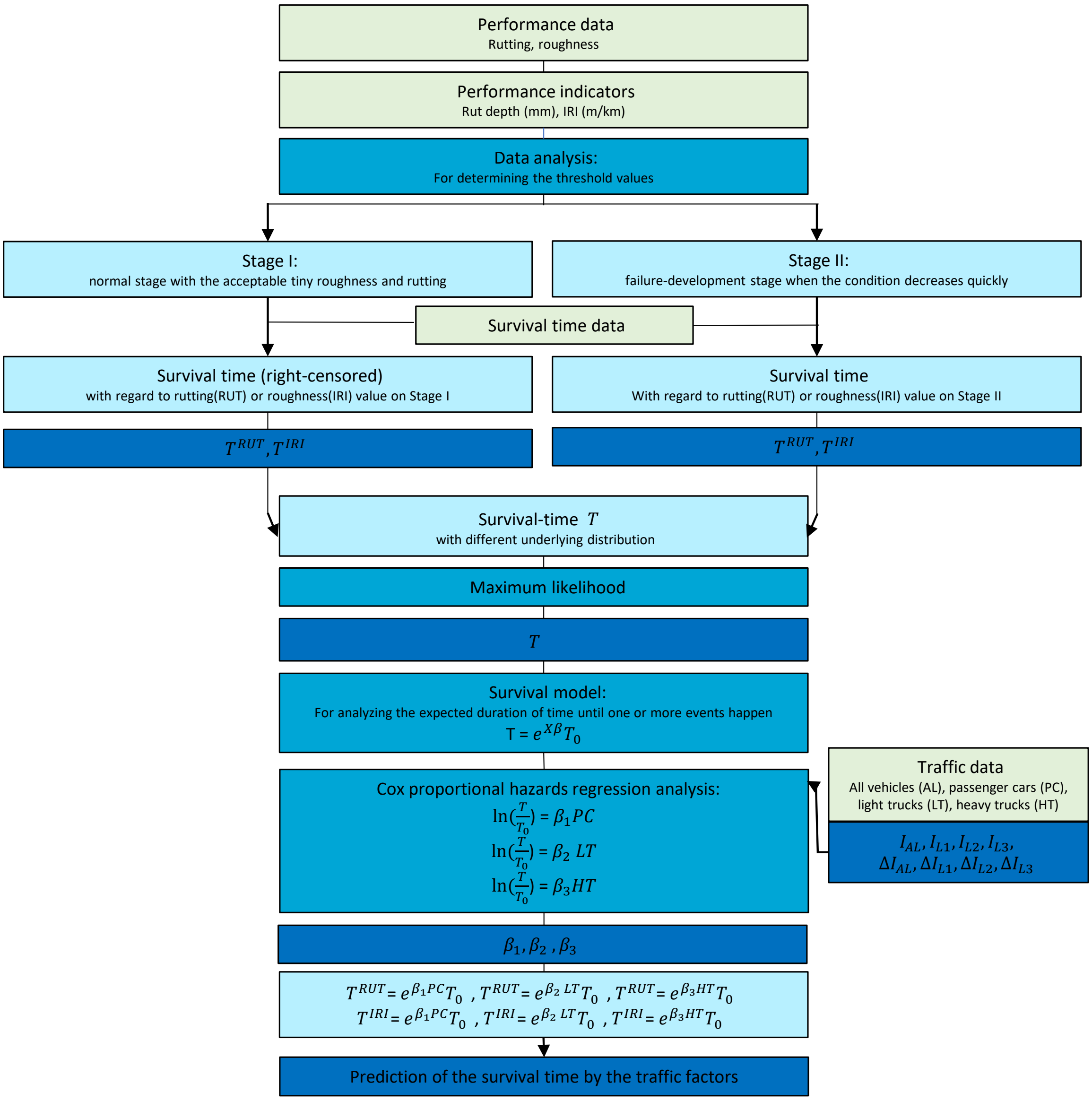


Figure 3.2: Methodology of survival model

annual measurement on the outer and median lanes. Secondly, it has been observed that the pavement deterioration would not appear for a long period, but once it occurs, it will develop dramatically in a short time, especially cracking (Wang et al., 2005). Therefore, the reasonable thresholds of the indicators are required, and they can make the resurfacing decisions standardized. The research divides the life-span into two stages: one is the normal stage with the acceptable tiny road roughness and rutting; the other is the failure-development stage when the condition decreases quickly. To find the thresholds, there are two applicable methods. The first one is employed by data analysis. Primarily, diagram the condition data and its corresponding change of next year, when the fluctuation of the line suddenly increases, it means there is a big probability of deteriorating to the next stage. The critical position can be determined by setting up several discrete ranges from continuous condition data and treating each range as a condition state, and then examining the transition probability of staying at the current state and the probability of deteriorating to one unit, two units until the maximum value difference of the indicators between the year and the next year. If the probability of staying at a certain condition state is low, it means pavement condition is unstable; and it would be appropriate to resurface the exiting pavement before its condition deteriorates to the next stage. Since it is possible that the same value of the roughness or the rut depth in different road sections have the different change in the next year, the study applies the median fitting to show the performance progression along the time. The second method to figure out the critical values is dependent on the empirical knowledge. The information is given in Section 1.6 as the damage assessment set up by RWS according to the comprehensive considerations of the impact of road roughness on serviceability and maintenance costs. The third step of setting up the survivor function is run in Matlab. With the uncertainty of the distribution of the survive time, all the classic distributions used in the previous literature will be tried, including normal distribution, Weibull distribution, logistic distribution, gamma distribution, log-logistic distribution, and lognormal distribution. The best-fitting form is selected by the maximum log likelihood. After defining the survivor function, the model then estimates how the expected failure time depends on traffic characteristics and other factors. It is a classic exponential regression approach, and r-squared and p-value are used to evaluate how goodness of fitting the model is. The last step of the modelling is the calibration for the censored data problem. There is the difficulty of unobserved failure events in a typical set of pavement condition data, because data collection surveys are typical for the limited duration. A few pavements have been already reached a serious unevenness on the first survey, most are not observed the failure events during the whole survey. If only the unevenness initiation events observed during the survey were included in a statistical analysis, important information about the stochastic and mechanistic properties of the phenomenon coming from the "before" and "after" events may be excluded, and thus cause a bias in the model. The most common solution to the censored data problem is introducing the maximum likelihood estimator, which is not applicable to the study because the censored data is made up of about 90 % of all the test data. The details will be discussed in Chapter 7.

3.5. Decision tree classifier

Decision tree classifier is a classic and fast classification model (Kaur & Garg, 2014). The classic machine learning algorithm has a big advantage that it can be easily interpreted as the natural "if..., then..., else..." structure compared to other machine learning methods, and it will be applied to the study. The methodology has been successfully used in multiple diverse areas, due to its capability to break down the complex decision-making process into the multistage simple decisions. It is a predictive model that represents a mapping between the object properties and the object values. The graphical method for intuitive use is the key of the analysis, and the output is a tree structure in which each internal node represents a test on an attribute, each branch represents a test output, and each leaf node represents a category. It is a kind of supervisory learning which requires the input of a bunch of samples, each of which has a set of attributes and a category. The categories should be determined in advance. Then by the learning process, a classifier has been set up by the segmentation of the source database and the data test. The process is to trim the tree recursively and when no further segmentation or a separate class can be applied to a branch, the recursion process is completed. The tree-shape output is able to predict the correct classification by the new attributes (Safavian & Landgrebe, 1991).

The approach of the model application in the study is shown in Fig. 3.3 The study sets up the model in Python with Scikit-learn package. There are two phases regarding the framework of the methodology in Fig. 3.3, the training phase and the predicting phase. The training phase as the first modeling application step has the input of the training data, the labels, and the categories. Based on these inputs, the learning method is trained and the tree structure is set up that will be used in the next phase, the predicting phase. Apart from the learning model formulated by the training process, the test data, as well as the labels are inputted to the second phase. The categories are the prediction results.

There are several requirements of the input data and information for the model:

- Both the training data and the test data are real numbers. Either continuous values or discrete values are acceptable.
- If a label is a category variable, it must be transferred as the discrete variables. For example, the data on SURFACE_LAYER of A15 is Special Pavement Surface Layer (COMBID),Close Asphalt Concrete (DAB) , Emulsion Asphalt Concrete (EAB), Open Asphalt Concrete (OAB), Special Asphalt Concrete (SMA) , ZOAB, New Zeer Open Asphalt Concrete (ZOAB+), Double-layer Zeer Open Asphalt Concrete (ZOABTW), or Zeer Open Emulsion Asphalt Concrete (ZOEAB). It can only be taken account in the model by the variables of SURFACE_COMBID, SURFACE_DAB, SURFACE_EAB, SURFACE_OAB, SURFACE_SMA, SURFACE_ZOAB, SURFACE_ZOAB+, SURFACE_ZOAB+, and SURFACE_ZOEAB with the bi-

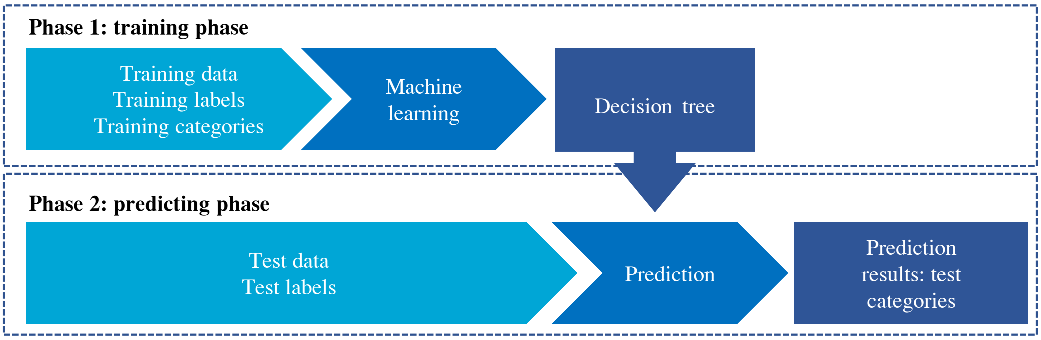


Figure 3.3: Modelling application process of the decision tree classifier

nary values of 1 or 0. 1 indicates the surface material of the road section is the type, while 0 means not.

- If a label is a date variable, the model automatically recognizes it as a digital form. Based on the common setting of the computer system, the date of 1st January, 1900 in the digital form is 1, the date of 2nd January, 1900 in the digital form is 2, and etc.

The learning method plays a key role in the model. The machine learning algorithm of the decision tree in the study has the functional forms that are not necessarily determined beforehand. The method can efficiently deal with large, complicated datasets without imposing a complicated parametric structure. Gini index is used as the splitting criterion. It indicates how good a split is by calculating the mix of classes in the two separated groups that can be created by the specific split. It favors large partitions of data. The worst split results in the gini index of 0.5 and the best gives the value of 0. Minimizing the index leads to the best split on each branch. The formula is 1 minus the quadratic sum of all the probabilities (p_i) of each class i in the given branch. The probabilities of a class in the branches are computed by the number of training samples (S) out of all that falls in the branch.

$$gini = 1 - \sum_{i=1}^{S} p_i^2 \tag{3.10}$$

3.6. Conclusion

The chapter defines the research methodology, which contains the model framework, the correlation analysis, and the performance models. The three kinds of performance models are all applicable to predict the pavement performance in theory. They can quantify the effects of various influence factors, in particular, the transport variables, on the road performance in different ways. This chapter finds the answer to sub question 3 as the following elaborates.

The thesis is on the foundation of the model framework proposed in the chapter. The correlation analysis, the regression models, the survival model, and the decision tree classifier defined in the chapter will all be applied in the remaining part of the research.

Sub question 3

What performance models are feasible to perform in this case and can achieve the research goal?

There are many types of data analysis models for road performance, which can be roughly divided into three categories: regression models, probability models, and classification models. Previous studies of the performance models have applied various forms of the three kinds of models. Each kind has the different assumption on the effects of traffic on road performance. The regression models assume that traffic load are the cause of the pavement performance progression, meaning the amount of traffic per unit can lead to a certain degree of the change in the road roughness or rutting. The probability models suppose that traffic flow is the reason for reducing the service life of the road, indicating the amount of traffic per unit can increase the likelihood that the road meets the maintenance requirements in advance. The classification models suggest that the relationship between traffic volumes and road performance may be in the form of transitions. That is to say, the traffic volume may affect the road characteristics only when it reaches a remarkable level or a critical value. In order to get a comprehensive and objective study result, the research selects at least one model in each category to quantify the effects of traffic flow on roughness and rutting. Based on the data that can be obtained in the study, five regression models are selected (three of which are predictive models of roughness, and two are for rutting), the survival model representing the probability model, and decision tree classifier, a classic classification model, as the research models.

4

Data Input

Chapter 3 defined the methodology and the model framework of the thesis. The framework as Fig.3.1 illustrates is used as a lens to look at the collaboration between the factors quantitatively during the service life of roadways. This chapter prepares data for the quantitative analysis of the relation between the pavement performance and the factors. Section 4.1 gives the introduction of the supportive database from RWS and DiTTlab. Various categories of data from the database of RWS and DiTTlab will be introduced in Section 4.2. The dependent and independent variables will be defined and formulated in Section 4.3. The chapter addresses the research question: SQ 4. Which pavement deterioration, its potential causes and traffic characteristics can be captured in the current data collection system? Which can not? The answer will be provided in Section 4.4.

4.1. Data source

Since 1987, Dutch Study Center for Road Construction (SCW) has established the visual inspection system for road performance evaluation (Stichting Studie Centrum Wegenbouw, 1987). The survey records were transferred to and joined in WINFRABASE. The joint database is expected to include the data of IRI, rut depth, SI, cracking, and ravelling every year (Rijkswaterstaat, 2006). But it is still under the construction. The current database contains the annual measurements of IRI and rut depth per 100 meters of all the highways in the Netherlands since 2009, except IRI in 2010. It also includes the yearly records of SI per 100 meter sections of all the Dutch freeways since 2012, expect the year 2016. As for ravelling and cracking, they were measured by 3D laser triangulation. The model for calculating the raveling values is being established, and for now only the raw data is accessible. The progress of cracking is that same as the raveling (Rijkswaterstaat, 2017b).

INWEVA stands for Intensities Wegvakken (Intensities of road sections, in English). Traffic intensity is measured on approximately 3,000 road sections. The other road sections are estimated on the basis of the traffic model. The intensities are measured on all lanes of the national road network in the Netherlands, including the on-ramps and off-ramps, the multiple lanes and the interchanges. The data gives the annual average traffic flow per road section of three vehicle classes. The vehicle classification in Database of Traffic Intensities of National Road Network in the Netherlands (INWEVA) is in Tab. 4.1. It is further divided into the weekdays, the weekends and the rush hours. Furthermore, the data is subdivided into morning, evening and night periods. The files are put into the different folders, listed in Tab. 4.2 (Rijkswaterstaat, 2012).

Table 4.1: Vehicle classification in INWEVA system (Rijkswaterstaat, 2012)

Classification	Description	Definition
L1	Passenger cars	Length less than 5.6 m
L2	Light trucks	Length between 5.6m and 12.2m
L3	Heavy trucks	Length longer than 12.2m

Table 4.2: Table of content in INWEVA system (Rijkswaterstaat, 2012)

Folder	Content	File format
A	Traffic intensities	.xlsx; Gis
B	Traffic intensities	.pdf
C	Various checks of the the traffic model results	.doc
D	Results of the samples	.xlsx
E	Methods and results of the independent testing	.doc; .xlsx
F	Updates of the road section numbers	.xlsx
G	Records of the processing and the response from the various departments in RWS	.doc
H	Traffic intensity visualisation	web
I	Regional flags	.png
J	Quality reports and methodology description	.doc

The other data source is DiTTlab. The research laboratory founded by TU Delft and CGI is set up for the transport data analysis and traffic simulation. With the support of DiTTlab, the following data that is relevant to the study is accessible:

- Speeds and flow from the loop detectors per minute on every 100-meter section of most of the highways, the distributor roads, and the access roads in the Netherlands since 25th July, 2016. They have being updated continuously.
- Average acceleration or deceleration, vehicle loss hours, and travel time derived from the data of speeds and flow
- The heated maps of the road segments derived from the data of speeds and flow
- Data from the dynamic speed limit system
- The network graph underneath with properties, like the number of lanes, RD Coordinates, etc.
- Weather data collected by all the weather stations in the Netherlands, including the temperature, the wind, and the precipitation every minute since 1st January, 2015. They have been updating continuously.
- Reports of the incidents and the accidents

4.2. Raw data

Aimed at the evaluation of the service quality, the measurement of the pavement performance is carried out once per year on average. The results have been recorded in WINFRABASE since 2009. It consists of rut depth, longitudinal flatness, and skid resistance. The article by Rijkswaterstaat (2017b) determines the measurement methods of the three performance indicators, and they are described respectively below.

Rut depth is the relative height difference between the central sensor and the sensors on both sides. The measurement is according to Dutch norm NEN-EN 13036-8: 2008 until 2018. It defines the rut depth as the average lane depth per 100-meter section. The measurement method firstly surveys the depth of a segment length of 2000 mm. The lane depth must be determined per lane between the length markings. Then, an average value per hectare track is determined per lane. The highest value is picked between both the sets. The indicator has the unit of mm.

Longitudinal flatness is indicated as IRI. According to Dutch norm NEN-EN 13036-5, it is the roughness index over 100 meters. The longitudinal flatness must be determined in the left and right driving lanes. An average value per hectare section is determined per lane, and the average value of all lanes is the indicator value of the hectare meter expressed in m/km. The change from IRI to HRI took place in 2017. The main difference is the averaging length.

SI is the assessment of the skid resistance of a road pavement surface by the measurement of the sideway-force coefficient. This is calculated per hectare on the basis of a weighted average of skid resistance. It is

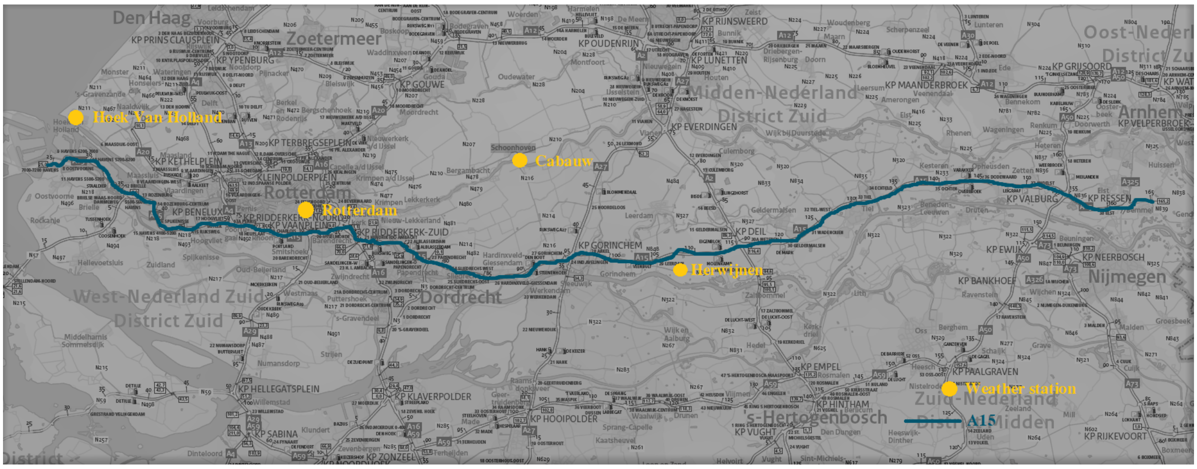


Figure 4.1: Map of A15 and weather stations

expressed as the difference from the standard value. For example, SI as 0.03 means the weighted average skid resistance is 0.03 higher than the norm value. SI is used in MJPV. There is an exception of the skid resistance measurement presented in Rijkswaterstaat (2017a). When the road maintenance is written in DBFM, the skid resistance is assessed by AS. It is calculated per hectare based on the average of two measurements taken immediately after each other. It is expressed as the difference with the norm values. Generally, SI is used in the most situations in a more cost-effective way, and AS can be equivalent assumed as SI. It should be noticed that the measurement method had the change in year 2016 and it was proved that the results did not meet the experts' expectation thus in year 2017, it was changed back.

Some facts of the performance indicators should be noticed. Rut depth were measured according to Dutch norm NEN-EN 13036-8: 2008 until 2018, and from the year of 2019 on, the measurement method is adjusted to apply for European norm. Before 2018, RWS estimated the longitudinal flatness of the roads by IRI, but from 2018 on, it uses HRI. For skid resistance, the distress model is not used in the maintenance planning system, and the engineers have found that the measurements are inaccurate partly. Therefore, considering the comparability and accuracy of the data, data of IRI from 2009 to 2018, and data of rut depth from 2009 to 2017 can be considered to use.

Besides the performance indicators, WINFRABASE includes the information of the road structure and the road network, as explained in Section 1.6. It also has the data from INWEVA system. It provides the traffic intensities since 2012 and updates annually. This study will use the data of the construction date, the surface layer materials, and the average annual traffic intensities of the road segments from the database.

Accordingly, the traffic and climate factors in the framework in 3.2 will be considered in the study. INWEVA gives the yearly-based traffic data, while DiTTlab provides the minute-based data. Due to the annual measurements of the pavement performance indicators, the yearly-based traffic data is used for figuring out the quantitative effects of the transportation on the road performance. In order to set up the real-time simulation of the pavement performance progression, the minute-based data is applicable. The climate data used in the study comes from the database of DiTTlab. To conclude, the available data and the unavailable data are distinguished in Fig. 4.2.

A prime criterion for selecting the test sections used in the thesis is the availability of sufficient historical distress data. The secondary is meeting both the academic interest and the practical interest of RWS. A15 (highway from Oostvoorne to Bemmel), with a total length of 140.2 km, including 75 segments listed in Appendix A, is selected in the project as the test roadway. The road network is shown in Fig. 4.1. The performance data in the past 10 years, the traffic data from the year 2012 on, and the climate data since 2015 from four weather stations near to A15 will be tested in the study. For all test sections, five types of data were collected: construction data, performance data, time data, traffic data, and climate data as listed in Tab. 4.3.

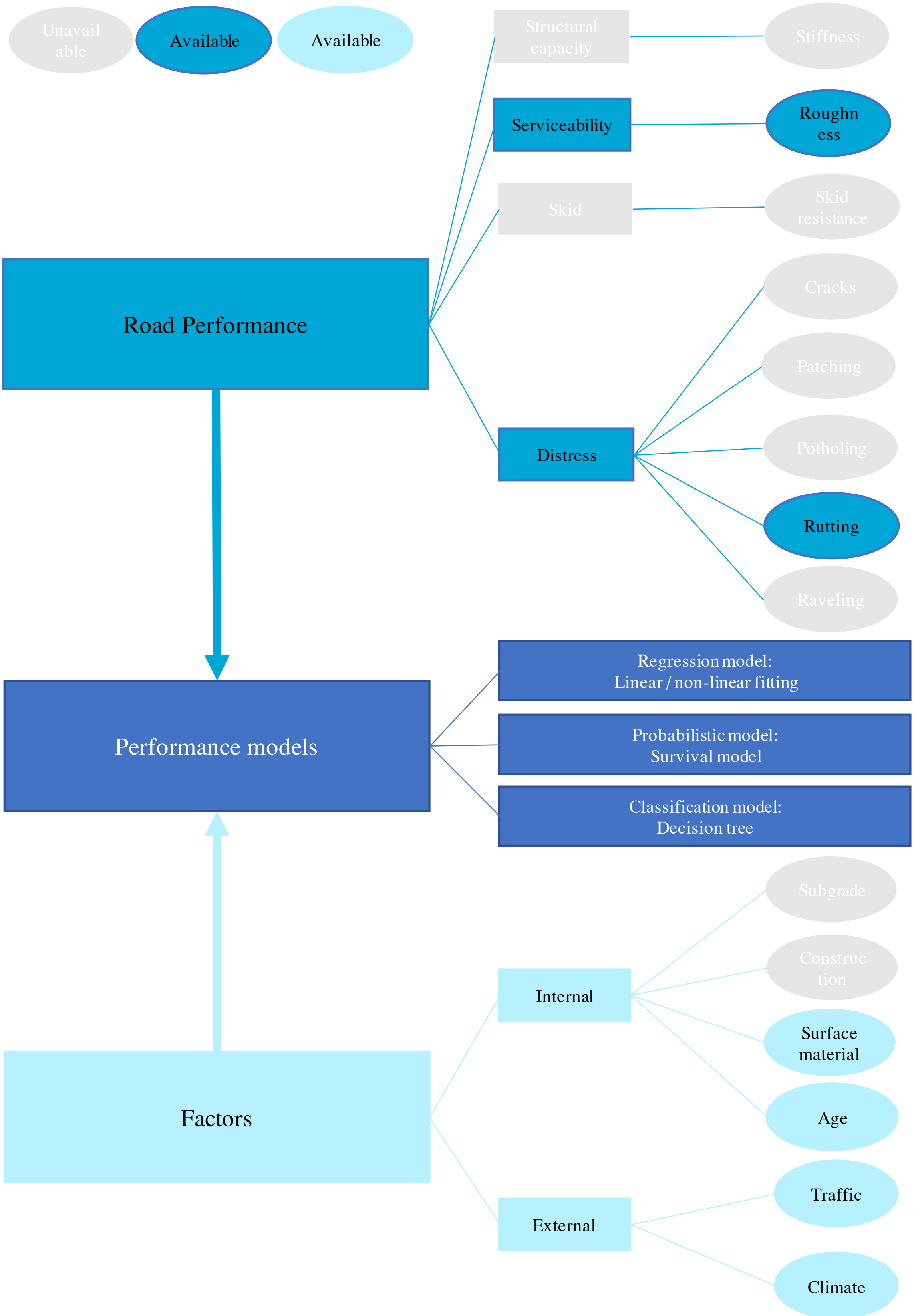


Figure 4.2: The available and unavailable data in the framework

Table 4.3: The input data

Data category	Label	Description
Construction	ROAD	Road name
	DIRECTION	Road direction
	LANE	Surveyed lane
	FROM_KM	Location of the start of the road segment
	TO_KM	Location of the end of the road segment
	CONSTR_DATE	Construction date
	SURFACE_LAYER	Surface layer material
Performance	IRI_VALUE	IRI by the survey
	RUT_VALUE	Rutting by the survey
Time	DATE_IRI	Date of the survey of IRI
	DATE_RUT	Date of the survey of rutting
Traffic	I_AL_Y	Average daily intensities of all the lanes in the survey year
	I_L1_Y	Average daily intensities of passenger cars on all the lanes in the survey year
	I_L2_Y	Average daily intensities of light trucks on all the lanes in the survey year
	I_L3_Y	Average daily intensities of heavy trucks on all the lanes in the survey year
Climate	STATION	Weather station
	TEMP	Temperature at the survey time
	PRECIPITATION	Precipitation at the survey time
	T_STA	Start time of the climate survey
	T_END	End time of the climate survey

4.3. Data process

The data gained from the databases cannot be inputted directly to the data analysis because of four problems. Firstly, The measurement data for each year was independently recorded in different tables. Therefore, it lacks the links. Secondly, the raw data set does not cover all the variables in the performance models, and some variables should be the derivatives thereof. Thirdly, the scales of various data types are different. There are three scales in the data sets, which are the regions, the big road sections, and the small sections. Four weather stations along the way divide A15 into four regions in Appendix D. The big road sections are defined as the roadways between the adjacent ramps, and the freeways have 75 in total, as listed in Appendix A. The small sections are defined as the 100-meter length, and there are 1402 small sections on A15. Weather data is recorded per region and traffic data is per big section, while construction, performance and time data is recorded per small road sections. Fourthly, some data on the road performance, the traffic, and the climate is missing. To solve the problems and get the values of the variables, the four steps of proceeding data are as follows. Due to the big size of the data, all the data processing is run by Matlab. Some sample codes are written in Appendix B, Appendix C, and Appendix E.

Step 1: Integrate the data sets from the different tables
As stated above, the tables of data are independent. The key of the solution is to find the corresponding label in the different tables. Both the databases from RWS and DiTTlab use the same location labels of RD Coordinates, and accordingly, the tables are integrated. One sample of the Matlab code of integrating the annual measurements of IRI from the different tables and plotting the progression is in Appendix B.

Step 2: Unify the unit of data of all the categories as per 100 meters
The problem of the different scales of the data will be solved by unifying the scale as the small road section, which is 100-meter long. The program starts with the original table of the road performance data which is per small section, and then creates the new columns and fills in the traffic intensities and climate data. It is obvious that a large number of small sections have the same traffic and climate data, when they are located in the same big road sections and the same regions.

Step 3: Fill in the missing data

Missing data is hard to avoid due to some instability in the collection system. In the databases, the performance, traffic, and climate data are missing to the different extents. Instead of focusing on the improvement of the data collection, the study solves the problem by filling some data in the blank cells according to some assumptions. The assumptions are made regarding the causes and the features of the missing data. The data of the roughness and the rut depth is missing mostly because the road was under construction during the measurement time. Thus, it is acceptable to leave the cells blank. The minute traffic data is supposed to include the traffic flow of all the 100-meter sections on A15 from the first minute on 25th July, 2016 which was the start time of data collection system to the last minute on 14th October, 2019 which is defined as the end time of the project. But due to some instability of the loop detectors, some data in the early collecting time was not accessed. To fill in the blank cells, the thesis assumes the traffic flow on the section where the traffic data is missing is the same as the flow on the nearest section. So if the traffic data on some sections is missing, it is filled in the values of the nearest section. Additionally, the weather data is expected to contain wind speeds, the temperature and the precipitation of every minute at the 4 weather stations from 1st January, 2015 which was the first date of the climate data collection system, to the last minute of 14th October, 2019 as the end time of the project. However, in fact, there are two occasions of the weather data missing. One is where the system did not work during the period from 1st August, 2015 to 31st December, 2015. The other is where some data is missing in the whole collecting period. The research has the assumption that the weather trend in the non-measurement time is the same as the measurement time. For example, the total time of the measurement in 2016 by the Cabauw station was 140.8 hours. Among them, there were 2.6 hours when the temperature was above 25 °C, 129.1 hours between 0 °C and 25 °C, and 9.1 hours below 0 °C. The rest 8619.2 hours were not measured. According to the assumption, the total duration when the temperature was above 25 °C is calculated as 2.6 * (140.8 + 8619.2) / 140.8 = 161.8 in Appendix E. The formulations of the other variables in the climate categories are applied to the same idea.

Step 4: Formulate the variables with the input data
The variables are selected by two criteria in the study: they should be included in the performance models in Section 2.4; they should be gotten or derived from the raw data in Tab. 4.3. Accordingly, the selection results are listed in Tab. 4.4. The formulations of all the variables are very simple, and based on general knowledge, in Tab. 4.5. However, there is difficulty in the calculations. The main challenge occurs when computing the cumulative intensities. Due to the different construction times of the road sections, the calculations of the starting year vary. To resolve the problem, this study initially divides the whole service lives into years, and then uses the average annual daily traffic intensities of every year times the service duration of that year and sum all up to the cumulative traffic intensities. A sample of coding the cumulative intensities is in Appendix C.

As a consequence, the complete set of data {construction, distress, time, traffic, and climate} is suitable for analysis. The ability to formulate this (time duration) variable was primarily due to the existence of a significant panel data set; when a pavement provides the bad service regarding roughness or rutting during the period of observation, it is assumed that it met the maintenance requirements at the measurement time. When the pavement section stays good until the end of the observation period, it is considered as right-censored ; similarly, when the pavement section has a high value of IRI or the rut depth at the first measurement, it means the bad condition exists before the initiation of the observation period, and it is considered as left-censored. Explanatory variables are external factors that are assumed to influence the performance and deterioration of the pavements; the data on explanatory variables were divided into four data groups: construction, time, traffic, and climate. Dependent variables are the performance indicators of the sections, and the data on the dependent variables are in the performance group in Tab. 4.4. Additionally, the input data is so big, especially some sections last more lane 10 years, the traffic and weather data contain the records of every minute in years. To improve the calculation efficiency, Matlab is used and the formulations are coded. All the variables are treated as the matrices in the program. However, the first step of reading the raw data from excel tables is the most time-consuming part in the total program. The solution to the problem is code a file that can read the excel tables and store in the format of mat (a kind of Matlab data). By this, it only requires reading the excel file once, and afterwards, if inputting the data, the program can call the mat file, which is much quick than reading the excel file. But due to the data process is not the key of the study, the report does not provide a excel-mat transforming code.

4.4. Conclusion

To conclude, the chapter investigates the data availability and sets up the feasible analytical framework for the study in Fig. 4.2. The formulas and the computation process of the variables including multiple Matlab codes play the key role in the chapter. The results will be used in the following parts, such as the correlation analysis in Chapter 5, the regression model application in Chapter 6, the survival model application in Chapter 7, and the decision tree model application in Chapter 8. Additionally, the answer to the sub question 4 is found and will be elaborated as follows.

Sub question 4

Which pavement deterioration, its potential causes and traffic characteristics can be captured in the current data collection system? Which can not?

As for the pavement deterioration, data of IRI and the rut depth is accessible from the databases of WINFRABASE. The measurement of skid index is not all accurate, so it is not chosen to study in the research. Cracking and raveling were measured by the 3D laser measurement tools, and in the system only the outputs of the tools were recorded. The raw data requires an additional data-processing program to compute the performance indices to study. Other performance indicators are not even measured.

Regarding the influencing factors, surface materials, the total service time of the road sections, traffic intensities of three vehicle classes, temperature and precipitation are all available, but design, subgrade types, and construction are not.

The traffic data collection system, which the study has the access to, known as INWEVA, contains the annual daily average traffic flow on every road section in the whole national road network in the Netherlands of three vehicle classes, the annual average traffic intensities of any vehicle type during the morning, evening or night period, the annual traffic intensities of any vehicle type during the weekdays, weekends or rush hours, and the percentage of heavy load traffic. In addition to the traffic intensities, WINFRABASE captures the relevant information of the locations in RD Coordinates and the directions of the traffic flow. The other data source from DiTTlab provides traffic speeds and traffic flow per minute on every 100-meter section of most roadways in the Netherlands, average acceleration or deceleration, vehicle loss hours, travel time of a route, traffic controls, roadway maps, and numbers of lanes.

The report will continue with the correlation analysis which will describe the correlations between the road performance and the factors. Those will be indicated as three correlation coefficients that has been defined before. That will be the end of Define phase.

Table 4.4: List of variables

Variable category	Label	Description
Construction variables	ROAD	Road name
	DIRECTION	Road direction
	LANE	Surveyed lane
	FROM_KM	Location of the start of the road segment
	TO_KM	Location of the end of the road segment
	CONSTR_DATE	Construction date
	SURFACE_LAYER	Surface layer material
Performance variables	IRI_VALUE_0	IRI at the last survey
	RUT_VALUE_0	Rutting at the last survey
	IRI_VALUE	IRI at the this survey
	RUT_VALUE	Rutting at the this survey
	IRI_VAR	Variance of IRI between the last survey and this one
	RUT_VAR	Variance of rutting between the last survey and this one
	IRI_CLASS	Class of IRI. 1 and 2 are acceptable; 3 and 4 are bad
	RUT_CLASS	Class of rutting. 1 and 2 are acceptable; 3 and 4 are bad
Time variables	AGE_IRI	Time between the construction and the survey of roughness
	AGE_RUT	Time between the construction and the survey of rutting
	DATE_IRI_0	Date of the last survey of IRI
	DATE_RUT_0	Date of the last survey of rutting
	T_IRI	Duration between this survey of IRI and the last one
	T_RUT	Duration between this survey of rutting and the last one
	DATE_IRI	Date of this survey of IRI
	DATE_RUT	Date of this survey of rutting
Traffic variables	I_AL	Total intensities from the first service day until the survey
	I_L1	Total number of passenger cars from the first service day until the survey
	I_L2	Total number of light trucks from the first service day until the survey
	I_L3	Total number of heavy truck from the first service day until the survey
	I_AL_INC	Number of vehicles from the last survey until the survey
	I_L1_INC	Number of passenger cars from the last survey until the survey
	I_L2_INC	Number of light trucks from the last survey until the survey
	I_L3_INC	Number of heavy trucks from the last survey until the survey
Climate variables	T_TEMP_25	Total time of the survey year when the temperature above 25°C, and the unit is hour
	T_TEMP_0	Total time of the survey year when the temperature between 0°C and 25°C, and the unit is hour
	T_TEMP_0_below	Total time of the survey year when the temperature below 0°C, and the unit is hour
	T_PRECIPITATION	Total time of rains in the survey year, and the unit is hour
Discrete variables	R_IRI_2015	Whether IRI meets the maintenance requirement for the first time in the 2015 survey. 1 is yes; 0 is no.
	R_IRI_2016	Whether IRI meets the maintenance requirement for the first time in the 2016 survey. 1 is yes; 0 is no.
	R_IRI_2017	Whether IRI meets the maintenance requirement for the first time in the 2017 survey. 1 is yes; 0 is no.
	R_IRI_2018	Whether IRI meets the maintenance requirement for the first time in the 2018 survey. 1 is yes; 0 is no.
	R_RUT_2015	Whether rutting meets the maintenance requirement for the first time in the 2015 survey. 1 is yes; 0 is no.

(To be continued)

Variable category	Label	Description
Discrete variables	R_RUT_2016	Whether rutting meets the maintenance requirement for the first time in the 2016 survey. 1 is yes; 0 is no.
	R_RUT_2017	Whether rutting meets the maintenance requirement for the first time in the 2017 survey. 1 is yes; 0 is no.
	R_RUT_2018	Whether rutting meets the maintenance requirement for the first time in the 2018 survey. 1 is yes; 0 is no.
	SURFACE_COMBID	Whether the surface layer material is COMBID or not. 1 is yes; 0 is no.
	SURFACE_DAB	Whether the surface layer material is DAB or not. 1 is yes; 0 is no.
	SURFACE_EAB	Whether the surface layer material is EAB or not. 1 is yes; 0 is no.
	SURFACE_OAB	Whether the surface layer material is OAB or not. 1 is yes; 0 is no.
	SURFACE_SMA	Whether the surface layer material is SMA or not. 1 is yes; 0 is no.
	SURFACE_ZOAB	Whether the surface layer material is ZOAB or not. 1 is yes; 0 is no.
	SURFACE_ZOAB+	Whether the surface layer material is ZOAB+ or not. 1 is yes; 0 is no.
	SURFACE_ZOABTW	Whether the surface layer material is ZOABTW or not. 1 is yes; 0 is no.
	SURFACE_ZOEAB	Whether the surface layer material is ZOEAB or not. 1 is yes; 0 is no.

Table 4.5: Formulations of the variables

Variable category	Label	Formulation
Construction variables	ROAD	Raw data
	DIRECTION	Raw data
	LANE	Raw data
	FROM_KM	Raw data
	TO_KM	Raw data
	CONSTR_DATE	Raw data
	SURFACE_LAYER	Raw data
Performance variables	IRI_VALUE_0	Raw data
	RUT_VALUE_0	Raw data
	IRI_VALUE	Raw data
	RUT_VALUE	Raw data
	IRI_VAR	IRI_VALUE - IRI_VALUE_0
	RUT_VAR	RUT_VALUE - RUT_VALUE_0
	IRI_CLASS	Derivation from IRI_VALUE, and the damage assessment
	RUT_CLASS	Derivation from RUT_VALUE, and the damage assessment
Time variables	AGE_IRI	DATE_IRI - CONSTR_DATE
	AGE_RUT	DATE_RUT - CONSTR_DATE
	DATE_IRI_0	Raw data
	DATE_RUT_0	Raw data
	T_IRI	DATE_IRI - DATE_IRI_0
	T_RUT	DATE_RUT - DATE_RUT_0
	DATE_IRI	Raw data
	DATE_RUT	Raw data
Traffic variables	I_AL	$\sum_{1}^{AGE} I\_AL\_Y$

(To be continued)

Variable category	Label	Formulation
Traffic variables	I_L1	$\sum_{1}^{AGE} I\_L1\_Y$
	I_L2	$\sum_{1}^{AGE} I\_L2\_Y$
	I_L3	$\sum_{1}^{AGE} I\_L3\_Y$
	I_AL_INC	$\sum_{1}^{T} I\_AL\_Y$
	I_L1_INC	$\sum_{1}^{T} I\_L1\_Y$
	I_L2_INC	$\sum_{1}^{T} I\_L2\_Y$
	I_L3_INC	$\sum_{1}^{T} I\_L3\_Y$
Climate variables	T_TEMP_25	Derivation from TEMP, and T_STA or T_END
	T_TEMP_0	Derivation from TEMP, and T_STA or T_END
	T_TEMP_0_below	Derivation from TEMP, and T_STA or T_END
	T_PRECIPITATION	Derivation from PRECIPITATION, and T_STA or T_END
Discrete variables	R_IRI_2015	Derivation from IRI_VALUE, and the damage assessment
	R_IRI_2016	Derivation from IRI_VALUE, and the damage assessment
	R_IRI_2017	Derivation from IRI_VALUE, and the damage assessment
	R_IRI_2018	Derivation from IRI_VALUE, and the damage assessment
	R_RUT_2015	Derivation from RUT_VALUE, and the damage assessment
	R_RUT_2016	Derivation from RUT_VALUE, and the damage assessment
	R_RUT_2017	Derivation from RUT_VALUE, and the damage assessment
	R_RUT_2018	Derivation from RUT_VALUE, and the damage assessment
	SURFACE_COMBID	Derivation from SURFACE_LAYER
	SURFACE_DAB	Derivation from SURFACE_LAYER
	SURFACE_EAB	Derivation from SURFACE_LAYER
	SURFACE_OAB	Derivation from SURFACE_LAYER
	SURFACE_SMA	Derivation from SURFACE_LAYER
	SURFACE_ZOAB	Derivation from SURFACE_LAYER
	SURFACE_ZOAB+	Derivation from SURFACE_LAYER
	SURFACE_ZOABTW	Derivation from SURFACE_LAYER
	SURFACE_ZOEAB	Derivation from SURFACE_LAYER

5

Correlation Analysis

Chapter 4 focuses on the availability of data and the computation of the variables. This chapter explores the strength of the relations between all the variables, including the road performance, the traffic, climate, and time factors. The statistics methods elaborated in Section 3.2 are used to analyse, and the values of Pearson's linear correlation coefficient, Kendall's tau coefficient and Spearman's rho describe the remarkableness of correlation between two testing variables. This correlation analysis has the result of how related the traffic variables that are captured in the current data collection system are to the pavement deterioration progression.

Before applying three correlation analytic methods formulated in Section 3.2, it is crucial to understand what types of data the methods can be analyzed. Among the three correlation analysis methods, Pearson's linear correlation coefficient has the most critical rules for the test data. Because rho^P is the ratio of the covariance to the standard deviation, the denominator cannot be 0, which means that the two variables have the continuous values. Spearman's method is based on the sorting position of the original data, which loosen the requirement for the input data that should have different values. If there is an abnormal number (extremely large or small) in the data, because rho^S reveals the rank correlation, its impact is far less in rho^S than in rho^P. The data that Kendall's method can analyze is the categorical variable that has the orders. In order to simplify the process of the correlation analysis, the study satisfies the strictest data requirements of the three correlation analysis and input the data table for the three correlation analysis, which is the analysis variables have continuous and various values. Thus the variables that do not have the features listed in Tab. 4.4 are not inputted for the correlation analysis, including all the construction variables (ROAD, DIRECTION, LANE, FROM_KM, TO_KM, CONSTR_DATE, and SURFACE_LAYER), 4 time variables (DATE_IRI_0, DATE_RUT_0, DATE_IRI, DATE_RUT), all the discrete variables, and IRI_CLASS and RUT_CLASS. The correlation analysis applies all the performance variables (IRI_VALUE, RUT_VALUE, IRI_VAR, and RUT_VAR, IRI_VALUE_0, RUT_VALUE_0), the four time variables (AGE_IRI, AGE_RUT, T_IRI, T_RUT), the traffic variables (I_AL, I_L1, I_L2, I_L3, I_AL_INC, I_L1_INC, I_L2_INC, I_L3_INC), and the climate variables (T_TEMP_25, T_TEMP_0, T_TEMP_0_below, T_PRECIPITATION).

It should be noticed that the time variables are the special factors. Theoretically, if the external factors, like traffic and climate can be all excluded, the effects of the time itself on the road performance is the process of oxidation, and as long as the new asphalt is laid down, the material is exposed to oxygen inevitably and the oxidation is beginning until the end of the lifespan of road sections. Although it is likely to result in the brittle asphalt, with significant loss of elasticity, bringing about a significant increase in the probability of failure, under the natural conditions (typically air is composed of 21% by volume of oxygen), it takes a very long time to cause changes in material properties (Asphalt Institute, 2003). Therefore, in the study, if the strong correlations between the time variables and the road performance are found, it does not indicate that the time itself is related to the development of the pavement performance, but means the combination of all the time-related variables both captured in the database, like the cumulative traffic intensities and the cumulative time of the specific weather conditions, and non-measured factors correlate with the road performance. Besides, in the survival model, the time variables are the dependent ones, which are defined as the duration from the construction time or the last repaving time until the road segments meet the maintenance requirements regarding the performance indices, in other words, the lifespans of the road sections. The meaning of the lifetime differs from the time variables input in the correlation analysis in this chapter. In the test data, about 90% on A15 road segments have the good service capability, and the time variables (AGE_IRI and AGE_RUT) counts the first operation day until the measurement date of the specific performance indicators (IRI and the rut depth in the study).

Formulated in Chapter 3.2, three correlation analytic methods are coded in Matlab in Appendix F. The chapter focuses on the dependence and independence of the variables, deriving from three correlation coefficients. It is necessary to clarify the correlation coefficients only indicate the correlation between the two variables in all current historical data for the A15 segments. The three correlation coefficients used in the report can only indicate the linear correlation and the rank correlation between the variables. So even if the three correlation coefficients are nearly 0, it does not mean that the two variables are absolutely irrelevant, but indicate the weak linear or rank correlations of two variables. The thesis cannot exclude the possibility that there are some other correlations between the two variables. Based on three correlation coefficients, it is able to determine the non-independence of variables. In statistics, the concept of irrelevance is less critical than independence. It is uncertain two variables are independent if three correlation coefficients of them are small, but it is definite that a big value of any correlation coefficient among the three indicates the non-independence.

In addition, this study cannot exclude the possibility that the data correlation cannot be observed on other roadways or rare data is incorrectly recorded in the database. In particular, there might be some occasions that the records were not updated during the construction time but afterwards, resulting a lag and inaccuracy in the database. The situation can result in the records of a fresh performance with all kinds of the factors in the rather long service time, like heavy cumulative traffic intensities. If it is the case, the actual correlations of transportation intensities and the road performance are most likely closer than the results of the data analysis.

5.1. Dependence of the variables

The section focuses on the linear correlation and the rank correlation between variables. It explores the dependence between the variables to some extent. The results of Pearson's linear correlation coefficients in Tab. 5.1 indicate how closeness to a line the correlation between two variables is. The closer the absolute value is to 1, the closer the two variables are to linear correlation. According to the results, the variable that has the strongest linear correlation with the roughness is the values of unevenness in the previous year ($rho^P = 0.78$). That is to say, approximately, if a road section in the first year has the high value of the surface unevenness, it will have a high one in the second year too, and if a low value, it will be about a small number as well. This can lead to a finding that the annual variance of roughness is a small amount. But that does not mean that the amount of variation is negligible, because even a tiny increase of the roughness of a road section which originally is close to the critical value can meet the maintenance requirement concerning the roughness. The development of the rut depth has a similar result, that is, the rut depth in the second year is near the first-year value. The evidence is that the linear correlation coefficient computed by the variables of RUT_VALUE and RUT_VALUE_0 is 0.68 that is close to 1. In addition, in the correlation analysis of A15 data, the roughness and the rut depth are a little bit linearly correlated and the coefficient is positive meaning a road section with a deep rut and a large unevenness at the same time. It is easy to understand by the definition of roughness and rutting. Road roughness is the comprehensive indicator, and a large rut depth certainly affect the profile flatness. The result supports the form of Paterson's performance model for roughness which considers all kinds of the distress as the independent variables, including cracking, rutting, patching, and potholing (Paterson, 1987). Only from the linear correlation coefficient computed by the data of A15, the linear correlation between traffic cumulative volumes of the various vehicle types and IRI is weak, and as well as the climate factors and IRI, which are the variables of the cumulative time when the temperature is above 25 °C, the cumulative time when the temperature is between 0 °C and 25 °C, and the duration when the temperature is below 0 °C in the study. The linear correlations between the roughness and other variables are not obvious by the analysis, where the absolute values of the coefficients are below 0.1.

Compared to the linear correlation of traffic flow and the roughness on A15, the linear relationship between traffic flow and the rutting of the roadway is a bit more obvious. rho^P formulated by IRI and the cumulative numbers of passenger vehicles, light trucks, and heavy trucks are 0.11, 0.13 and 0.14 respectively, while the coefficients by the rut depth and the traffic flow of the three categories are 0.18, 0.23 and 0.25 respectively. It is worth noting that in the three types of vehicles, the linear correlation coefficient of the numbers of the heavy trucks and the rutting is the largest, the same as the roughness. It is not to say that the heavy trucks have a greater impact on the development of the rut depth or the surface unevenness than the other vehicle types, but indicates that the number of heavy trucks is more linearly related to the rut depth or the roughness compared with the number of the other vehicle types.

Besides, according to the analysis of the correlation between the temperature variables and the rutting data on A15, the characteristics of linearity is a little more apparent compared to the relation between the temperature variables and the roughness. The absolute values of rho^P between the temperature variables and the rutting have the range from 0.31 to 0.38, while the coefficients of the temperature variables and IRI have the absolute values between 0.15 to 0.19.

The variances of IRI are most linearly related to the corresponding time (rho^P = 0.49) and the total service time of the road segment (rho^P = 0.3). This does not mean that the time itself is related to the development of the pavement performance, but indicates the combination of all the time-related variables both captured in the database and non-measured factors has a correlation with the road performance. The most linearly-related variable of the variance of the rut depth is the rut depth in the last measurement (rho^P = 0.46) by the data analysis of A15. In summary, the closeness of the linear correlations between the dependent and independent variables are shown in Fig. 5.1. The wider the lines in the figure, the more close the relations between the two variables linked are to the linear form.

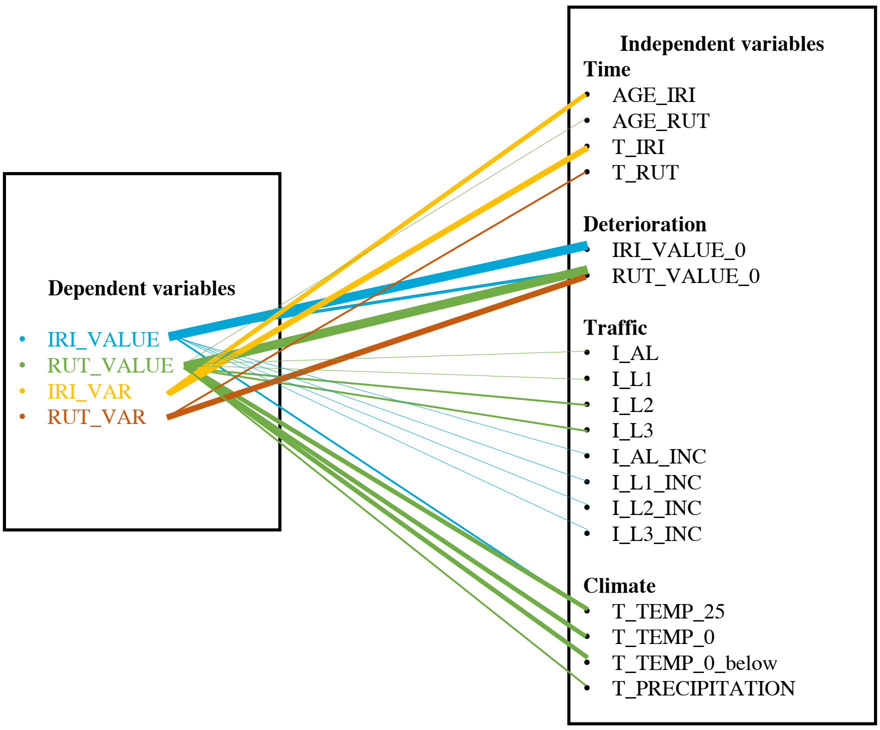


Figure 5.1: Closeness of the linear correlation between the variables analyzed by A15 data

When all the input data are the continuous variables, the analytic results of Kendall's tau coefficient and Spearman's rho are basically similar, because both are to work out the ordinal association between the test two quantities. The more the coefficients approach 1, the more similarity of the orderings of the data that is ranked in each of the quantities. In contrast, the closer the coefficient is to -1, the closer the rank correlation of the two variables is to fully inverse. If the ranking is completely independent, the coefficients will be 0. According to the results of Kendall's tau coefficient in Tab. 5.2 and Spearman's rho in Tab. 5.3, the rank correlation of the input variables is found. The closest the ranking of the data to the variable of IRI of this year is the IRI data in the last measurement of all A15 segments (rho^K = 0.58 and rho^S = 0.75). In other words, based on the historical data, the ordering of a road segment concerning its roughness among all the A15 sections does not have a big change in the measurement and the following one. The time intervals between twice measurement are approximately one year in the study. There is the similar situation with regard to the rut depth. The order of the rut depth of a certain pavement section in all A15 sections is close to that in the second year Another finding is that in the historical data of the A15, the weather condition (in the study, represented by the four variables: the cumulative time over 25 °C throughout the year, the cumulative time between 0 and 25 °C throughout the year, and the cumulative time below 0 °C , as well as the total rain time throughout the year, has the rank correlation with the rut depth to some extent. The total time when the temperature between 0 and 25 °C is negatively associated with the rut depth by their respective rankings. The absolute values of Kendall's tau coefficients between the four weather variables and the rut depth range from 0.23 to 0.31, while those of Spearman's rhos are between 0.28 to 0.37. As for the rank correlation between the traffic intensities of three classes and the pavement performance (IRI and rutting in the research), the correlation analysis of A15 data finds the stronger ordinal association between them and the rutting than that of the traffic intensities and the roughness. Although among the three vehicle categories (passenger vehicles, light trucks, and heavy trucks), the largest correlation coefficients are computed by the variables of the rut depth and the total number of the light trucks (rho^K= 0.20 and rho^S=0.27), the coefficients of the traffic flow of three vehicle classes and the rut depth are very close.

Additionally, a common question is whether the values of the correlation coefficient computed respectively by two different variables and the same variable indicate which variable affects the variable more. For example, if the correlation coefficient between variable A and variable B is 0.7 and that of variable A and variable C is 0.3, does it mean that variable B has a greater influence on variable A than C? The answer is not necessarily. Because, firstly, the correlation coefficient describes the correlation of the data on the two test variables, neither the cause and effect of the variables nor the amount of the change of one variable caused by the change of the other one. Secondly, the correlation coefficient is to indicate the strength of a kind of association between variables. It is likely that a strong correlation between the two variables of the other form, but the correlation coefficient applied in the study is very low. For instance, x and y have a strong relationship as xy=1, but the linear correlation coefficient is low. The comparison of the correlation coefficients by one variable and the different variables can indicate that, according to the input data of the study, the correlation between variable A and variable B is greater than that of variable A and variable C.

5.2. Independence of the variables

Apart from the linear correlations and the rank correlations between the input variables, the chapter focuses on the independence between the variables. As stated in Chapter 2, the pavement performance indicators and the influencing factors are formed in a complex and complicated situation. They can be (in)dependent or (un)correlated from each other. It should be critical to consider the (in)dependence of the variables when setting up the performance model, especially in the linear form. Since Pearson's linear correlation coefficient, Kendall's tau coefficient, and Spearman's rho can only indicate the linear correlations and the rank correlations, so the other forms of the association cannot be derived from in the chapter, and the small values of three correlation coefficients of two test variables do not mean that the two variables are absolutely irrelevant, but only show the weak correlation in the linear form or by ordering. But according to the results of rho^P, rho^K, and rho^S, it is able to determine the non-independence between the variables. If any correlation coefficient of two variables have the value close to or equal to 1, the non-independence between them is definite. Therefore, there are 7 non-independent pair of the variables which at least one of the absolute value of rho^P, rho^K, and rho^S is above 0.98 as follows. In the performance modelling application in Develop phase, the variables of those pairs will not show in one model at the same time.

- AGE_IRI & AGE_RUT
- T_IRI & T_RUT
- I_AL & I_L1
- I_AL_INC & I_L1_INC
- T_TEMP_25 & T_TEMP_0
- T_TEMP_25 & T_TEMP_0_below

- T_TEMP_0 & T_TEMP_0_below

The analytic results of non-independence in the last paragraph make sense. The non-independence of the time variables, the pair of AGE_IRI and AGE_RUT, and the pair of T_IRI and T_RUT, is due to the fact that the measurements of the unevenness and the rut depth of a road section on A15 usually was carried out on the same day. According to the historical data of A15, only a very few measurements of the performance characteristics in some years on some sections were taken on different dates. Thus, the data on the time variables formulated by the measurement time of the roughness is almost identical with that derived from the measurement time of the rut depth on the same pavement segment. In Develop phase, the application of the performance models concerning the roughness uses data of AGE_IRI and T_IRI as the independent variables, and the predictive models for the rut depth have the time-associated independent variables as AGE_RUT and T_RUT.

As for the traffic variables, AL (the sum of all the vehicle classes) and L1 (the passenger vehicles) has high dependence. In fact, by the definition, these two equations are always true: (1) I_AL = I_L1 + I_L2 + I_L3; (2) I_AL_INC = I_L1_INC + I_L2_INC + I_L3_INC. That is to say, those traffic variables in the equations are dependent. The passenger cars are the majority of the vehicles so that among the correlation coefficients of the various traffic variables, the pairs of I_AL and I_L1 and I_AL_INC and I_L1_INC have the highest values. Accordingly, in the model application, either the total traffic intensities (I_AL) or the total traffic intensity increments (I_AL_INC) is used, or the total numbers of the three vehicle types (I_L1, I_L2, and I_L3) or the increment numbers of the three vehicle types (I_L1_INC, I_L2_INC, and I_L3_INC) are used. Because one of the study goal is to figure out the effects of the various vehicles on the road performance, the research applies the variables of each class of the vehicles in Chapter 6, Chapter 7, and Chapter 8 rather than the sum.

The non-independence has been observed between the climate variables as well. By the definition, there is an equation of these variables: T_TEMP_25 + T_TEMP_0 + T_TEMP_0_below = T, where T is the constant representing 365 or 366 days. So it is definite that the variables are dependent. It is worth noting that the rank correlations of the three weather variables are 1 or -1, meaning any location of A15 has the same position in the ranking of all the locations on A15 with regards to the high-temperature time (above 25 °C) as in the ranking of the low-temperature time (below 0 °C) among all places on the roadway. The rankings are just inverse of the one concerning the normal-temperature time (between 0 and 25 °C) Out of the sheer coincidence, the study only considers data from four weather stations near A15. In other words, only four data of a weather variable are sorted, and the number of so-called "all the locations on A15" as phrased before is only four. Although the test road section where has higher temperature also has lower temperature, it is because the size of the study regions. Supposing the study looks at the weather in the past decades all over the Netherlands by inputting the data of a large number of the weather stations throughout the country, the calculation results of rho^K or rho^S must be not equal to 1 or -1, and the coincidence can be avoided. In summary, the coincidence occurs mainly due to the limited inputting weather data.

According to the literature study in Chapter 2, the effects of the influence factors are not simple. For example, only the factor of time has little impact on the road performance. But once in hot conditions and under the slow-moving heavy load for hours, the asphalt tends to behave like viscous liquid, while in cold climates with rapid loading for a while, the material has a tendency to behave as elastic solids (Asphalt Institute, 2003). Since the main focus of the study is to figure out the influence of the transport flow on the pavement performance, one reasonable solution is comparing the road performance of various sections where have the identical conditions except the traffic flow. The solution has some feasible issues because at first, the study cannot capture all kinds of factors. As stated in Chapter 4, the design, the construction, and the upgrade are not recorded in the database that supports the research, and the climate data is accessible of every region but the difference between the specific locations in a region is assumed as none. Secondly, if selecting the road sections that have all the other variables captured in the current data collection system but different traffic flow, there are very few comparable road segments in this way. The root reason is as for A15 basically only the adjacent parts are constructed on the same date, and the traffic flow are the same if there is no ramp between them. That is to say, once the construction time is the controlling variable, it is highly possible to have the same value of the traffic flow and the research on the influence of the traffic intensities cannot achieve. In summary, although the experiment method of controlling variables is a common way of studying the effects of a single variable on a dependent variable, its effectiveness is limited under the study conditions.

5.3. Conclusion

To conclude, the chapter computes three common correlation coefficients, and shows the linear and rank correlations between the variables that have the continuous data in the collection system in numbers. The dependence and independence of the variables are derived from the coefficients, and the analytic results will be used to set up the performance models in the next phase.

According to the correlation analysis, the traffic variables that are captured in the data collection system relate to the pavement deterioration progression. Derived from Pearson's linear correlation coefficient, based on A15 data from 2015 to 2018, the analysis results show that the traffic flow and the road roughness are slightly-weak linear correlated, and the traffic flow and the rut depth are weak linear correlated. The range of Pearson's linear correlation coefficients between the variables of traffic flow and the roughness is from 0.11 to 0.14, while the coefficients of every kind of the traffic-flow variables and the rut depth are between 0.18 and 0.25. According to Kendall's tau coefficients and Spearman's rhos, the data analysis of A15 data from 2015 to 2018 indicates that the traffic flow and the road roughness have slightly-weak rank correlation, and the traffic flow and the rutting have weak rank correlation. The range of Spearman's rhos computed by the variables of the traffic flow and the road roughness is 0.15 to 0.17, and the range of Kendall's tau coefficients of the traffic flow and the road roughness is 0.10 to 0.11. The range of Spearman's rhos of the variables of the traffic flow and the rut depth is 0.23 to 0.27, and Kendall's tau coefficients of the traffic flow and the rut depth have the range of 0.17 to 0.20. The study does not test other forms of data association, except linear correlation and rank correlation. So it cannot exclude the possibility that transport flow is strongly related to road roughness and rutting in other kinds of relation. The test data only contains A15 data from 2015 to 2018, so the results are not accountable for the relation of the transport flow and the road roughness or the rutting on other roadways during other observation time. Besides, there might be occasions that the records were not updated during the construction time but afterwards, resulting a lag and inaccuracy in the database. The situation can result in the records of a fresh performance with all kinds of the factors in the rather long service time, like heavy cumulative traffic intensities. If it is the case, the actual correlation of the transportation intensities and the road performance is most likely closer than the results the data analysis finds.

The chapter is the end of Define phase. The report will continue with the modelling application for predicting the pavement performance by the factors that can be captured in the current data collection system in the next phase.

Table 5.1: Pearson's linear correlation coefficients

	AGE _IRI	AGE _RUT	IRI_ VALUE _0	RUT_ VALUE _0	T_IRI	T_RUT	I_AL	I_L1	I_L2	I_L3	I_AL _INC	I_L1 _INC	I_L2 _INC	I_L3 _INC	IRI_ VALUE	RUT_ VALUE	IRI _VAR	RUT _VAR	T_ TEMP _25	T_ TEMP _0	T_ TEMP _0_ below	T_ PRECI-PITA-TION
AGE _IRI	1.00	0.99	-0.03	0.10	0.53	0.58	0.24	0.24	0.25	0.26	0.10	0.09	0.16	0.13	0.05	0.18	0.30	0.23	0.11	-0.15	0.15	0.04
AGE _RUT	0.99	1.00	-0.03	0.10	0.51	0.58	0.25	0.25	0.26	0.27	0.10	0.10	0.17	0.13	0.05	0.19	0.29	0.23	0.11	-0.15	0.15	0.04
IRI_ VALUE _0	-0.03	-0.03	1.00	0.19	-0.05	-0.05	0.11	0.11	0.14	0.12	0.10	0.09	0.11	0.08	0.78	0.24	-0.07	0.05	0.21	-0.17	0.16	-0.11
RUT_ VALUE _0	0.10	0.10	0.19	1.00	0.00	0.02	0.22	0.21	0.24	0.26	0.02	0.02	0.08	0.05	0.23	0.68	0.11	-0.32	0.33	-0.35	0.35	0.04
T_IRI	0.53	0.51	-0.05	0.00	1.00	0.88	-0.02	-0.02	-0.01	-0.01	-0.12	-0.12	-0.11	-0.10	0.06	-0.03	0.49	0.23	-0.01	0.01	-0.01	-0.01
T_RUT	0.58	0.58	-0.05	0.02	0.88	1.00	-0.01	-0.01	0.00	0.00	-0.09	-0.09	-0.07	-0.07	0.02	-0.02	0.42	0.25	0.01	-0.01	0.01	-0.01
I_AL	0.24	0.25	0.11	0.22	-0.02	-0.01	1.00	1.00	0.95	0.94	0.38	0.37	0.42	0.28	0.11	0.19	0.01	-0.02	0.35	-0.35	0.35	0.02
I_L1	0.24	0.25	0.11	0.21	-0.02	-0.01	1.00	1.00	0.94	0.93	0.39	0.39	0.42	0.27	0.11	0.18	0.01	-0.03	0.34	-0.35	0.34	0.02
I_L2	0.25	0.26	0.14	0.24	-0.01	0.00	0.95	0.94	1.00	0.94	0.29	0.28	0.38	0.26	0.13	0.23	0.01	0.00	0.36	-0.36	0.35	-0.03
I_L3	0.26	0.27	0.12	0.26	-0.01	0.00	0.94	0.93	0.94	1.00	0.30	0.28	0.40	0.33	0.14	0.25	0.05	0.00	0.34	-0.36	0.36	0.04
I_AL _INC	0.10	0.10	0.10	0.02	-0.12	-0.09	0.38	0.39	0.29	0.30	1.00	1.00	0.89	0.81	0.13	0.00	0.02	-0.04	0.28	-0.14	0.10	-0.30
I_L1 _INC	0.09	0.10	0.09	0.02	-0.12	-0.09	0.37	0.39	0.28	0.28	1.00	1.00	0.86	0.77	0.13	0.00	0.02	-0.04	0.28	-0.14	0.11	-0.28
I_L2 _INC	0.16	0.17	0.11	0.08	-0.11	-0.07	0.42	0.42	0.38	0.40	0.89	0.86	1.00	0.89	0.13	0.04	0.01	-0.05	0.34	-0.21	0.18	-0.32
I_L3 _INC	0.13	0.13	0.08	0.05	-0.10	-0.07	0.28	0.27	0.26	0.33	0.81	0.77	0.89	1.00	0.14	-0.02	0.05	-0.09	0.18	-0.01	-0.02	-0.41
IRI_ VALUE	0.05	0.05	0.78	0.23	0.06	0.02	0.11	0.11	0.13	0.14	0.13	0.13	0.13	0.14	1.00	0.26	0.57	0.06	0.19	-0.16	0.15	-0.09
RUT_ VALUE	0.18	0.19	0.24	0.68	-0.03	-0.02	0.19	0.18	0.23	0.25	0.00	0.00	0.04	-0.02	0.26	1.00	0.10	0.46	0.31	-0.37	0.38	0.13
IRI _VAR	0.30	0.29	-0.07	0.11	0.49	0.42	0.01	0.01	0.01	0.05	0.02	0.02	0.01	0.05	0.57	0.10	1.00	0.13	0.03	-0.03	0.03	0.00
RUT _VAR	0.23	0.23	0.05	-0.32	0.23	0.25	-0.02	-0.03	0.00	0.00	-0.04	-0.04	-0.05	-0.09	0.06	0.46	0.13	1.00	0.01	-0.06	0.07	0.12
T_ TEMP _25	0.11	0.11	0.21	0.33	-0.01	0.01	0.35	0.34	0.36	0.34	0.28	0.28	0.34	0.18	0.19	0.31	0.03	0.01	1.00	-0.93	0.90	-0.26
T_ TEMP _0	-0.15	-0.15	-0.17	-0.35	0.01	-0.01	-0.35	-0.35	-0.36	-0.36	-0.14	-0.14	-0.21	-0.01	-0.16	-0.37	-0.03	-0.06	-0.93	1.00	-1.00	-0.09
T_ TEMP _0_ below	0.15	0.15	0.16	0.35	-0.01	0.01	0.35	0.34	0.35	0.36	0.10	0.11	0.18	-0.02	0.15	0.38	0.03	0.07	0.90	-1.00	1.00	0.17
T_ PRECI-PITA-TION	0.04	0.04	-0.11	0.04	-0.01	-0.01	0.02	0.02	-0.03	0.04	-0.30	-0.28	-0.32	-0.41	-0.09	0.13	0.00	0.12	-0.26	-0.09	0.17	1.00

Table 5.2: Kendall's tau coefficients

	AGE _IRI	AGE _RUT	IRI_ VALUE _0	RUT_ VALUE _0	T_IRI	T_RUT	I_AL	I_L1	I_L2	I_L3	I_AL _INC	I_L1 _INC	I_L2 _INC	I_L3 _INC	IRI_ VALUE	RUT_ VALUE	IRI _VAR	RUT _VAR	T_ TEMP _25	T_ TEMP _0	T_ TEMP _0_ below	T_ PRECI- PITA- TION
AGE _IRI	1.00	1.00	0.06	0.19	-0.24	-0.23	0.76	0.75	0.79	0.76	0.15	0.15	0.21	0.13	0.07	0.20	0.06	0.00	0.32	-0.32	0.32	0.22
AGE _RUT	1.00	1.00	0.06	0.19	-0.24	-0.23	0.76	0.75	0.79	0.76	0.16	0.15	0.21	0.13	0.07	0.20	0.06	0.00	0.32	-0.32	0.32	0.22
IRI_ VALUE _0	0.06	0.06	1.00	0.10	-0.08	-0.07	0.09	0.09	0.10	0.10	0.10	0.09	0.11	0.08	0.58	0.16	-0.02	0.07	0.13	-0.13	0.13	-0.04
RUT_ VALUE _0	0.19	0.19	0.10	1.00	-0.15	-0.14	0.19	0.18	0.20	0.19	0.05	0.05	0.10	0.08	0.13	0.47	0.08	-0.25	0.27	-0.27	0.27	0.13
T_IRI	-0.24	-0.24	-0.08	-0.15	1.00	0.99	-0.16	-0.16	-0.21	-0.15	0.30	0.30	0.15	0.33	0.00	-0.28	0.12	-0.10	-0.49	0.49	-0.49	-0.51
T_RUT	-0.23	-0.23	-0.07	-0.14	0.99	1.00	-0.16	-0.16	-0.21	-0.15	0.30	0.30	0.15	0.33	0.00	-0.28	0.12	-0.10	-0.49	0.49	-0.49	-0.51
I_AL	0.76	0.76	0.09	0.19	-0.16	-0.16	1.00	0.97	0.90	0.85	0.36	0.35	0.38	0.25	0.10	0.17	0.06	-0.03	0.27	-0.27	0.27	0.10
I_L1	0.75	0.75	0.09	0.18	-0.16	-0.16	0.97	1.00	0.88	0.82	0.36	0.36	0.38	0.24	0.10	0.17	0.06	-0.03	0.27	-0.27	0.27	0.10
I_L2	0.79	0.79	0.10	0.20	-0.21	-0.21	0.90	0.88	1.00	0.87	0.30	0.29	0.37	0.23	0.11	0.20	0.06	-0.02	0.31	-0.31	0.31	0.16
I_L3	0.76	0.76	0.10	0.19	-0.15	-0.15	0.85	0.82	0.87	1.00	0.31	0.30	0.38	0.31	0.11	0.19	0.08	-0.03	0.30	-0.30	0.30	0.11
I_AL _INC	0.15	0.16	0.10	0.05	0.30	0.30	0.36	0.36	0.30	0.31	1.00	0.96	0.73	0.59	0.12	0.02	0.08	-0.03	-0.02	0.02	-0.02	-0.32
I_L1 _INC	0.15	0.15	0.09	0.05	0.30	0.30	0.35	0.36	0.29	0.30	0.96	1.00	0.70	0.56	0.12	0.02	0.09	-0.03	-0.03	0.03	-0.03	-0.31
I_L2 _INC	0.21	0.21	0.11	0.10	0.15	0.15	0.38	0.38	0.37	0.38	0.73	0.70	1.00	0.69	0.13	0.06	0.08	-0.05	0.13	-0.13	0.13	-0.26
I_L3 _INC	0.13	0.13	0.08	0.08	0.33	0.33	0.25	0.24	0.23	0.31	0.59	0.56	0.69	1.00	0.12	-0.01	0.11	-0.09	0.03	-0.03	0.03	-0.36
IRI_ VALUE	0.07	0.07	0.58	0.13	0.00	0.00	0.10	0.10	0.11	0.11	0.12	0.12	0.13	0.12	1.00	0.17	0.42	0.05	0.10	-0.10	0.10	-0.04
RUT_ VALUE	0.20	0.20	0.16	0.47	-0.28	-0.28	0.17	0.17	0.20	0.19	0.02	0.02	0.06	-0.01	0.17	1.00	0.09	0.33	0.31	-0.31	0.31	0.23
IRI _VAR	0.06	0.06	-0.02	0.08	0.12	0.12	0.06	0.06	0.06	0.08	0.08	0.09	0.08	0.11	0.42	0.09	1.00	0.01	0.03	-0.03	0.03	0.02
RUT _VAR	0.00	0.00	0.07	-0.25	-0.10	-0.10	-0.03	-0.03	-0.02	-0.03	-0.03	-0.03	-0.05	-0.09	0.05	0.33	0.01	1.00	0.04	-0.04	0.04	0.10
T_ TEMP _25	0.32	0.32	0.13	0.27	-0.49	-0.49	0.27	0.27	0.31	0.30	-0.02	-0.03	0.13	0.03	0.10	0.31	0.03	0.04	1.00	-1.00	1.00	0.41
T_ TEMP _0	-0.32	-0.32	-0.13	-0.27	0.49	0.49	-0.27	-0.27	-0.31	-0.30	0.02	0.03	-0.13	-0.03	-0.10	-0.31	-0.03	-0.04	-1.00	1.00	-1.00	-0.41
T_ TEMP _0_ below	0.32	0.32	0.13	0.27	-0.49	-0.49	0.27	0.27	0.31	0.30	-0.02	-0.03	0.13	0.03	0.10	0.31	0.03	0.04	1.00	-1.00	1.00	0.41
T_ PRECI- PITA- TION	0.22	0.22	-0.04	0.13	-0.51	-0.51	0.10	0.10	0.16	0.11	-0.32	-0.31	-0.26	-0.36	-0.04	0.23	0.02	0.10	0.41	-0.41	0.41	1.00

Table 5.3: Spearman's rho

	AGE _IRI	AGE _RUT	IRI_ VALUE _0	RUT_ VALUE _0	T_IRI	T_RUT	I_AL	I_L1	I_L2	I_L3	I_AL _INC	I_L1 _INC	I_L2 _INC	I_L3 _INC	IRI_ VALUE	RUT_ VALUE	IRI _VAR	RUT _VAR	T_ TEMP _25	T_ TEMP _0	T_ TEMP _0_ below	T_ PRECI-PITA-TION
AGE _IRI	1.00	1.00	0.09	0.27	-0.28	-0.28	0.90	0.89	0.91	0.88	0.22	0.22	0.29	0.16	0.10	0.26	0.08	0.01	0.41	-0.41	0.41	0.28
AGE _RUT	1.00	1.00	0.09	0.27	-0.28	-0.28	0.90	0.89	0.91	0.88	0.22	0.22	0.29	0.16	0.10	0.26	0.08	0.00	0.41	-0.41	0.41	0.28
IRI_ VALUE _0	0.09	0.09	1.00	0.16	-0.10	-0.10	0.13	0.13	0.15	0.14	0.14	0.14	0.16	0.11	0.75	0.23	-0.03	0.10	0.17	-0.17	0.17	-0.06
RUT_ VALUE _0	0.27	0.27	0.16	1.00	-0.20	-0.20	0.27	0.26	0.29	0.28	0.08	0.07	0.15	0.11	0.19	0.60	0.12	-0.36	0.35	-0.35	0.35	0.17
T_IRI	-0.28	-0.28	-0.10	-0.20	1.00	0.99	-0.21	-0.21	-0.27	-0.18	0.38	0.38	0.21	0.47	0.00	-0.36	0.18	-0.14	-0.61	0.61	-0.61	-0.63
T_RUT	-0.28	-0.28	-0.10	-0.20	0.99	1.00	-0.21	-0.21	-0.27	-0.18	0.38	0.38	0.21	0.47	0.00	-0.36	0.17	-0.14	-0.61	0.61	-0.61	-0.63
I_AL	0.90	0.90	0.13	0.27	-0.21	-0.21	1.00	1.00	0.99	0.97	0.48	0.48	0.53	0.36	0.15	0.24	0.08	-0.05	0.36	-0.36	0.36	0.13
I_L1	0.89	0.89	0.13	0.26	-0.21	-0.21	1.00	1.00	0.98	0.96	0.49	0.49	0.53	0.35	0.15	0.23	0.08	-0.05	0.35	-0.35	0.35	0.13
I_L2	0.91	0.91	0.15	0.29	-0.27	-0.27	0.99	0.98	1.00	0.97	0.42	0.41	0.51	0.33	0.16	0.27	0.08	-0.03	0.40	-0.40	0.40	0.20
I_L3	0.88	0.88	0.14	0.28	-0.18	-0.18	0.97	0.96	0.97	1.00	0.44	0.43	0.53	0.43	0.17	0.25	0.11	-0.04	0.38	-0.38	0.38	0.13
I_AL _INC	0.22	0.22	0.14	0.08	0.38	0.38	0.48	0.49	0.42	0.44	1.00	1.00	0.89	0.77	0.18	0.03	0.13	-0.04	-0.03	0.03	-0.03	-0.42
I_L1 _INC	0.22	0.22	0.14	0.07	0.38	0.38	0.48	0.49	0.41	0.43	1.00	1.00	0.87	0.74	0.18	0.02	0.13	-0.04	-0.04	0.04	-0.04	-0.42
I_L2 _INC	0.29	0.29	0.16	0.15	0.21	0.21	0.53	0.53	0.51	0.53	0.89	0.87	1.00	0.86	0.19	0.09	0.12	-0.07	0.17	-0.17	0.17	-0.34
I_L3 _INC	0.16	0.16	0.11	0.11	0.47	0.47	0.36	0.35	0.33	0.43	0.77	0.74	0.86	1.00	0.18	-0.01	0.16	-0.13	0.04	-0.04	0.04	-0.46
IRI_ VALUE	0.10	0.10	0.75	0.19	0.00	0.00	0.15	0.15	0.16	0.17	0.18	0.18	0.19	0.18	1.00	0.24	0.58	0.07	0.14	-0.14	0.14	-0.05
RUTT_ VALUE	0.26	0.26	0.23	0.60	-0.36	-0.36	0.24	0.23	0.27	0.25	0.03	0.02	0.09	-0.01	0.24	1.00	0.13	0.45	0.37	-0.37	0.37	0.28
IRI _VAR	0.08	0.08	-0.03	0.12	0.18	0.17	0.08	0.08	0.08	0.11	0.13	0.13	0.12	0.16	0.58	0.13	1.00	0.02	0.03	-0.03	0.03	0.02
RUT _VAR	0.01	0.00	0.10	-0.36	-0.14	-0.14	-0.05	-0.05	-0.03	-0.04	-0.04	-0.04	-0.07	-0.13	0.07	0.45	0.02	1.00	0.05	-0.05	0.05	0.13
T_ TEMP _25	0.41	0.41	0.17	0.35	-0.61	-0.61	0.36	0.35	0.40	0.38	-0.03	-0.04	0.17	0.04	0.14	0.37	0.03	0.05	1.00	-1.00	1.00	0.32
T_ TEMP _0	-0.41	-0.41	-0.17	-0.35	0.61	0.61	-0.36	-0.35	-0.40	-0.38	0.03	0.04	-0.17	-0.04	-0.14	-0.37	-0.03	-0.05	-1.00	1.00	-1.00	-0.32
T_ TEMP _0_ below	0.41	0.41	0.17	0.35	-0.61	-0.61	0.36	0.35	0.40	0.38	-0.03	-0.04	0.17	0.04	0.14	0.37	0.03	0.05	1.00	-1.00	1.00	0.32
T_ PRECI-PITA-TION	0.28	0.28	-0.06	0.17	-0.63	-0.63	0.13	0.13	0.20	0.13	-0.42	-0.42	-0.34	-0.46	-0.05	0.28	0.02	0.13	0.32	-0.32	0.32	1.00

III

Develop

According to the previous part, three kinds of well-applied models are chosen, which are the regression models, the survival model, and the decision tree classifier model. And the variables both dependent and independent are defined and computed from the test data.

In this phase, the performance models are applied to quantify the effects of various influencing factors according to the test data. However, not all the variables in the pre-defined models are accessible in the study, as stated in Chapter 4. Therefore, some assumptions are made in the study. The results of the model application will be used in Deliver phase to design a simulation tool for improving road management. The model application is based on the methodology of the pre-defined performance models. And this phase applied the models to the case study of A15. It consists the model application of three kinds of models. The main research question can be answered mainly by the study of this phase.

6
Regression Models

The chapter is the modeling application of the regression models, that has been defined in Chapter 3. The application uses the collected data of the entire roadways of A15 during years from 2015 to 2018, and the results indicate the effects of the transport flow on the road performance quantitatively in terms of the regression forms. The linear regression models for predicting roughness and rutting in Section 6.1, and the non-linear results are put in Section 6.2. At last, two sub-questions are answered concerning the regression models, which are how to apply the regression models, and what are the results of the regression models by the test data.

6.1. Linear regression models

The section focuses on two linear regression models: one for the road roughness, and the other about the rutting, defined in Chapter 3.3. To avoid the comprehensive confusion that the interspersed introduction may bring, the two models will be elaborated separately. The roughness-relevant one takes the first place and the linear regression model of rutting follows. The elaboration on each model contains the setting of the model, the application of the model and the model results, and they will be explained in this order.

According to Chapter 3.3, the linear regression model of the roughness is in the form that the change of IRI in a period is the function of the cumulative traffic load passing on the road section during the same duration. There are two problems when applying the pre-defined model in the study. At first, the load is not an available variable by the database. Secondly, the linear associations between the roughness increment and the transport flow are weak according to Pearson's linear correlation coefficients in Tab. 5.1. ESAL, a cumulative traffic load summary statistic, represents a mixed traffic stream determined by the different axle load and the axle configurations. It is an equivalent number of 80 kN single axle load summed over the period. The idea of the equivalent computation is actually to simplify the causing of a variety of vehicle classes, and it assumes that every one unit of load of any vehicle kind has the same effects on the road roughness progression.

With the data support of the numbers of the different vehicle classes that are categorised by the vehicle lengths, the project is able to study the effects of various vehicle classes on road damage. Therefore, the cumulative numbers of the different vehicle categories are taken as the independent variables in the linear regression models. In this way, the assumption of the original model can be tested. Additionally, the relationship between the roughness increment and the transport flow is the weak similarity to the linear association, but the relationship between the roughness in the prediction year and the original roughness is most close to a line (rho^P=0.78). To get a better fitting of the regression model, the linear regression model takes the roughness as the dependent variable and the original roughness as the independent variable. In the original model as Eq. 3.4, there is only the single independent variable (NE_t). The adjusted model has multiple independent variables (the original road roughness and the traffic flow), with different units and magnitudes. If using the raw data with a variety of units and different ranges of the values for regression analysis, it may overestimate the indexes which have large quantities and relatively reduce the influence of the indicators the values of which are small. Therefore, in order to ensure the reliability of the results, the original indicator data needs to be normalized. The data normalization is a process to scale the data down to a small specific interval, and the converted data is commonly a dimensionless pure value (Dutka & Hanson, 1989). It is often used for removing the unit and making the various physical quantities comparable. At present, there are various methods for data normalization, such as linear methods (extreme value method, standard deviation method), polyline method (tri-fold method), and curve method (semi-normal distribution). The most typical one is min-max normalization, that is, the data is uniformly mapped to the interval between 0 and 1. The data normalization function is that for the new data sequence (the dimensionless quantity) y1, y2, etc. is the ratio of the difference between the original data sequence x1, x2, etc. and the minimum value and the difference between the maximum value and the minimum one. The data normalization of the study applies the method. Thus, the applicable linear regression model the study uses is as Eq. 6.1. In the model, IRI and IRI_0 are the dimensionless quantities computed according to the field measurement of a year and the year before respectively. I_L1, I_L2, and I_L3 are the dimensionless quantities computed by the cumulative traffic flow of the passenger vehicles, the light trucks, and the heavy trucks respectively. α_1, α_2, α_3, and α_4 are the estimated parameters in the regression model. u is the constant value. Although in terms of the concept of the model, it indicates without any influence factors (IRI_0, I_L1, I_L2, I_L3 all equal 0), IRI is identical to u. However, it conflicts with the condition that IRI_0 is 0. Thus, u in the model represents all the effects of the other factors that are not taken into account in the model.

$$IRI = \alpha_1 IRI_0 + \alpha_2 I\_L1 + \alpha_3 I\_L2 + \alpha_4 I\_L3 + u \tag{6.1}$$

After setting up the model as Eq. 6.1, the application runs in Matlab. To focus on the effects of the traffic factors, the climate is considered as the control variable. The program code, shown in Appendix G, contains the min-max normalization and the regression. The values of all the parameters are estimated by the regress module in Matlab, as well as the goodness of fitting.

The estimation of all the multiple linear regression coefficients and the indicators of the goodness of the fitting are shown in Tab. 6.1. R-square has a range between 0.55 and 0.76, and the p-value of 0.0000 with the default significance level of 0.05. Because the value of r-square is relatively close to 1, and the p-value is less than the default significance level of 0.05, the goodness of fitting is acceptable. Fig. 6.1 gives the result of the fitting intuitively by drawing the two independent variables and the road roughness. Since all the nodes are near to the regression plane, the fitting result is rather good.

Table 6.1: Regression results of the linear regression model of roughness based on A15 data (No. test data in Climate scenario 1 = 1604; No. test data in Climate scenario 2 = 2524; No. test data in Climate scenario 3 = 1200; No. test data in Climate scenario 4 = 5964)

	Estimation of parameters					Goodness of fitting	
	α_1	α_2	α_3	α_4	u	R^2	p-value (significance level=0.95)
Climate scenario 1	0.9397	0.0464	-0.0120	-0.0966	0.0904	0.76	<0.0001
Climate scenario 2	0.7713	-0.4210	0.1013	0.2861	0.0224	0.70	<0.0001
Climate scenario 3	0.8482	0.8640	-0.5137	-0.3992	-0.0472	0.67	<0.0001
Climate scenario 4	0.7336	-0.3774	0.0882	0.3512	0.0727	0.55	<0.0001

According to the model result in Tab. 6.1, among the four independent variables considered in the regression model, the most significant influence on the dependent variable is the initial value of the road roughness, because the coefficient of the variable is always the largest under the different weather conditions, expect climate scenario 3. Even as for climate scenario, the coefficient is the second larger that is close to the largest value. The significance of the traffic flow' effects on the road roughness is verified, especially under climate scenario 2, 3 and 4. According to the signs of the parameters of the traffic flow, the way that a certain vehicle type affects the road roughness, either negative or positive, is not always certain. Under some weather conditions, with the number of a vehicle class increases, the pavement roughness increases, but under some weather conditions, the road roughness decreased by the reduction of the traffic flow of the vehicle type. Based on the model results, it is observed that the development direction in which the trucks (both heavy and light) influence on the surface flatness is always the same, and it contrary to the cause by the passenger cars. Three possible explanations for the model results are: (1) the development direction of the road roughness affected by the traffic flow is different under the various weather conditions (2) compared with the traffic flow, the climate factor on the road roughness is more decisive (3) the linear form is not desirable for capturing the effect of traffic flow on the road unevenness. The values of the parameters are dependent on the test data. Based on the model results, the coefficient of IRI_0 is not only the largest but much larger than the other parameters in every scenario, indicating that the variable is the decision variable among the independent variables of the model. There is no observation that the coefficients of the three traffic flow have a determined sequence of magnitudes in every scenario. So the model does not find which vehicle type has the greatest impact on road flatness and which less. In summary, according to the fitting indexes, the model can be used to predict the unevenness of the road, but it cannot be used to study the impact of various traffic flow on road roughness.

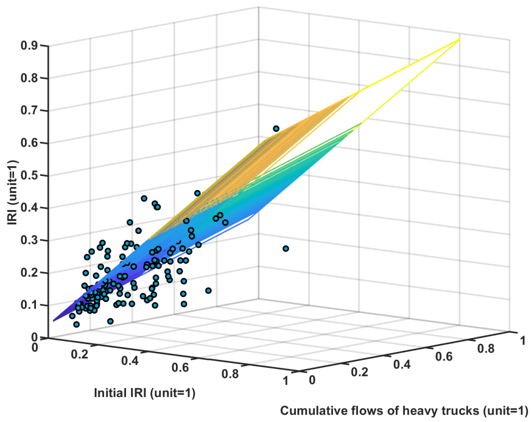


Figure 6.1: Graphical representation of the linear regression model of roughness on A15

The other linear regression model is to predict the rut depth. The setting of the pre-defined model in Chapter 3.3 assumes the linear relationship between the increment of the rutting and the cumulative traffic load (NE_t) over the same period. However, there are two problems to apply the identical model to the study. One is that the project aims to figure out the effects of various vehicle categories, where the pre-defined model ignores the difference of effects caused by the amount of load of the various vehicle types. The other is that based on A15 data and the computation result of Pearson's linear correlation coefficients, the relationship between the change of the rutting and the traffic flow are not close to a line. In other words, the pre-defined linear regression model of the rutting will not result in a good fitting, if applied. Similar to the adjustment of the linear regression model of the road roughness as stated above, the setting of the model takes the rut depth (RUT) as the dependent variables, and the initial rutting (RUT_0), traffic flow of three vehicle types (passenger cars I_1, light trucks I_2, and heavy trucks I_3) as the independent variables. Due to the different units and the magnitudes of the variables, the normalization process is required and min-max normalization (Dutka & Hanson, 1989) is applied. The setting of the regression model is as Eq. 6.2, where the variables (RUT, RUT_0, I_1, I_2, and I_3)are the unitfree items, β_1, β_2, β_3, or β_4 is the estimated parameter of each variable in the regression model, and w is the constant value, indicating the effects of the other factors. It should be noticed that w does not equal to the start value of the rutting. Because the initial value of the rut depth of a section is taken as a variable (RUT_0) in the model.

$$RUT = \beta_1 RUT_0 + \beta_2 I\_L1 + \beta_3 I\_L2 + \beta_4 I\_L3 + w \tag{6.2}$$

The model is solved in Matlab by the code in Appendix G. There computation process contains four parts orderly. At first, reading the data from the excel file, which is the result of the data processing in Chapter 4. Secondly, define the variables, both dependent ones and independent ones. Thirdly, formulate the dependent and independent variables by the min-max normalization method. At last, run the regression module of Matlab and estimate the parameters and the statics of the goodness of fitting.

The regression results in the different weather scenarios are shown in Tab. 6.2. The coefficients of the variables indicate the relationship between the factors and the rut depth, and the sign of each coefficient indicates the direction of the relationship between a predictor variable and the response variable. A positive sign indicates that as the predictor variable increases, the response variable also increases, while a negative sign indicates that as the predictor variable increases, the response variable decreases. Therefore, according to the data analysis of A15 data from 2015 to 2018, the high values of the original rutting makes the progression of the rutting towards deeper, but the effects of the traffic flow vary from the different climate conditions. According to the literature study in Chapter 2 as well as the common sense, vehicles regardless of passenger cars, light trucks, and heavy trucks are the causes of the rutting development. This was confirmed by the correlation analysis in Chapter 5. In the results of the multivariate linear model, most of the parameters of the traffic flow are estimated as the positive values under all kinds of weather conditions. But there are two values that are negative. It does not indicate that the traffic flow are beneficial to the road rutting under a certain weather condition, but it is very likely because of the extreme values in the data. The signs of the constant values are always positive, indicating there is a big possibility that some factors which are not taken into account in the model have a negative influence on the rutting progression.

Table 6.2: Regression results of the linear regression model of rutting based on A15 data (No. test data in Climate scenario 1 = 1622; No. test data in Climate scenario 2 = 2524; No. test data in Climate scenario 3 = 1200; No. test data in Climate scenario 4 = 5964)

	Estimation of parameters					Goodness of fitting	
	β_1	β_2	β_3	β_4	w	R^2	p-value (significance level=0.95)
Climate scenario 1	0.4710	0.3906	0.0247	0.2949	0.1444	0.30	<0.0001
Climate scenario 2	0.6053	0.0232	0.7502	-0.0223	0.1017	0.52	<0.0001
Climate scenario 3	1.373	1.6132	0.6324	0.3370	0.0616	0.65	<0.0001
Climate scenario 4	0.5116	-0.2721	0.1818	0.5687	0.1342	0.38	<0.0001

The coefficient value signifies how much the mean of the dependent variable changes given a one-unit shift in the independent variable while holding other variables in the model constant. This property of holding the other variables constant is crucial because it allows assessing the effect of each variable in isolation from the others. According to the model result, the mean response value that determined by the min-max normalization of A15 data on the rut depth increases by the original rutting increases. But the certain sequence of the magnitude of the traffic flow on A15 is not observed under various weather conditions. The constant values are significant since the response variables have the range between 0 and 1, meaning the factors that are not considered in the model have the nonnegligible influence on the road rut depth.

The goodness of fitting of the model is indicated by r-square and p-value. The values of r-square are rather low, and the p-values of 0.0000 are less than the default significance level of 0.05. Since the good fitting result requires both r-square is close to 1 and the p-value is less than the default significance level, the fitting of the model is not very good. The visual representation of the model is drawn in Fig. 6.2. All the nodes are separated around the regression plane. There is a series of nodes which has the basically same cumulative flow of the heavy vehicles and the similar value of the rutting depth, but their initial value of the rutting vary, which makes the fitting not very ideal.

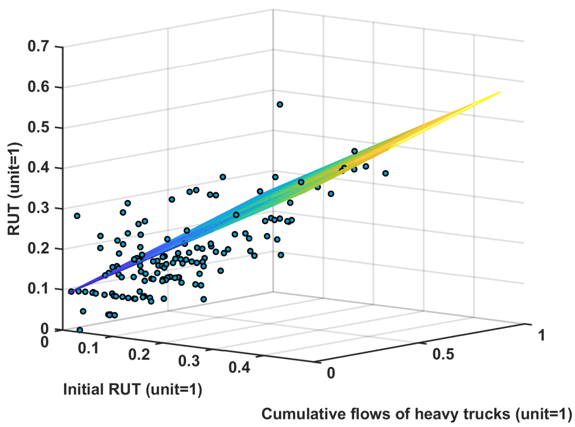


Figure 6.2: Graphical representation of the linear regression model of rutting on A15

To summarize, the model is not a good predictive model for the rut depth, because the fitting is rather bad, and some significant influence factors might be not captured in the model. To some extent, it confirms the undesired effects of the traffic flow on the rut depth, but it does not result in a certain ranking regarding the influence of the different vehicle types on the rutting progression based on A15 data.

6.2. Non-linear regression models

The performance models proposed by the previous papers not only are the linear functions, but also have the non-linear forms. The section contains 3 non-linear regression models defined in Section 3.3, two of which are for predicting the road roughness as Eq. 3.6 and Eq. 3.7 indicate, and one is about the rutting defined as Eq. 3.8. To illustrate every non-linear model clearly, the section elaborates 4 models one by one in the order in which the formulas appear in Section 3.3.

The first non-linear regression model of the road roughness assumes the influence factors of the road roughness are the initial road roughness, the cumulative traffic loading, and time, and the relationship between the response variable and the predictor variables is the exponential function in which the time variable occurs as an exponent, and the initial road roughness and the cumulative traffic loading make up the multiplier. There are two key assumptions of the model. Firstly, the road roughness progression is hypothesized that the variation is small at the start stage, but the roughness increases sharply in the late stage. The rate of growth is determined by the initial value of IRI, the cumulative traffic flow, and the coefficient of the time variable that is estimated by regression. Secondly, the variables of the traffic flow and time are independent. In the case of A15, the fluctuation of the road unevenness has been observed by the data analysis of the test data from 2015 to 2018 as Fig 6.3 illustrates. That's to say, the test data in the case does not meet the first key assumption. As for the second one, according to the correlation analysis in Chapter 5, the correlation coefficients between the cumulative total traffic flow and the ages of the road sections does indicate the high dependence according to the rank correlations (rho^K=0.76 and rho^S=0.90). Since the test data of the study does not satisfy the key assumptions the non-linear model sets, the model is not applicable to the case study.

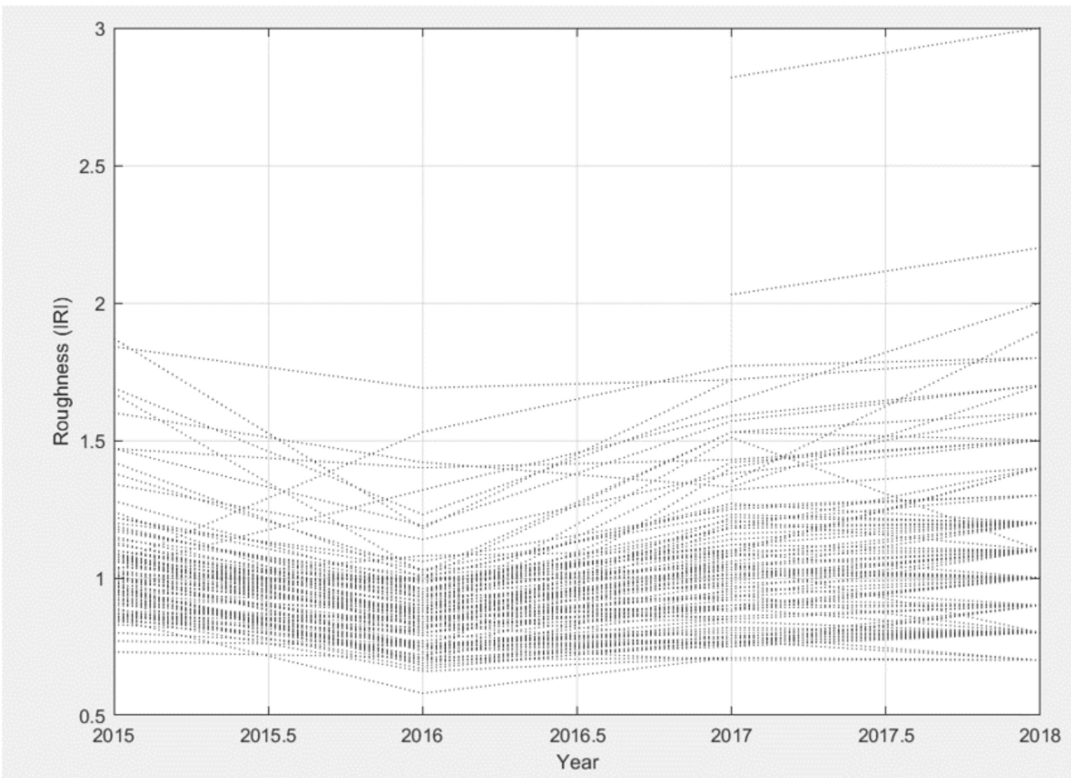


Figure 6.3: IRI progression on A15 from 2015 to 2018

The second non-linear regression model of the road roughness is the function of the change of IRI that computed by the increment of the traffic flow, the change of various performance indexes, including the cracking, rutting, patching, and potholing, and the value of IRI over the same period on a road section. There are two difficulties in applying the pre-defined model. First of all, the data of cracking, patching, and potholing are not included in the database that the thesis uses. Secondly, as defined in Eq. 3.7, the traffic variable in the pre-defined model is the design parameter of the traffic volumes (commonly as ESAL). The computation of ESAL is based on the principle that the mixed traffic would cause the same road damage as the repeating standard axle load does, meaning the model assumes the effects of an amount of load of various vehicle types have no difference regarding the road roughness. The hypothesis is set up due to a lack of real measurement data on each vehicle category. But the data is available in this research. Accordingly, to solve the problems of applying the pre-defined model, the report proposes the adjusted one as Eq. 6.3 shows, which excludes the items in the original model that formulated by the cracking, patching, and potholing, and used the increments of the traffic flow of different vehicle types as the traffic variables (ΔI_L1, ΔI_L2, ΔI_L3). T is the service time of the road section. ΔRUT and IRI are the increments of the rut depth and the road roughness. α_1, α_2, α_3 indicates the effects of the various vehicle types on the change of the road roughness (ΔIRI) during the period (t). The four parameters as well as n, ω_1, and γ can be estimated by regression.

$$\Delta IRI = (\alpha_1 \Delta I\_L1 + \alpha_2 \Delta I\_L2 + \alpha_3 \Delta I\_L3) e^{nT} + \omega_1 \Delta RUT + \gamma IRIt \tag{6.3}$$

It is primarily estimated that the model application to this case probably results in a bad fitting. There are two reasons. Firstly, the model assumes that the amount of change in road roughness is equal to the sum of all kinds of items, which are determined by the predictor variables respectively. The one related to the traffic flow is in the natural exponential function, in which the use time of the roadway is considered as the exponent and the coefficients are made up of the traffic flow and the estimated parameters. According to the feature of an exponential function, the increment of the response variable will increase by the increase of the predictor variable as the exponent. In this case, it means the change of road roughness has been increasing as the service time of the roadway goes by. However, the kind of development is not observed in the test data, as Fig. 6.3 illustrates. Secondly, because the three types of distress data are not available, the exclusion of the variables may make the model inapplicable.

Considering the negative expectation of the model application result, the case study applies preliminary testing by inputting A15 data from 2016 to 2017 under the climate scenario 1. The min-max data normalization (Dutka & Hanson, 1989)is applied. The non-linear function cannot be solved by the “curve fitting tools” in Matlab, because it is not one of the forms that are initially installed in the program. The solution for the calculation is converting the regression problem to the optimization problem. To achieve a good fitting, maximizing r-square is set as the objective. The constraint is that n is positive. The parameters are computed in the “optimization tool” in Matlab.

The model application result of the preliminary testing is as Eq. 6.4. The R-square of 0.1756, much less than 1, indicates that the model fits poorly, as expected. The model result does not include the items determined by the traffic flow and the use time, is most likely because the correlation of the roughness and the service time is not in the form of exponential. In summary, the nonlinear model is not suitable for this case.

$$\Delta IRI = 0.0011 \Delta RUT + 0.0096 IRI_0 t \tag{6.4}$$

The third non-linear regression model is for predicting the rut depth. It assumes that the rut depth is the exponential function of the cumulative traffic load and the service time of a roadway. There are two problems of applying the pre-defined model to the study. The one is that the original model takes ESAL as the traffic variable which hypothesizes the same effects of various vehicle types on the rut depth as long as the amount of the load is identical. In the research, the data that contains the total number of each vehicle category, is able to simulate the traffic conditions more similar to reality. Thus, the model in the study is taking the percentages of three vehicle types (the passenger cars, the light trucks, and the heavy trucks) as the traffic variables (p_1, p_2, and p_3), each of which have a parameter (β_1, β_2, or β_3), that indicates the influence of each vehicle type on the rut depth. The formula of the model is in Eq. 6.5. The other parameters and variables in the model have the same meanings as the pre-defined one, where T, the age of a road section, is the base and its exponent is x, and the cumulative total traffic volume (I_Al) as the base has the exponent of y.

$$RUT = (\beta_1 p_1 + \beta_2 p_2 + \beta_3 p_3) T^x I\_AL^y \tag{6.5}$$

The other problem is the dependence of the two predictor variables, the cumulative traffic flow and the service time. According to the correlation analysis in Chapter 5, the linear correlation coefficients of the traffic flow and the total service time range from 0.24 to 0.26, but the rank correlation coefficients, especially Spearman's rho, have the high values of between 0.88 and 0.91, since the maximum is 1. The results make sense because as for a road section, the longer it is in the operation stage, the larger the cumulative traffic flow it has, but the sections out of a number of the pavements which have the long service duration do not necessarily have the higher total traffic volumes. Due to the rather weak linear correlation and the strong rank correlation between the variables, they are not independent to some extent by the test data.

$$lgRUT = lg(\beta_1 p_1 + \beta_2 p_2 + \beta_3 p_3) + xlgT + ylgI\_AL \tag{6.6}$$

In order to confirm if the correlation of the variables of I_AL and T does influence the model result, the logarithm method is used to convert the model to a multivariate function as Eq. 6.6 shows. The independent variables in a multivariate linear fitting model must be linearly independent, otherwise, in terms of the goodness of fitting, the model is always ideal. The predictor variables of the logarithms of I_AL and T in the model are drawn in Fig. 6.4. It confirms the clear linear correlation between the two variables. That's to say, even if the model results in good fitting, it is likely because the independent variables are dependent. Therefore, the non-linear regression model is not applicable to the case study.

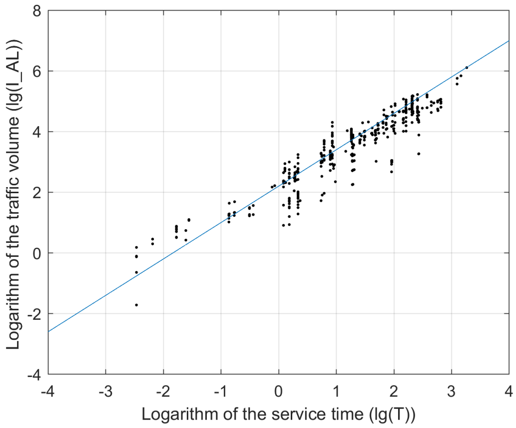


Figure 6.4: Logarithms of the variables of I_AL and T

In the section, the three nonlinear models are all predicted to be unsuitable for the study case. The prediction is derived from that the measurement data on A15 does not support the key assumptions of the models. In particular, as for the first and third regression models elaborated above, it is very clear that the measurement data the case uses is against the key assumptions the models set. Therefore, the model application result is not displayed in the report. About the second nonlinear model, although it is assessed that the regression would not result to be ideal, the study provides the results of the model application of the experimental data to prove the prediction. Actually, in the study process, the test data was first applied to the model, and it resulted in the non-logical outcomes. Then the reason was derived, which was the case data did not meet the key assumptions of the models. However, it should have been figured out in the pilot assessment of the applicability of the models. So as to a clear elaboration that conforms to the general thinking habits, the section is organised as firstly giving the prediction of the applicability followed by the proof, which inverses the actual model application process.

6.3. Conclusion

The chapter is to apply the regression models for predicting pavement performance of A15. In the last chapter, five models (3 for roughness and 2 for rutting) have been defined with the predictor variable of the cumulative traffic load based on previous studies. Due to one research goal of studying the influence of different vehicle types on road performance, on the foundation of the pre-defined regression models, the study replaces the total traffic flow with the sum of the cumulative traffic volumes of each vehicle type, and run the model application. The outcome is that the two multivariate linear regression models have better fitting results for the A15 highway based on the measurement data from 2015 to 2018, but the three nonlinear regression models are not suitable for the case study. The reasons for the inapplicability contain that the road roughness progression on A15 during the four study years does not meet the model hypothesis; several independent variables in the model have no measurement data; and some independent variables in the model are not independent on basis of A15 data. The two linear models in this study result in good fitting mainly because the roughness and rutting in the predictive year are strongly correlated to the original values in the A15 data, especially the roughness. It requires further study to clarify whether considering the traffic flow of various vehicle categories can achieve the better prediction. In summary, according to the model results, the linear models proposed in the study can predict the road roughness and the rut depth on A15, and different vehicle types have various effects on road performance development under different weather conditions.

Apart from the model application results of the regression models, the chapter finds the answer to sub question 5a and 6a as elaborated in the following.

Sub question 5a

How to apply the regression models?

The application of the regression model in the study is made up of 4 steps. At first, a new model is established by replacing the total amount of traffic with the sum of the traffic volumes of the different types of vehicles in the pre-defined models. Secondly, it is to assess the applicability of the models, according to the features of A15 data. Thirdly, the test data is input in the models and the regression coefficients are computed, in which the different physical quantities are converted to the dimensionless quantities by min-max normalization. At last, the effects of the factors on the road performance are analyzed according to the model results.
In the application process of each model, not all the steps are completed. The application of two non-linear models (one for unevenness and one for rut depth) does not proceed when they are determined that the test data does not support the key model assumptions. The other non-linear model of the increment of the road roughness results in a poor fitting by using the test data that meets the pilot assessment. Thus, it is unable to provide the analysis results of the effects of the factors. In the study, the two linear models have the complete model application process, and the model results to some extent quantify the effects of the traffic flow as well as the other influencing factors on the road performance characteristics.

Sub question 6a

What are the results of the regression models by the test data?

The study finds two regression models that are applicable to A15 road performance progression prediction. The models are in the linear forms as Eq. 6.1 and Eq. 6.2 define. According to the application results of A15 data from 2015 to 2018, the important factors that affect the road roughness are the initial roughness of the road and the traffic flow. The cumulative number of each vehicle type has different effects on road roughness under varied weather conditions. Even if under a certain climate condition, it is inconclusive that the traffic has the positive or negative effects on the road roughness. The development direction of road roughness to good or to bad may be jointly determined by the traffic flow of the mixed traffic and the weather conditions. The traffic volumes of an individual vehicle type cannot be determined to be beneficial or detrimental for a roadway with regards to the roughness. It is worth noting that, according to the results of the A15 data analysis, the impact of trucks (both the light type and the heavy type) on the development direction of pavement roughness is consistent under any kind of weather condition, and that is contrary to the effects of the passenger cars. However, whether the finding suggests a certain pattern or it is a coincidence caused by the limited test data requires further research to clarify.
(To be continued in next page)

Sub question 6a

What are the results of the regression models by the test data?

(The answer continues here)
According to the results of the linear roughness model of the rut depth of A15, the initial value of the rut depth and the traffic flow of the three kinds of vehicle types have an important influence on the road rut depth. The pattern in the test data has been found that the increase of the traffic volumes of any type of vehicles will make the rut depth increase, under all kinds of climate scenarios. But the quantities of the effects of various vehicle categories on the rutting vary from the different climate scenarios. Among the four weather situations in the study, the most significant traffic-flow factors of the rut depth can be the number of the passenger cars (for climate scenario 1 and 3), the number of the light trucks (for climate scenario 2), or the number of the heavy trucks (for climate scenario 4). One explanation is that the truck is supposed to have a large impact on the rut depth of the pavement due to its heavy load, but the passenger cars may have a large impact due to the accumulation of the numbers. Under various weather conditions, the share of each vehicle category on the A15 sections may lead to the ranking of the influence of various vehicle types on the rut depth, which requires to be confirmed by further study.
The goodness of fitting in the two linear fitting models proposed in this project are both acceptable. However, the three non-linear regression models in the study do not result in the ideal outcomes. The main reason for the inapplicability of each non-linear models is respectively: (1) the road roughness progression on A15 during the four study years does not meet the model hypothesis; (2) several independent variables in the model have no measurement data; (3) some independent variables in the model are not independent on the basis of A15 data.

The thesis will continue to apply the methodology defined in the chapter. The model applications of the survival model and the decision tree classifier will be in the remaining content of Develop phase.

7

Survival Model

This chapter applied the survival model to estimate the traffic impacts on the initiation of bad roughness and rutting. The model has been defined in Section 3.4. The test data for the application contains the entire A15 data from 2015 to 2018 provided by the database as Chapter 4 introduces. The outcome gives the effects of the traffic volumes of the different vehicle categories on the survival time of the roadway concerning the pavement performance. As stated in Section 3.4, the process of applying the survival analysis contains five steps: (1) selecting the performance indicators; (2) defining the threshold values of the performance indicators; (3) modelling the survivor function; (4) estimating how the expected failure time depends on traffic characteristics; and (5) calibration. Each step becomes a section of the chapter. The last section is to summarize the findings of the chapter and provides the answer to SQ. 5b how to apply the survival model and SQ. 6b what are the results of the survival model by the test data.

7.1. Step 1: selecting the performance indicators

Pavement conditions of A15 sections are evaluated by distress, roughness, and skid resistance measured on the outer and median lanes. The research initially selected cracking and raveling as the candidate performance indicators. Because the two are mainly concerned in real cases, and the previous research confirms the applicability of the survival model for predicting the cracking (Loizos & Karlaftis, 2005). However, limited by the accessibility of data, the two variables have no values in the database the study uses. Therefore, the two performance indicators are not selected. As stated in Section 3.4, the previous study applied the survival model to predict the occurrence of the bad condition points of the pavements (Wang et al., 2005). Condition Points (CD) is a comprehensive performance indicator determined by cracking, subgrade failure, raveling, swell and patches. However, it is actually a biased procedure of setting up the weights for every performance indicator.

The research selects roughness and rutting as the performance indicators, because in an academic view, there is a lack of the survival model application regarding the performance indicators, and in a practical view, the database via IVON provided by RWS provides more complete data of roughness and rutting and also they are two important indexes to assess the roadway performance indicators.

7.2. Step 2: defining the threshold values of the performance indicators

It has been observed that pavement deterioration, like the initiation of the cracking, would not appear for a long period, but once it occurs, it will develop dramatically in a short time (Wang et al., 2005). As for road roughness, it may fluctuate all the way. It is observed in many A15 sections that the more uneven the road is, the greater fluctuation in IRI annually as Fig. 6.3 shows. The progression of the rutting is a process that is affected by traffic loading, pavement strength, climate and drainage condition (Alaswadko & Hassan, 2018). The development of rutting can be divided into three phases in Fig. 7.1 (Freeme, 1983). Initial densification or bedding-in phase is that the densification occurs on the newly constructed pavement after it is opened to traffic. Its degree depends on the level of compaction during construction and the applied traffic load. In the second stage, gradual or stable rate of deformation phase, the deformation rate constantly increases with traffic. This rate depends on several factors such as traffic loads, pavement strength and environment. The accelerating deformation phase as the final one has the rapid increasing deformation rate depending on traffic loading, pavement strength and environment.

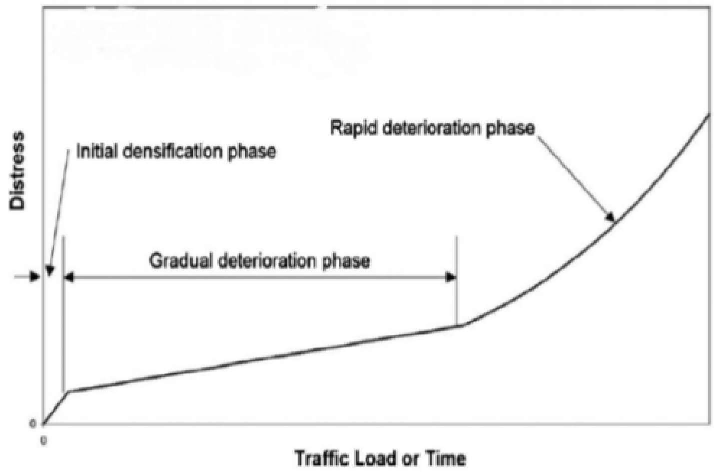


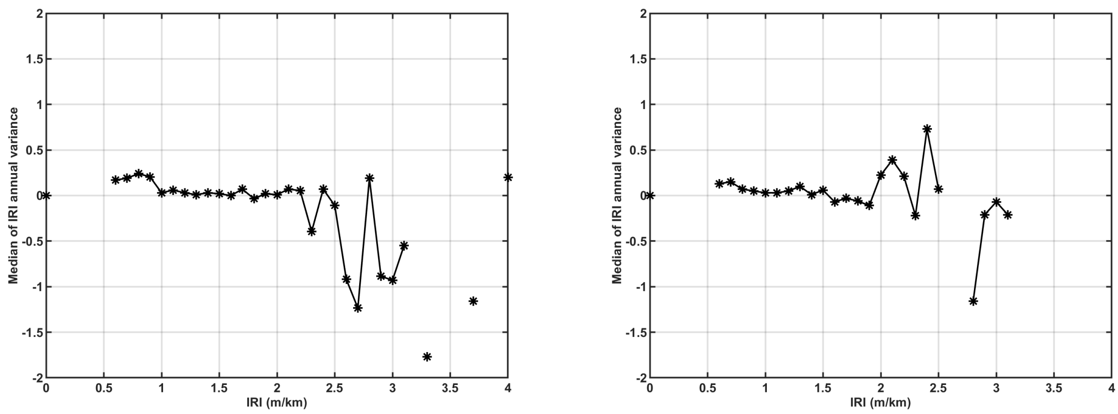
Figure 7.1: Phases of rutting progression (Freeme,1983)

Accordingly, the study assumes the threshold of the pavement deterioration where road sections with the values of the deterioration index below the threshold and above the threshold have the different development regarding the the distress. Thus the life-span of a road section is divided into two stages in the research: one is the normal stage where pavement is in the good condition with the acceptable unevenness and rutting; the other is the failure-development stage when the condition is bad and it potentially develops towards worse direction rapidly. The estimation of the thresholds of the performance indicators is not only required, but also the values are of great significance in the resurfacing decisions.

To estimate the thresholds, there are two applicable methods. The first one is employed by data analysis. In previous study, the critical value concerning CD was defined by diagramming CD as the x axle and the annual increment of CD as the y axle (Wang et al., 2005). When the fluctuation of the polyline suddenly increased, there was a big deteriorating to the next stage. The study applies basically the same procedure to find the thresholds of the rut depth and the road roughness on A15. The only slight change is selecting the medium as the annual change of the certain value of the performance indicator. Because, in the big data set, the road sections which have the same values of the performance indicators may have the different change of the next year. The median, a value that is not easily affected by extreme values and can represent a medium level of value, is used as an indicator of the annual growth.

The progression of road roughness of A15 in two districts are respectively in Fig. 7.2a and Fig. 7.2b. The computation code in Matlab is in Appendix H. Most of the unevenness of the test road sections are basically above 0.5 m/km, but there are some points with the value of 0 m/km, which are the cases of lack in measurement data. As Fig. 7.2a and Fig. 7.2b illustrate, when the value of IRI is relatively small, the median annual change swings around 0 m/km. When IRI reaches a high value, the annual change becomes considerable. In the diagrams, there are some discrete points and discrete polylines in the area of high-values of IRI, indicating that no test data fell in the intervals. According to the data analysis on A15, the road roughness fluctuates when the roadway has a slight unevenness. Of the pavement in WNZZ district, after the road roughness approaches 2.3 m/km, the annual variation of the unevenness on the roadway becomes remarkable. As for in ONZ district, the critical number is about 1.9 m/km.

The same method of determining the threshold of rutting on A15 is programmed in Matlab. Limited by the length of the report, the code is not displayed in the report. Fig. 7.3a and Fig. 7.3b diagram the progression of rutting of A15 in two districts respectively. It is noticed that there is a piece of line unconnected to the continuous polyline in Fig. 7.3b. It is because that there is no test data belonging to the intervals of 12 mm to 13 mm

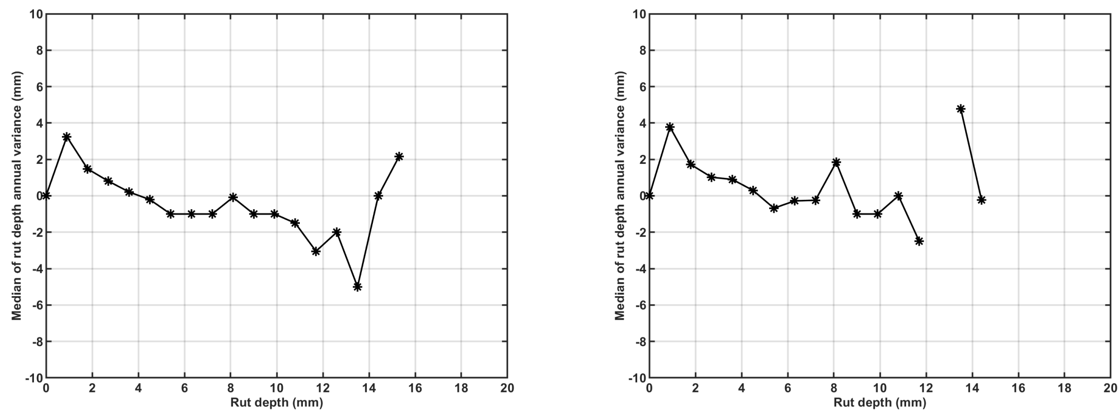


(a) ZOAB road roughness progression from 2015 to 2018 of A15 in WNZZ district (No. test data = 2556)

(b) ZOAB road roughness progression from 2015 to 2018 of A15 in ONZ district (No. test data = 1372)

Figure 7.2: ZOAB road roughness progression from 2015 to 2018 of A15

of the rut depth. According to the data analysis of ZOAB sections on A15 in the two districts, when the rutting is slight but not at the initial densification phase, the median annual change is relatively small, and when the rutting is deep, the median annual change becomes considerable. In WNZZ district, after the roadway has a rut depth above 10 mm, the annual variation of the rutting on the roadway becomes remarkable. The critical value in ONZ district is about 10 mm, as observed. Based on the analysis results, most of the measurement data of rutting is in the initial densification phase or the gradual deterioration phase. The lack of data on the rapid deterioration phase may cause some decreasing slopes in Fig. 7.3a and Fig. 7.3b, which are not in line with knowledge.



(a) ZOAB road rutting progression from 2015 to 2018 of A15 in WNZZ district (No. test data = 2556)

(b) ZOAB road rutting progression from 2015 to 2018 of A15 in ONZ district (No. test data = 1372)

Figure 7.3: ZOAB road rutting progression from 2015 to 2018 of A15

The second method is dependent on empirical knowledge to figure out the threshold values of the performance indicators . As stated in Section 1.6, RWS established up the damage assessment by the comprehensive considerations of the impact of road roughness on serviceability and maintenance costs. The good condition of a pavement regarding IRI is below 2.6 m/km, and the intervention level is above 3.4 m/km. As for rutting, the critical values respectively are 10 mm and 17 mm.

The thresholds of the rut depth found by the two methods are identical, while those of the road roughness are similar. The development of the pavement deterioration in the different regions on A15 has the similar shape. It means that the roughness and the rutting of ZOAB road sections in the different regions have commonalities.

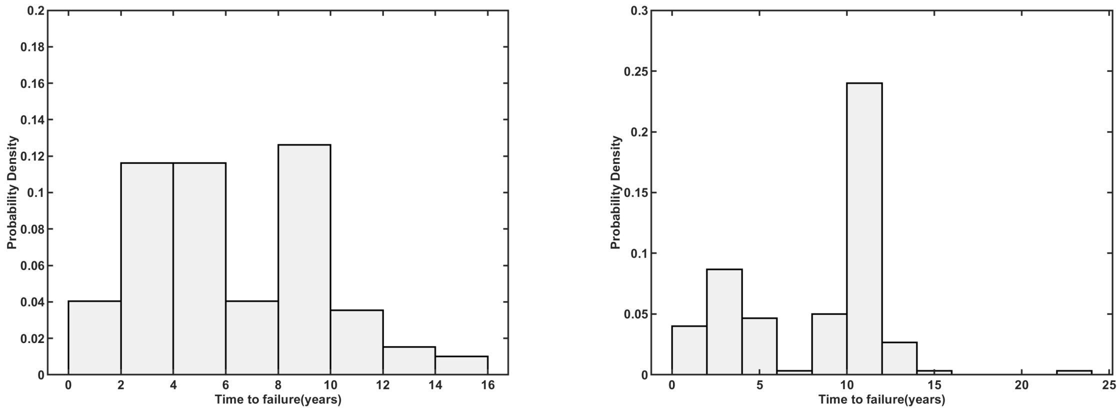
In terms of being practical, the second method should be selected. However, A15 performance from 2015 to 2018 was mostly in good condition. There is few data in the failure-development stage if defined by the damage assessment. Tab. 7.1 gives the data proportions of multiple classes. The following study process may be impacted due to the insufficient data. Accordingly, the study uses the results of the first method to define the threshold. As for the rutting, the two methods result in the same outcomes, which makes it no necessary to compare or select the method concerning the rutting in the study.

Table 7.1: Data distribution of the different damage assessments of IRI

	Stage	
	I	II
Data analysis of A15	IRI<2.3 m/km	IRI ≥ 2.3 m/km
Data distribution	98.12%	1.88%
Damage assessment of RWS	IRI<2.6 m/km	IRI ≥ 2.6 m/km
Data distribution	98.78%	1.22%

7.3. Step 3: modelling the survivor function

According to the results in the last section, a road section with the value of IRI above 2.3 m/km or the value of the rut depth larger than 10 mm is considered as the poor condition. The period from its construction date to the survey date of the poor condition is defined as the survival time. Based on A15 data from 2015 to 2018, the distribution of the survival time of ZOAB sections regarding roughness and rutting is diagrammed respectively in Fig. 7.4a and Fig. 7.4b. The computation code in Matlab is shown in Appendix I.



(a) ZOAB road survival-time distribution concerning IRI based on A15 data from 2015 to 2018 (No. test data = 3928, No. observation data = 99)

(b) ZOAB road survival-time distribution concerning rutting based on A15 data from 2015 to 2018 (No. test data = 3928, No. observation data = 150)

Figure 7.4: ZOAB road survival-time distributions based on A15 data from 2015 to 2018

The distribution of the survival time concerning IRI based on the input data of A15 indicates that most of the ZOAB road sections which fail during the observation period are between 2 and 10 years. To further study what form the test data is distributed in, multiple classic distribution functions are tested, including normal distribution, Weibull distribution, logistic distribution, gamma distribution, log-logistic distribution and lognormal distribution. The goodness of fitting of the distribution is evaluated by the log-likelihood, as the classic statistics method. The maximum log-likelihood indicates the best-fitting distribution. The mean, the variance, and the log-likelihood of every distribution are given in Tab. 7.2. The best fitting function, in this case, is Weibull distribution according to the likelihood ratio test. The fit of the Weibull functional form, diagramed in Fig. 7.5a, suggests that the probability of the critical road roughness occurring during the next period is higher when it has not done so up to approximately 12 years, with the maximum probability of roughness at 6.11 years, and then becomes low from 16 years and on. It should be noticed that the numbers are derived from the mathematical analysis, but because there is few data about 16 or more years, it requires critical consideration when applying the numbers.

Table 7.2: Distribution functions of ZOAB road survival time concerning IRI based on A15 data from 2015 to 2018 (No. test data = 3928, No. observation data = 99)

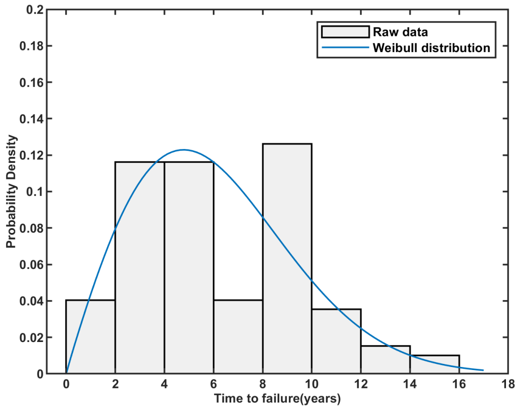
Distribution	Mean	Variance	Log-likelihood
Normal	6.15	10.06	-254.23
Weibull	6.11	10.57	-253.30
Logistics	6.00	11.28	-256.52
Gamma	6.15	14.37	-258.65
Log-logistic	7.01	49.99	-263.72
Lognormal	6.90	42.03	-277.22

According to the measurement data of the rut depth of ZOAB pavements on A15 from 2015 to 2018, there were two peaks of the failure event occurrence. One is at the service time of about three years and the other is approximately 11 years. In particular, there is an extreme failure peak at around the 11th year. The most widely-used functional forms for the survival models are used, including normal distribution, Weibull distribution, logistic distribution, gamma distribution, log-logistic distribution and lognormal distribution, and the likelihood ratio test was used to compare the goodness of fitting of the different functional forms, as the same method as the study of the distribution of the survival time regarding IRI above. The results of the mean, the variance, and the log-likelihood of every distribution are listed in Tab. 7.3. The likelihood ratio test indicates that the normal form is the most appropriate for the test data. The normal distribution indicates that ZOAB sections of A15 have the highest probability of failure concerning the significant rutting in 8.09 years. Although the normal distribution is the best fitting distribution of the case study with regards to the log-likelihood, the gap between the raw data and the normal distribution is large, as Fig. 7.5b illustrates. The main reason is that the original data has two peak points, but whether the normal distribution or other probability distribution forms used in the study has the feature of the only one peak. In general, the goodness of fitting of the distribution functions of ZOAB road survival time concerning rutting based on A15 data from 2015 to 2018 is not ideal. Actually, the difficulty has been proposed in the previous study that the model results are tightly correlated to the availability and nature of the data and it is difficult to find the appropriate fit, a simple distribution, in a particular case sometimes (Ebrahimi et al., 2019).

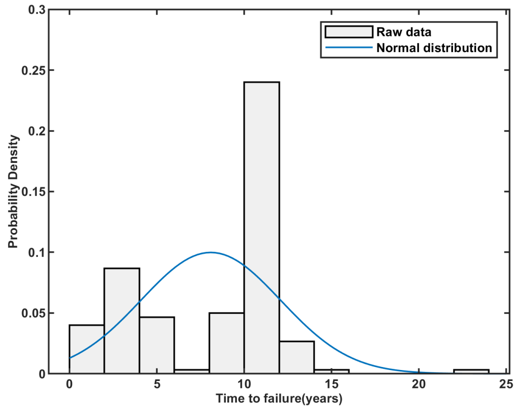
In summary, this step finds the survival-time distribution of the ZOAB road on A15 based on the data from 2015 to 2018. The one concerning surface roughness is in the form of Weibull distribution with the expected survival time of 6.11 years. As for the rut depth, the normal distribution is found as the appropriate distribution to the case, and its mean value of the survival time is 8.09 with the unit of years. In the dataset that the study uses, the majority of ZOAB sections were in good conditions and only 1.88% (in Tab. 7.1) met the maintenance requirements related to the performance indicators the research selected. That is to say, the survivor functions the step models are determined by a small amount of data.The problem will be further discussed and the solution will be proposed in Section 7.5.

Table 7.3: Distribution functions of ZOAB road survival time concerning rutting based on A15 data from 2015 to 2018 (No. test data = 3928, No. observation data = 150)

Distribution	Mean	Variance	Log-likelihood
Normal	8.09	15.96	-420.10
Weibull	7.97	17.84	-427.72
Logistics	8.44	17.89	-423.32
Gamma	8.09	28.96	-440.60
Log-logistic	10.26	234.36	-456.33
Lognormal	9.50	109.75	-473.18



(a) Best-fitting distribution of ZOAB road survival time concerning IRI based on A15 data from 2015 to 2018 (No. test data = 3928, No. observation data = 99)



(b) Best-fitting distribution of ZOAB road survival time concerning rutting based on A15 data from 2015 to 2018 (No. test data = 3928, No. observation data = 150)

Figure 7.5: Best-fitting distributions of ZOAB road survival time based on A15 data from 2015 to 2018

7.4. Step 4: estimating how the expected survival time depends on traffic characteristics

The survival model defined in Eq.3.9 is an accelerated life model, which assumes the effect of covariates on the survival time is multiplicative. There are two quantities of the predictor variables in the function. One is the baseline survival-time distribution (T_{s0}) that has been defined in the last step, and the other one is the exponential sum of $f\ \lambda$, where f is a matrix of covariates and λ is a vector of regression coefficients measuring the impact of covariates. The way of estimating how the expected survival time depends on traffic characteristics in the previous literature is taking the traffic volumes as one column of the covariate matrix (Ebrahimi et al., 2019), or taking ESAL and/or the traffic levels in the covariate matrix (Wang et al., 2005; Loizos and Karlaftis, 2005). In the study, the effects of the three vehicle categories (passenger cars, light trucks, and heavy trucks) are focused. Because the three variables have some dependence according to the correlation analysis in Chapter 5. In the application, the predictor variable of f is one of the cumulative numbers of the three type of traffic flow. The test data is divided into WNZZ and ONZ district, where are respectively located in the coverage of Rotterdam weather station and Herwijnen weather station. In this way, the climate factors are controlled.

To estimate how the expected failure time depends on the multiple traffic characteristics, the cox proportional hazards model (Kartsonaki, 2016) is utilized. Because the traffic intensities as the predictor variables in the study are the time-dependent covariates, it is necessary to convert survival data to the counting process form at first (Mathworks, 2006a). The method is executing the code that converts the survival time (T_s) to a time interval (T_{str}, T_{stp}]. The estimation results of the coefficients are listed in Tab. 7.4. Coefficients with the negative values indicate that the existence of the corresponding variables reduce the expected road survival time. The ratio of the time reduction is defined by the hazard ratio, which is the exponential of the coefficient in the survival model. Standard errors, z-values and p-values describe the characteristics of the coefficient estimates. When the p-value is greater than 0.05, the null hypothesis cannot be rejected. It means that the difference between the samples selected out of the population is caused by the sampling error. In this case, it happens due to the small population data.

According to the survival analysis results of A15 ZOAB data from 2015 to 2018, three types of vehicles all have a negative influence on road health determined by IRI, but the degrees of their influence are different. The intensities of the passenger cars shorten the mean of the survival time by 4.33% based on WNZZ data and by 4.05% derived from the data in ONZ district. The effects of the volumes of the light trucks on A15 in the WNZZ and ONZ districts during the observation year is to decrease the expected survival time of IRI by 52.08% and 29.69% respectively. The impact of light trucks on ZOAB road survival time defined by IRI is significantly different in the two regions. According to the high value of p-value that is above 0.05, there is less significance

Table 7.4: Estimation results for ZOAB road roughness on A15 (No. test data in WNZZ = 2556; No. test data in ONZ =1372)

Variable	Coeffcient	Hazard ratio	Standard error	z-value	p-value (significance level=0.95)
(WNZZ district)					
I_L1	-0.0443	0.9567	0.0085	-5.2006	<0.0001
I_L2	-0.7357	0.4792	0.1660	-4.4318	<0.0001
I_L3	-0.5084	0.6015	0.1467	-3.4656	<0.0001
(ONZ district)					
I_L1	-0.0413	0.9595	0.0200	-2.0603	0.0395
I_L2	-0.3522	0.7031	0.1928	-1.8264	0.0678
I_L3	-0.4693	0.6254	0.2095	-2.2407	0.0250

Table 7.5: Estimation results for ZOAB road rutting on A15 (No. test data in WNZZ = 2556; No. test data in ONZ =1372)

Variable	Coeffcient	Hazard ratio	Standard error	z-value	p-value (significance level=0.95)
(WNZZ district)					
I_L1	-0.0325	0.9680	0.0057	-5.7227	<0.0001
I_L2	-0.3667	0.6930	0.0596	-6.1529	<0.0001
I_L3	-0.2776	0.7576	0.0477	-5.8163	<0.0001
(ONZ district)					
I_L1	-0.0926	0.9156	0.0438	-2.1131	0.0346
I_L2	-1.5288	0.2168	0.7321	-2.0881	0.0368
I_L3	-1.4397	0.2370	0.7320	-1.9669	0.0492

in the statistical test based on the ONZ data so that it cannot be rejected that the null hypothesis is valid. So the result is invalid. As for the influence of the heavy-truck flow on the survival time analysed in the case, they reduced the life expectancy of ZOAB roads in the WNZZ area by 39.85%, compared to 37.46% in the ONZ area. Apparently, the impacts of trucks (both light and heavy) on the expected survival time regarding IRI are significantly considerable.

The same method is employed to estimate how the expected survival time concerning the rut depth depends on the intensities of the three vehicle categories. In the study, the road section is defined as "fail" when the rut depth on it was measured above 10 mm. The critical number is defined in Section 7.2. The estimation results are in Tab. 7.5. All types of vehicles have an negative influence on the ZOAB pavement survival time determined by the rut depth. The intensity of each vehicle category affects the expected survival time regarding the rutting in a different scale. The traffic flow of the passenger cars decrease the mean of the survival time by 3.2% in the WNZZ district and 8.44% in the ONZ district. The effects of the light trucks on the road health are significantly indicated by the hazard ratios. The class shortens the expected survival time by 30.7% and 78.32% according to the data analysis of WNZZ and ONZ data respectively. As for the heavy trucks, the estimation results indicate that the number of the vehicle class reduces the expected survival time of ZOAB pavements of A15 by 24.24% in the WNZZ region, and the number of 76.3%. Comparing the estimation results calculated by the data in the WNZZ district and the data in the ONZ area, the influence of any vehicle type on the expected survival life duration of the ZOAB pavements derived from the WNZZ region is greater than that of ONZ. Especially, the two types of trucks shorten the pavement life expectancy in the ONZ area by an extremely large proportion, which is nearly 4/5. The finding is not comparable with the existing literature as well as common sense. Apparently, the traffic impact on the ONZ sections is significantly overestimated. The reasonable explanation for this is due to less observation data. For the same reason, the thesis can not confirm or reject the likelihood that the traffic impact on the survival time estimation in the WNZZ area has also been overestimated.

Although the hazard ratios for the survival time of ZOAB concerning roughness and rutting that associated with the trucks are larger than that of passenger cars, it is not a critical conclusion that the impacts of the trucks on ZOAB road life expectation is greater than that of passenger cars. Because the survival model defined in this study has a single predictor variable. The computation of the coefficient is based on the assumption that only the variable influences the survival time distribution, while the other potential influencing factors have none effects. That is to say, the study of the impact of the passenger car flow on the ZOAB road survival time is according to the hypothesis that the impact of light truck traffic on the survival time concerning the roughness and the rut depth of the ZOAB pavements is 0, and neither does the heavy trucks. The way of studying the effect of multiple variables on the expected survival time is the stratified model. The model has the same function as Eq. 3.9, where f is a matrix containing k columns, where the k columns correspond to the predictor variables using the name-value pair argument. The unstratified model is applicable to the study, rather than the stratified form, because the three traffic variables are not independent.

It is noticed that the tests on the WNZZ district are valid but some on the ONZ district are not, according to the p-values. The main reason is that from the observation year of 2015 to 2018, there is few sections with the IRI value higher than 2.3 m/km (as defined in Section 7.2) on On A15 from 2015 to 2018. The number of observations in the WNZZ district is 39, but only 8 data in the ONZ district. As for the rutting, 48 ZOAB sections had the measurement data of the rut depth above 10 mm, the threshold value. Only 12 in the ONZ area were observed as "fail" determined by the rutting from 2015 to 2018. The consequences of a lack of observation data are (1) some p-values are close to or even larger than the significance default level, indicating the null hypothesis is not rejected; (2) it is likely that the effects of the traffic characteristics on the expected survival time are overestimated; (3) the limited observation data results in a right censoring problem. The solution to the problem is discussed in the next section.

7.5. Step 5: calibration

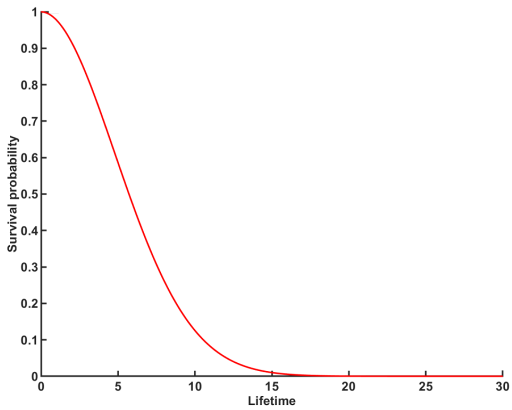
Censored data problem in the survival analysis has been observed in both the previous studies (Paterson, 1987; Ebrahimi et al., 2019) and the model application steps of this case above. It is caused by the limited survey time, and it indicates no observation of the failure events in the measured sample(s). There are three common types of censored data problem (R. Miller & Halpern, 1982):

- Left censoring, means the failure event occurred before the observation starts.
- Interval censoring, indicates the failure event happened in an interval during the whole survey duration, but the specific time of the occurrence of the event can not be determined
- Right censoring, means the failure event will happen after the observation ends.

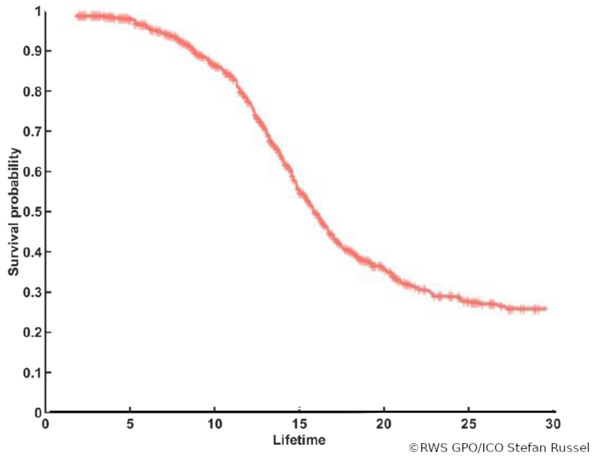
In this case, the three types of censored data problem are involved. The survey time is from 2015 to 2018, and the measured objects are all the A15 sections whose surface coating material is ZOAB. Among these subjects, a small number of roadways were observed that the failure event happened in the first measurement in the year of 2015. The failure event was defined as the road roughness is above 2.3 m/km, or the rut depth is above 10 mm, as Section 7.2 finds. The actual failure time was earlier than the beginning of the observation. The data is left censored. Besides, the road performance (roughness and rut depth, in this case) is measured once a year. When the road was observed to meet the failure criteria for the first time in the measurement of 2016, 2017, or 2018, the real failure time was at some point between the last measurement and the measurement in which the failure event was observed. This part of the data has the interval censoring problem. As of the last measurement in 2018, more than 90% of the ZOAB sections on A15 did not meet the failure criteria specified in this report (IRI> 2.3m/km, rut depth>10 mm). The failure time of these sections cannot be estimated, which caused a considerable right censoring problem for the study.

The solution to the left censoring problem is converting the survival time data to the survival interval, which is defined by two variables, T_{str} and T_{stp}. It indicates the failure event happened on the road section between the start time (T_{str}) and the end time (T_{stp}). As for left-censored data, T_{str} is 0, and T_{stp} is the duration between the construction time and the survey date of 2015. The method has been employed in the model application above.

Similar to the processing of dealing with left-censored data, the interval censoring problem can be solved by converting the data of the survival time to the interval of the survival time too. The indicators of the survival interval, T_{str} and T_{stp}, are the number of days between the construction date and the last measurement , and that between the construction date and the survey when the failure event was observed respectively. The processing way has been used in the model application above.



(a) Survivor function of ZOAB pavements concerning IRI based on A15 data from 2015 to 2018 (Rijkswaterstaat, 2019)



(b) Survivor function of ZOAB pavements based on the long term measurement of the roadways in the Netherlands

Figure 7.6: Survivor functions of ZOAB pavements

The solution to the right censoring problem is adding a column of the binary variable that indicates whether the road section meets the failure criteria as defined in the case study to the matrix f. The value of 1 of the variable means the road performance of roughness or rutting reached the threshold level by the measurement in 2018, which was the end of the observation in the case. The value of 0 of the variable describes the road section stayed good until the survey in 2018. The method is the common solution to the right censoring problem, but it is not applicable to this research. Because, as observed in the test data, 90% are right-censored. That is to say, it is not possible to find the survivor function by the large number of right-censored data. The other solution is extending the observation time, and studying the larger amount of the population. The comparison of the survivor function of this study based on A15 ZOAB data from 2015 to 2018 and that of all ZOAB roadways in the Netherlands measured in a long term by MJPV (Rijkswaterstaat, 2019) are illustrated in Fig. 7.6. It is a derivation from this comparison that because of the right censoring problem, the case study significantly underestimates the expectation for the road survival time. The road survival time based on a large amount of data is expected to be 12.09 years, and about 10% of the ZOAB sections still maintain good performance within 30 years (Rijkswaterstaat, 2019). By the limited research time, the project does not apply the data of a longer term and more test roadways. But it is confirmed that bigger observation data can help to solve the right censoring problem, and makes the model results closer to reality.

7.6. Conclusion

The chapter applies the survival model to the case study of ZOAB pavement on A15 from the measurement of 2015 to 2018. The survival of a road section is defined as its performance concerning the road roughness and the rut depth stays good. The survival time, as the response variable of the model, is the duration from the construction time or the last repaving to the measurement date, when the bad performance is observed. By the data analysis, the three types of vehicles all have a negative influence on road health, and the influence is quantified as by how much the volumes of the vehicle class on A15 in the WNZZ or ONZ district during the observation years decreased the expected survival time. However, due to the censored data problem, the impacts of the traffic factors must be overestimated, and the solution to that is to extend the observation time or the test area. This chapter finds the answers to sub question 5b and 6b elaborated as follows.

Sub question 5b

How to apply the survival model?

The application of the survival model contains 5 steps: (1) selecting the performance indicators; (2) defining the threshold values of the performance indicators; (3) modelling the survivor function; (4) estimating how the expected failure time depends on traffic characteristics; and (5) calibration. The first two steps are to define the failure event of the pavements. The selection criterion of the performance indicators are at first, there is the measurement data of the performance indicator, and its progression at the stage before reaching the critical level and afterwards must be a discernible difference. To define the thresholds, both data analysis and empirical knowledge are considered. The outcomes from data analysis are generally more critical than the empirical knowledge. This study applied the the thresholds deriving from the data analysis because then there is more poor-performance data to analyze. After defining the failure event, the next step is modelling the survivor function, followed by the estimation of how the expected survival time depending on the traffic characteristics. Because of the censored data problem, especially left censoring and interval censoring in the case, the survival time is not the exact time point, but a survival interval. The interval is determined by the start time (T_{str}) and the end time (T_{stp}), when the failure event occurs in the period. The survivor functions are computed by the " Survivor function" module in Matlab, and the best fitting function of the distribution is found according to the max log-likelihood method. The estimation of the influence of the traffic characteristics on the expected survival time is by the coz proportional hazards model. Because the cumulative numbers of the three types of vehicle categories are taken as the predictor variables in the model, they are time-dependent. To correctly compute the coefficients, it is necessary to convert the survival data to the counting process form at first. It is the same method as processing the left or interval censored data, that is completed in the last step. The reliability of the estimation results is analyzed by the p-values. If the p-value is larger than the significant default level (in the study, 5%), the null hypothesis can not be rejected and the estimation of the coefficient makes less sense. The last step of the model application is to calibrate the case study with the findings in the previous study and the observation in reality.

Sub question 6b

What are the results of the survival model by the test data?

The study selects IRI and the rut depth as the indicators of road performance. According to the analysis of the measurement data of the A15 ZOAB sections from 2015 to 2018, when IRI of pavement is small, the annual unevenness change fluctuates around 0 m/km. When the critical value is reached, the median value of the annual change is considerable. The threshold value is 2.3 m/km in the WNZZ area and 1.9 m/km in the ONZ area. As for the development of rutting, according to the analysis of the same dataset, its annual median growth at the minimum level (the rut depth is approximately 1 mm) is considerable. After that, as the rut depth increases, the annual growth rate was flat. Up to a critical value of 10 mm, the annual increment becomes significant. Accordingly, the research defines the failure event of pavement as road roughness is greater than 2.3 m/km or the rut depth is larger than 10 mm. Hence, the survival time refers to the duration from the construction time or the time of the last maintenance to the observation of the failure event. With the processing of the survival time, the survivor function is modelled. The expected survival time of ZOAB sections concerning the roughness is 6.11 years, or 8.09 years regarding the rut depth. The best-fitting distribution of the survival time related to roughness is Weibull distribution, which indicates the probability of the critical situation occurring during the next period is higher when it serves less than 12 years approximately, and becomes low from the 16th year on.
(To be continued in next page)

Sub question 6b

What are the results of the survival model by the test data?

(The answer continues here)
The best fitting distribution of the survival time concerning the rutting is normal distribution according to the log-likelihood. But the gap between the raw data and the fitted normal distribution is considerable. That is to say, the study does not find the appropriate fit to the test data. Even though, the study still applies the findings of Weibull distribution and normal distribution to estimate how the expected survival time depends on traffic flow. The analysis results show the intensities of three vehicle categories have a negative influence on road health. The influence of each vehicle class in the different districts vary. The intensities of the passenger cars shorten the expected survival time of the roughness by 4.33% and 4.05% in two regions, and regarding the rut depth it is decreased by 3.2%, and 8.44% respectively. The effects of the number of the light trucks reduce the mean of the survival time of roughness by 52.08% and 29.69% in the WNZZ and ONZ district, and of rutting by 30.7% and 78.32% respectively. The heavy trucks decrease the life expectancy of the ZOAB pavements by 39.85% and 37.46% about the road roughness in the two areas, and as for the rutting in the WNZZ and ONZ district, the numbers are 24.24% and 76.3% respectively. Because the survival analysis in the report assumes that with the controlling variables of the climate conditions, the failure of the pavement performance is totally dependent on the traffic flow of any one vehicle class, the analysis results significantly overestimate the impact of traffic factors on road health. Besides, in the raw data, less than 10% was observed the occurrence of the failure event. The solution to calibrate the analysis is to include more observation data and consider more influence factors. The study finds the outcomes of the survival model are very data sensitive.

Apart from the regression model and the survival model, the classification model is also defined to study the effects of the traffic characteristics on road performance. The model application of the decision tree classifier is in the next chapter. It is the last part of Develop phase.

8

Decision Tree Classifier

In this chapter, the DTC algorithm is applied for predicting the future road performance and the prediction results are evaluated by using the historical data of the entire A15 from 2015 to 2018. As defined in Section 8, the model is made up of two phases: the training phase and the predicting phase. Accordingly, the chapter is organised in the way that Section 8.1 presents the training phase, and Section 8.2 shows the predicting phase. The end of the chapter is the conclusions and the answers to two sub questions of SQ. 5c how to apply the decision-tree classifier model and SQ. 6c what are the results of the decision-tree classifier model by the test data, in Section 8.3.

8.1. Training phase

According to the model framework as illustrated in Fig. 3.3, the input data of the machine learning contains two sets. One is to run the model, and the other is to test the accuracy of the model. Thus, it is necessary to process the historical data to the individual tables. From a complete data set, part of the data is used to build the model, and the remaining is used to train the model. There is no definition of the split share of the data. To get a good model, in general, the number of the data to run the model is larger than that for testing the model. In the case, the measurement data of 2015, 2016 and 2017 is for the establishment of the decision-tree model, while the measurement data of 2018 are to test it. The data used for the case study is originated from IVON of RWS, which consists of 8484 rows and 29 columns whereby every row represents information of a 100-meter road section and the columns provide the measurement results of the section. All of the information is retrieved from the database of IVON system and saved as excel format.

DTC algorithm is programmed in Python with the support of "Scikit-learn" module. The code of DTC is in Appendix. J. The program of the training phase is made up of five parts. The first one is calling the relevant modules the model uses, including pandas, numpy, sklearn, matplotlib, seaborn, graphaviz, pydotplus, io, scipy, and IPython. Among these, the library of pandas and numpy are the basic ones in Python. The library of sklearn is the core of building up the decision tree. The others are all related to plotting the outcomes. The second is loading the two data tables that have been prepared. The data format in the study is in the excel table. But for Python, the other forms of data, like txt, json are also readable. The sklearn module in Python requires no missing values of all the cells in the input tables, thus the study excludes those rows which have missing data and 92.5% data is used. The third is giving the label of each column of the data. The model will then use the label as the name of the variable, and output the decision criteria as a label is greater or smaller than a value. In the study, the predictor variables of the model contain the service time of the pavement (AGE_IRI, AGE_RUT), the material of the surface (SURFACE_COMBID, SURFACE_DAB, SURFACE_EAB, SURFACE_OAB, SURFACE_SMA, SURFACE_ZOAB, SURFACE_ZOAB+, SURFACE_ZOABTW, SURFACE_ZOEAB), the initial value of the performance indicator (IRI_VALUE_0, RUT_VALUE_0), the cumulative traffic flow (I_L1, I_L2, and I_L3]), and the climate data (T_TEMP_25, T_TEMP_0, T_TEMP_0_below, T_TPERCIPITATION). The response variables are the classification of IRI (R_IRI) and the rut depth (R_RUT). If the road roughness or rutting meets the maintenance requirement defined by RWS as Tab. 1.1 indicates, the variable has the value of 1; otherwise, 0. The fourth module in the model application is to call DTC function

Table 8.1: Confusion matrix of DTC prediction of A15 roughness in 2018

N = 2786	Prediction: 0	Prediction: 1
Actual: 0	2695	56
Actual: 1	11	24

(a) Minimum sample split=10

N = 2786	Prediction: 0	Prediction: 1
Actual: 0	2685	66
Actual: 1	7	28

(b) Minimum sample split=20

N = 2786	Prediction: 0	Prediction: 1
Actual: 0	2685	66
Actual: 1	7	28

(c) Minimum sample split=30

N = 2786	Prediction: 0	Prediction: 1
Actual: 0	2723	28
Actual: 1	8	27

(d) Minimum sample split=40

installed in the module of sklearn. The last is plotting the decision tree as well as the prediction accuracy that computed according to the comparison of the prediction results by DTC and the real measurement.

The establishment of DTC is according to the minimum sample split, which is the number of minimum samples required on the node to decide of node whether to continue splitting process or stop on that node. The decrease in minimum samples split increases the risk of overfitting. The lower minimum samples split makes the tree grow to be more fit for the train data. The minimum value of the minimum sample split is the number of the smallest class in the sample. It is significant to estimate the appropriate minimum sample split. Apart from the minimum sample split, in the algorithm of DTC, the random state is required to define. Setting the random state to a certain number ensures that the permutation of features used in building the trees will always be the same. It does not matter which number is chosen as the random state. Fixing the random state is useful as it facilitates in reproducing the experiments with the same results. In the study, the candidate minimum sample splits are 10, 20, 30, and 40, and the setting of the random state is 0. The outcomes of four trained decision tree models of the road roughness on A15 run by the different settings of the minimum sample splits are illustrated in Fig 8.1, Fig. 8.2, Fig. 8.3, and Fig. 8.4 respectively. Decision tree models with different minimum samples split have different tree shapes. When the minimum number of samples required for a split is lower, the resulting decision tree is deeper, because more splits, more branches are needed.

To determine the appropriate minimum sample split, the evaluation of how accurately the DTCs derived from the different settings of minimum sample splits can predict the testing data are compared. There are two kinds of accuracy to be explored in order to evaluate the accuracy of the predictions made by the model: the overall accuracy rate and the class accuracy number. The overall accuracy rate is the ratio of the correct number of the prediction of the class and the total number of the prediction results. The class accuracy number is based on the quantity of the correct prediction among the test data from a specific class. The problem investigated in this research is a binary classification problem with the two classes ”0” and ”1”. The value of 0 of the target variable indicates the road section has not met the maintenance requirement regarding to road roughness at the determined prediction time, and the other class with the value of 1 is that the road section has met the interval level concerning the surface roughness. There are two kinds of correction of the model prediction, as well as two types of errors. Two types of correct predictions are that roads that shall not meet the maintenance requirements in 2018 by the prediction of DTC did not meet the maintenance requirements based on the actual measurements in 2018, and the road that shall meet the maintenance requirements in 2018 predicted by DTC did meet the maintenance requirements according to the actual measurement in 2018. There are two scenarios of the prediction mistakes. One is that the prediction of DTC gives the result that the pavement section shall meet the maintenance requirements in 2018, but the measurement data of the road performance indicated the good condition. The other is that DTC predicts that the road section shall not meet the maintenance requirements in 2018, but the actual measurement does not support the prediction. In the case, it is more significant that the road section shall be intervened but judged as good by the model. Thus, the key indexes to evaluate the model prediction accuracy and determine the appropriate value of the minimum sample split are the overall accuracy rate and the class accuracy number of the class of "1" that is mistakenly predicted as "0". The model prediction accuracy of the different settings of the minimum sample split about the road roughness according to A15 data from 2015 to 2018 is listed in Tab. 8.1.

According to the comparison of the prediction results of the training models and the actual measurement

Table 8.2: Confusion matrix of DTC prediction of A15 rut depth in 2018

N = 2786	Prediction: 0	Prediction: 1
Actual: 0	2448	286
Actual: 1	8	45

(a) Minimum sample split=10

N = 2786	Prediction: 0	Prediction: 1
Actual: 0	2410	324
Actual: 1	4	49

(b) Minimum sample split=20

N = 2786	Prediction: 0	Prediction: 1
Actual: 0	2413	321
Actual: 1	3	50

(c) Minimum sample split=30

N = 2786	Prediction: 0	Prediction: 1
Actual: 0	2484	250
Actual: 1	3	50

(d) Minimum sample split=40

data of 2018, the overall accuracy rates derived from the different settings of the minimum sample split are all above 97%, which indicates a rather high model prediction accuracy. It confirms that DTC is an applicable model to predict the road performance of the case. Based on the class accuracy number in Tab. 8.1, the setting of the minimum sample split as 20 or 30 results in the minimum errors of recognizing the class "1" to the class "0", meaning the least ignorance of the bad conditions. Compared the decision tree modelled by setting 30 as the minimum sample split in Fig. 8.3, the outcome derived by setting 20 as the minimum sample split Fig. 8.2 illustrates have more branches to classify the class "0" into smaller categories, which is not the concern of the study. Therefore, the appropriate setting of the parameter of this case is 30. Accordingly, the predicting phase will use DTC with the setting of the minimum sample split as 30. The overall accuracy rate of the model regarding the road roughness is 97.4%.

The same methodology is applied to find the appropriate setting of the parameter in DTC models to predict the rut depth. The candidate settings are 10, 20, 30, and 40. The random state is always set as 0. Each input results in the different decision trees, as Fig. 8.5, Fig. 8.6, Fig. 8.7, and Fig. 8.8 respectively. Similar to the model results about the road roughness, the decision trees have a slight variance between the shapes with the different minimum samples split. When the minimum number of samples required for a split is lower, the resulting decision tree is deeper, and vice verse. The overall accuracy rates and the class accuracy numbers as elaborated above are computed as Tab. 8.2 shows.

The general prediction of DTC model on the road rutting is more than 88% accurate in the case study. The majority of errors are mistakenly predicting the actual class of "0" as the class of "1". Among 2787 prediction results, not larger than 8 road sections are wrongly assessed as the good condition when the real situation is opposite, under 4 testing sets of the minimum sample split. Although there is about 10% pavement staying good but is judged to be in need of repair by the model, it might lead to a more cautious measurement strategy. The focus of the study is on the prediction results of the class of "1".

The appropriate value of the minimum sample split is determined according to the overall accuracy rate and the class accuracy number of the class of "1" that is mistakenly predicted as "0". Based on the computation results in Tab.8.2, there are the similar overall accuracy results for the four settings, and either the setting of the minimum sample split as 30 or 40 has the least errors of mistaking the class "1" into the class "0". Since the setting of 40 has the high overall accuracy rate, the study finds it is the appropriate setting of the parameter. Accordingly, the DTC model to predict the rutting in the case study is illustrated in Fig. 8.8, and the predicting phase will use it. The overall accuracy rate of the model regarding the road rutting is 90.9%.

8.2. Predicting phase

According to the outcomes of the training phase in the last section, the DTC models trained by the measurement data of 2015, 2016 and 2017 for predicting the road roughness and the rutting in 2018 have been found. The appropriate settings of the minimum sample split has been determined (30 for modeling the roughness, 40 for modeling the rutting). In the prediction phase, the measurement data of 2015, 2016, 2017 and 2018 is applied as the training data to establish the DTC model, and the data on the period from the measurement survey of 2018 (mostly on the date of May 15th, 2018) to January 1st, 2019 is used as the test data. Based on the training model and the test data, the prediction of the road performance at the date of January 1st, 2019 is given as the consequence. The list of the pavement which meets the maintenance requirement according to the prediction will be delivered to RWS for getting noticed in the next measurement.

The general program process of the predicting phase contains 5 parts. The first one is calling 10 Python modules the model uses, including pandas, numpy, sklearn, matplotlib, seaborn, graphaviz, pydotplus,io, scipy, and IPython, where sklearn is the core of building up the decision tree. The second is loading the two data tables that have been prepared. In this study, the measurement data of 2015, 2016, 2017 and 2018 are applied as the training data to establish the DTC model, and the data on the period from the measurement survey of 2018 to January 1st, 2019 is used as the test data. The third is to name each column of the data. In the case, the training data and the test data in the predicting phase have the different contents. The training data contains the predictor variables including the service time of the pavement (AGE_IRI, AGE_RUT), the material of the surface (SURFACE_COMBID, SURFACE_DAB, SURFACE_EAB, SURFACE_OAB, SURFACE_SMA, SURFACE_ZOAB, SURFACE_ZOAB+, SURFACE_ZOABTW, SURFACE_ZOEAB), the initial value of the performance indicator (IRI_VALUE_0, RUT_VALUE_0), the cumulative traffic flow (I_L1, I_L2, and I_L3]), and the climate data (T_TEMP_25, T_TEMP_0, T_TEMP_0_below, T_TPERCIPITATION), and the response variables that are the classification of IRI (R_IRI) and the rut depth (R_RUT). But the test data only include the predictor variables, and the response variables are the prediction outcomes. If the road roughness or rutting meets the maintenance requirement defined by RWS (IRI > 2.6 m/km, the rut depth > 10 mm), the response variable has the value of 1; otherwise, 0. The fourth part in the model application is to establish DTC by the setting of the minimum sample split determined in the training phase and DTC function of the module of sklearn. The last is outputting the road sections which would meet the maintenance requirement by the model prediction as well as its characteristics.

The illustration of the model of road roughness is in Fig. 8.9 The common features of the failed road sections are inducted into three condition sets as follows.

Condition set I

- The value of IRI in the last survey is not more than 2.135 m/km.
- The cumulative traffic volumes of the heavy trucks are more than 73622230 vehicles.

Condition set II

- The value of IRI in the last survey is larger than 2.135 m/km, but not more than 2.435 m/km.
- The cumulative time when the temperature is over 25 °C at the location of the year does not exceed 194.767 hours.
- The total service time of the pavement is longer than 19.64 years.

Condition set III

- The value of IRI in the last survey is above 2.58 m/km.
- The cumulative rainy time at the location of the year is not longer than 1638.178 hours.
- The cumulative traffic volumes of the heavy trucks are more than 43825 vehicles.
- The total service time of the pavement is longer than 5.69 years, or less than 3.65 years.

According to the prediction by DTC established in the case study, up to January 1st, 2019, there were 66 road sections on A15 that met the conditions, where 1 road section in the condition set 1, 0 in the condition set 2, and 65 in the condition set 3. These road sections as well as their characteristics are listed in Appendix K. From the model results, it can be derived that the initial value of IRI is an important criterion to classify the road performance class regarding the road roughness. The effects of the traffic factors found by the model are that when the measurement of the road section in the previous year was less than 2.135, the road sections with the very high cumulative intensities of the heavy trucks were supposed to be in a bad condition regarding the roughness in the next year. When the unevenness value measured in the previous year was close to the intervention level, even with a slight value of the cumulative number of the heavy vehicles, the road section can be predicted to fail by the model found in the case study. The findings are based on the training data of A15 from 2015 to 2018. The report can not confirm if the same finding shall be achieved by the other input data set. For example, the decision trees on the dependent of A15 from 2015 to 2017 do not show the importance of the traffic intensities of the heavy trucks to the classification of the road roughness, but the volumes of the light trucks and the passenger cars play some roles, as Fig. 8.1, Fig. 8.2, Fig. 8.3, and Fig. 8.4 illustrate. In addition, the high temperature (over 25 ° C), rain, and the total service time are also selected as the classification conditions in the different condition sets according to the case study.

The DTC of the road rut depth is modelled with the same input of the training data and the test data. The setting of the minimum sample split is 40, as the finding of the training phase above. The random state is set as 0. The model results are drawn in Fig. 8.10, and the pavement which is predicted to meet the maintenance requirement concerning the rut depth (in the study, the rut depth is above 10 mm) is included in Appendix. K. The outcome is delivered to RWS so that these road sections will get plenty of attention in the next measurement. The features of the road sections which met the maintenance requirements are categorised into 12 condition sets in the following.

Condition set I

- The surface material can be ZOAB, ZOAB+, COMBID, DAB, EAB, OAB, or SMA.
- The value of the rut depth in the last survey is not more than 7.075 mm.
- The cumulative time when the temperature is over 25 °C at the location of the year exceeds 215.635 hours.
- The total service time of the pavement is not longer than 1.78 years.

Condition set II

- The surface material can be ZOAB, ZOAB+, COMBID, DAB, EAB, OAB, or SMA.
- The value of the rut depth in the last survey is between 6.035 mm and 7.075 mm.
- The cumulative time when the temperature is over 25 °C at the location of the year does not exceed 215.635 hours, but longer than 162.033 hours.
- The cumulative traffic volumes of the light trucks are not less than 4279932 vehicles, but the cumulative number of the passenger cars is not bigger than 53343516.

Condition set III

- The surface material can be ZOAB, ZOAB+, COMBID, DAB, EAB, OAB, or SMA.
- The value of the rut depth in the last survey is between 7.075 mm and 9.025 mm.
- The cumulative traffic volumes of the light trucks are more than 39399850 vehicles.

Condition set IV

- The surface material type is DAB.
- The value of the rut depth in the last survey is between 7.09 mm and 9.025 mm.
- The cumulative time when the temperature is over 25 °C at the location of the year exceeds 192.218 hours.
- The cumulative traffic volumes of the light trucks do not exceed 39399850 vehicles.
- The total service time of the pavement is longer than 8.69 years.

Condition set V

- The surface material type is ZOABTW.
- The value of the rut depth in the last survey is not larger than 9.025 mm.
- The cumulative time when the temperature is over 25 °C at the location of the year exceeds 191.255 hours.
- The total service time of the pavement is longer than 3.02 years.

Condition set VI

- The surface material type is ZOEAB.

- The value of the rut depth in the last survey is not larger than 9.025 mm.
- The cumulative time when the temperature is over 25 °C at the location of the year is longer than 194.767 hours.

Condition set VII

- The value of the rut depth in the last survey is larger than 9.91 mm, but not reach 10.06 mm.
- The cumulative time when the temperature is over 25 °C at the location of the year is between 131.375 hours and 194.767 hours.

Condition set VIII

- The value of the rut depth in the last survey exceeds 10.06 mm.
- The cumulative time when the temperature is over 25 °C at the location of the year is not more than 162.033 hours, and the cumulative rainy time at the location of the year is longer than 1842.258 hours.
- The total service time of the pavement does not exceed 4.83 years.

Condition set IX

- The surface material can be ZOAB, ZOAB+, COMBID, EAB, OAB, SMA, ZOABTW, or ZOEAB.
- The value of the rut depth in the last survey exceeds 9.025 mm.
- The cumulative time when the temperature is over 25 °C at the location of the year is more than 194.767 hours.

Condition set X

- The value of the rut depth in the last survey exceeds 10.06 mm.
- The cumulative time when the temperature is over 25 °C at the location of the year is not more than 194.767 hours, and the cumulative time when the temperature is below 0 °C at the location of the year is less than 535.271 hours.
- The total service time of the pavement exceeds 4.83 years.

Condition set XI

- The surface material type is SMA.
- The value of the rut depth in the last survey exceeds 10.06 mm.
- The cumulative time when the temperature is over 25 °C at the location of the year is not more than 194.767 hours. But the cumulative time when the temperature is below 0 °C at the location of the year is more than 535.271 hours.
- The cumulative traffic volumes of the heavy trucks do not exceed 28024675 vehicles.
- The total service time of the pavement exceeds 4.83 years.

Condition set XII

- The surface material can be ZOAB, ZOAB+, COMBID, DAB, EAB, OAB, ZOABTW, or ZOEAB.
- The value of the rut depth in the last survey exceeds 10.06 mm.
- The cumulative time when the temperature is over 25 °C at the location of the year is not more than 194.767 hours. But the cumulative time when the temperature is below 0 °C at the location of the year is more than 535.271 hours.
- The cumulative traffic volumes of the light trucks and the heavy trucks do not exceed 8351482, and 28024675 vehicles respectively. But the traffic flow of the passenger cars are more than the number of 49586086.
- The total service time of the pavement exceeds 4.83 years.

The model predicts 259 road sections would meet the maintenance requirement concerning rutting on January 1st, 2019, the majority of which is in the condition set IX. There are 0, 9, 2, 0, 33, 29, 17, 1, 101, 50, 1, and 16 road sections in the condition set from 1 to 12 respectively as assessed to fail by the model, which are listed in Appendix K. The significant classification factors of the rutting are, the material, the initial value of the rutting at the last measurement, the hot and freezing weather, the traffic intensities, and the whole serving time. As for the effects of the traffic-associated influencing factors, among 12 condition sets listed above, 5 of them have the traffic flow as one of the decision criteria, which are the condition set II, III, IV, XI and XII. In the various condition sets, the critical values of traffic flow are different. In the five condition sets that select the traffic flow as part of the decision criterion, the critical values of traffic flow have a range between 4 million and 50 million, which are considerable numbers.

8.3. Conclusion

The chapter applies the DTC model to predict which A15 sections would meet the maintenance requirements concerning the road performance about the roughness and the rutting on January 1, 2019. The input data for training the model is the measurement data from 2015 to 2018, in the case. To estimate the appropriate setting of the key parameter for establishing the model, the minimum sample split, the training phase is executed at first. It determines the setting of the parameter by the prediction accuracy, that is computed according to the prediction result of the road performance in 2018 by the model based on the measurement data of 2015, 2016 and 2017 and the actual observation of A15 performance in 2018. The appropriate setting of the minimum sample split meets two criteria. One is that the overall accuracy rate is high, and the other is the class error number, especially the actual class of 1 that is mistakenly predicted as 0, is less. Accordingly, the appropriate settings of the parameters to predict the road roughness class and the rutting class are 30, and 40 respectively. The DTC modelled by the appropriate setting of the minimum sample split and the measurement data of A15 from 2015 to 2018 indicates the common features of all A15 sections which have the severe roughness or the rut depth where the maintenance requirements are reached. The features are classified into 15 sets. The significant classification factors of the roughness are the initial value of the roughness at the last measurement, the hot and rainy weather, the traffic intensities, and the whole serving time. As for the rutting, they contain the material, the initial value of the rutting at the last measurement, the hot and freezing weather, the traffic intensities, and the whole serving time. By the prediction of the model, 66 road sections of A15 would meet the intervention level of roughness (IRI > 2.6 m/km) on January 1, 2019, and 259 road sections of A15 would meet the intervention level of rutting (rut depth > 10 mm) at that date. In addition, this chapter finds the answers to sub question 5c and 6c in the following.

Sub question 5c

How to apply the decision tree classifier model?

The model application of the decision tree classifier contains 2 phases, the training phase and the prediction phase. The training phase is to test the feasibility of the model for the case study, in which the appropriate setting of the key parameter to establish the model and the prediction accuracy of the model are estimated. The program of the training phase consists of 5 parts. Firstly, the code calls the relevant modules the DTC model requires, including pandas, numpy, sklearn, matplotlib, seaborn, graphaviz, pydotplus,io, scipy, and IPython. Then, two data tables are loaded, the training data and the test data. The input data in the study is in the form of the excel tables. It should be noticed that the sklearn module in Python requires no missing values of all the cells in the input tables, thus the rows which have missing data are excluded in advance. In the case study, the training data contains all the measurement data of 2015, 2016, and 2017, while the test data includes the measurement data of 2018. Both of the datasets have the predictor variables are the response variables. The predictor variables defined in the research contains the service time of the pavement, the material of the surface, the measurement value of the performance indicators at the last survey, the cumulative traffic flow of three vehicle classes, and the climate data. The response variables are the classification of road performance and the rut depth, which have the value of 1 if the road roughness or rutting meets the maintenance requirement; otherwise, 0.
(To be continued in next page)

Sub question 5c

How to apply the decision tree classifier model?

(The answer continues here)
The third part of the model application of the training phase is assigning the labels to each column of the data, including AGE_IRI, AGE_RUT for the service time of the pavement; SURFACE_COMBID, SURFACE_DAB,SURFACE_EAB, SURFACE_OAB, SURFACE_SMA, SURFACE_ZOAB, SURFACE_ZOAB+, SURFACE_ZOABTW, SURFACE_ZOEAB for the surface material; IRI_VALUE_0 and RUT_VALUE_0 indicating the measurement value of the performance indicators at the last survey; I_L1, I_L2, and I_L3 for the cumulative traffic flow of passenger cars, light trucks and heavy trucks respectively; T_TEMP_25, T_TEMP_0, T_TEMP_0_below,T_TPERCIPITATION describing the climate occasions; and R_IRI and R_RUT for the classification of IRI and the rut depth. The fourth part in the training phase program is modeling the DTC with the different settings of the minimum sample split, which determines how detailed the decision tree classifies the data. At last, the visualisation of DTC as well as the prediction accuracy computed by the comparison of the prediction results of 2018 and the actual measurement data in the year.
The predicting phase follows, which aims to predict the classification of the road performance of the pavement on the determined date. The estimation of the minimum sample splits in the training phase is the input for the prediction phase. The program of this phase consists of 5 parts, where the first and the third modules are identical to the ones in the training phase, but there are some differences in the other three parts. In the second step of the program of the predicting phase, the training data and the test data are still loaded, but they contain different variables. The training data table has both the predictor variables and the response variables as defined as the same in the training phase, but the test table has the predictor variables only, while the response variables are the outputs of the model. The fourth module in the predicting phase applies the DTC with the settings of the minimum sample split determined by the training phase. The last part is to output all the road sections on A15 which would meet the maintenance requirements regarding the road roughness and the rutting by the model prediction at a determined time.

Sub question 6c

What are the results of the decision tree classifier model by the test data?

The prediction of the road performance of A15 up to January 1, 2019 by the DTC model shows that 66 road sections of A15 would meet the intervention level of roughness (IRI > 2.6 m/km) and 259 road sections of A15 would meet the intervention level of rutting (rut depth > 10 mm) at that date. It indicates the common features of A15 sections where the severe roughness or the rut depth were observed during the years from 2015 to 2018. The common features are classified into 15 condition sets. The significant classification factors of the roughness are the initial value of the roughness at the last measurement, the hot and rainy weather, the traffic intensities, and the whole serving time. As for the rutting, they contain the material, the initial value of the rutting at the last measurement, the hot and freezing weather, the traffic intensities, and the whole serving time.
Comparing the DTC predicting results of the road performance in 2018 and the actual measurement data in the year, 97.6% of the A15 sections' road roughness class are predicted correctly with the setting of 10 as the minimum sample split, and the overall prediction accuracy are 97.4%, 97.4%, and 98.7% when setting the minimum sample split as 20, 30 and 40 respectively. In the study, the road roughness class is either 1 where IRI of the road section is measured above 2.6 m/km or 0 where the value of IRI of the pavement is not more than the critical number. As for the prediction of the rutting, 89.5%, 88.2%, 88.4% and 90.9% of the A15 sections' road rutting class are predicted correctly with the settings of 10, 20, 30 and 40 as the minimum sample split respectively. The rutting class of 1 indicates the road section meets the maintenance requirement concerning the rut depth (bigger than 10 mm, defined by RWS); while 0 means good condition with regards to rutting.
(To be continued in next page)

Sub question 6c

What are the results of the decision tree classifier model by the test data?

(The answer continues here)
The high prediction accuracy confirms the feasibility of the DTC model to the project. In addition, according to the class accuracy number, the appropriate setting of the parameters to predict the road roughness class and the rutting is 30, and 40 respectively of the case.
DTC model confirms the important influence of the traffic intensities on the road performance (roughness and rutting), since the traffic flow is selected as one of the conditions to predict whether the road will meet the maintenance requirements in multiple decision trees. Among three condition sets to classify the road performance, two have the traffic flow as the decision criteria, while about the rutting, 5 out of 12 condition sets select the total traffic volumes as the classification criteria. The way that DTC model quantifies the effects of the traffic intensities on the road performance is transitional. That is to say, the model assumes only when the traffic flow reaches a certain value, it has an impact on the road performance.
In the case study, when the measurement of the road roughness in the previous year was less than 2.135 m/km, the road sections with the cumulative intensities of the heavy trucks above 73622230 vehicles were predicted to be in bad condition regarding the roughness in the next year. When the unevenness value measured in the previous year was close to the intervention level, even with a slight value of the cumulative number of the heavy vehicles that are more than 43825 vehicles, the road section would be predicted to fail. As for the effects of the traffic factors on the road rutting the model finds, the critical numbers of the traffic intensities in the multiple decision trees are different, that vary from 4 million to 50 million. Moreover, in the multiple decision trees the study runs, the types of traffic flow (passenger cars, light trucks, and heavy trucks) selected as the conditions for predicting whether roads meet maintenance requirements are different. Therefore, the critical level of the traffic volumes of each vehicle class are very sample-data sensitive, and can change when inputting the different period of the measurement data and the minimum sample split.

The chapter is the end of Develop phase, which focuses on the case study of the decision tree classifier model. The model application and results of the regression models and the survival analysis in the front part of Develop phase as well as DTC in this chapter will be compared in the following, Deliver phase. The most applicable model among the three will be applied to design a simulation tool in the next part of the thesis.

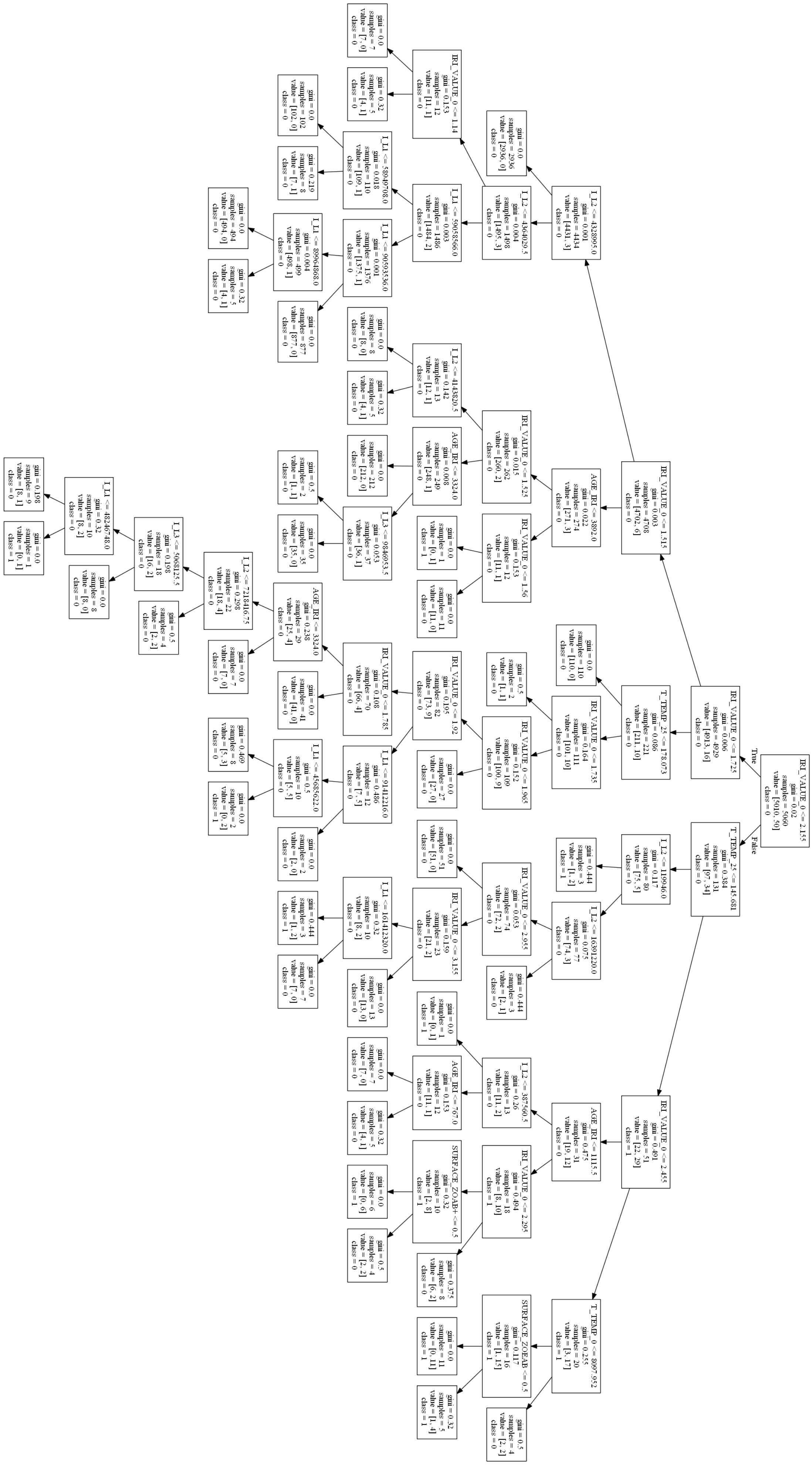


Figure 8.1: Decision tree classifier of road roughness trained by A15 data from 2015 to 2017 (Minimum sample split=10)

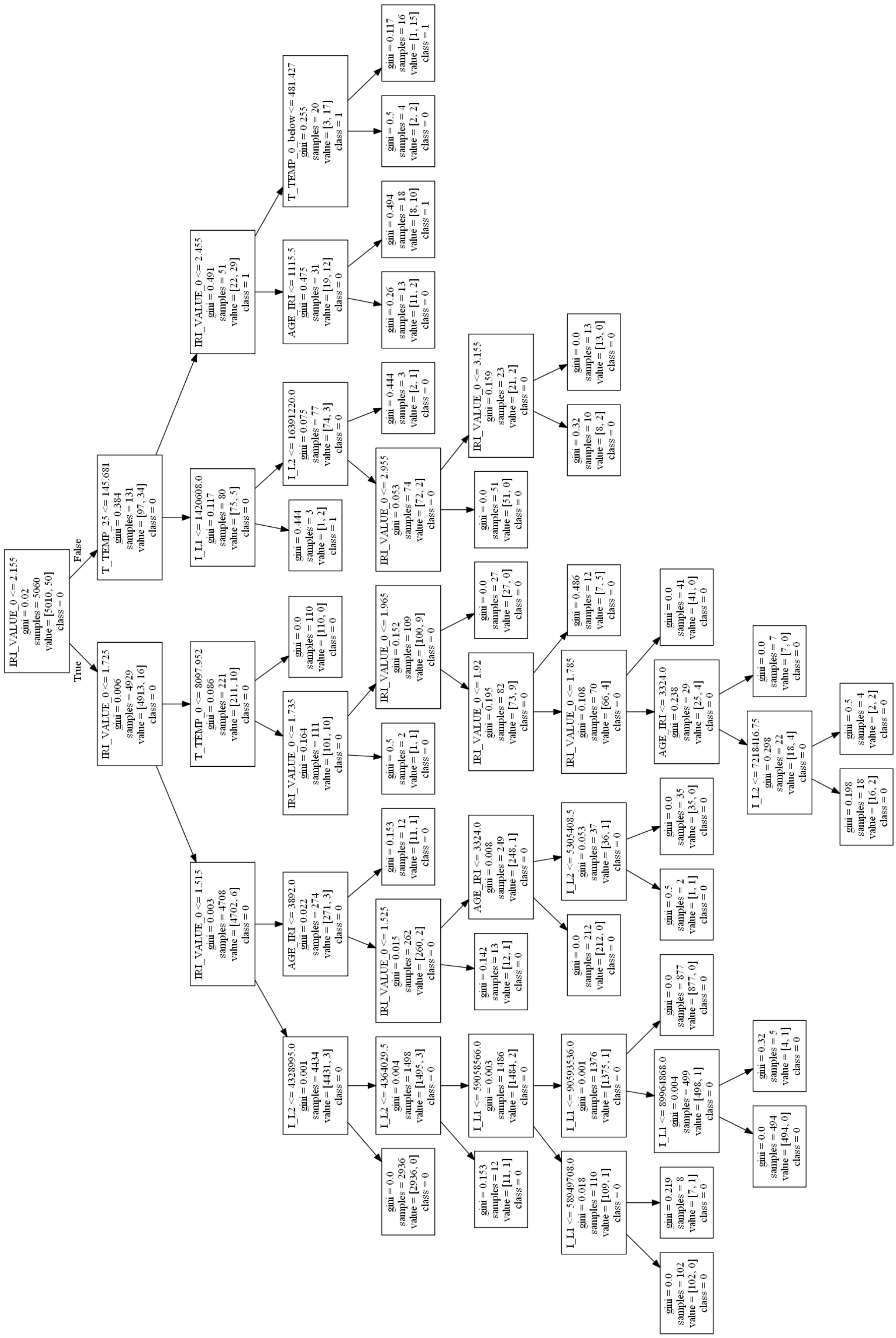


Figure 8.2: Decision tree classifier of road roughness trained by A15 data from 2015 to 2017 (Minimum sample split=20)

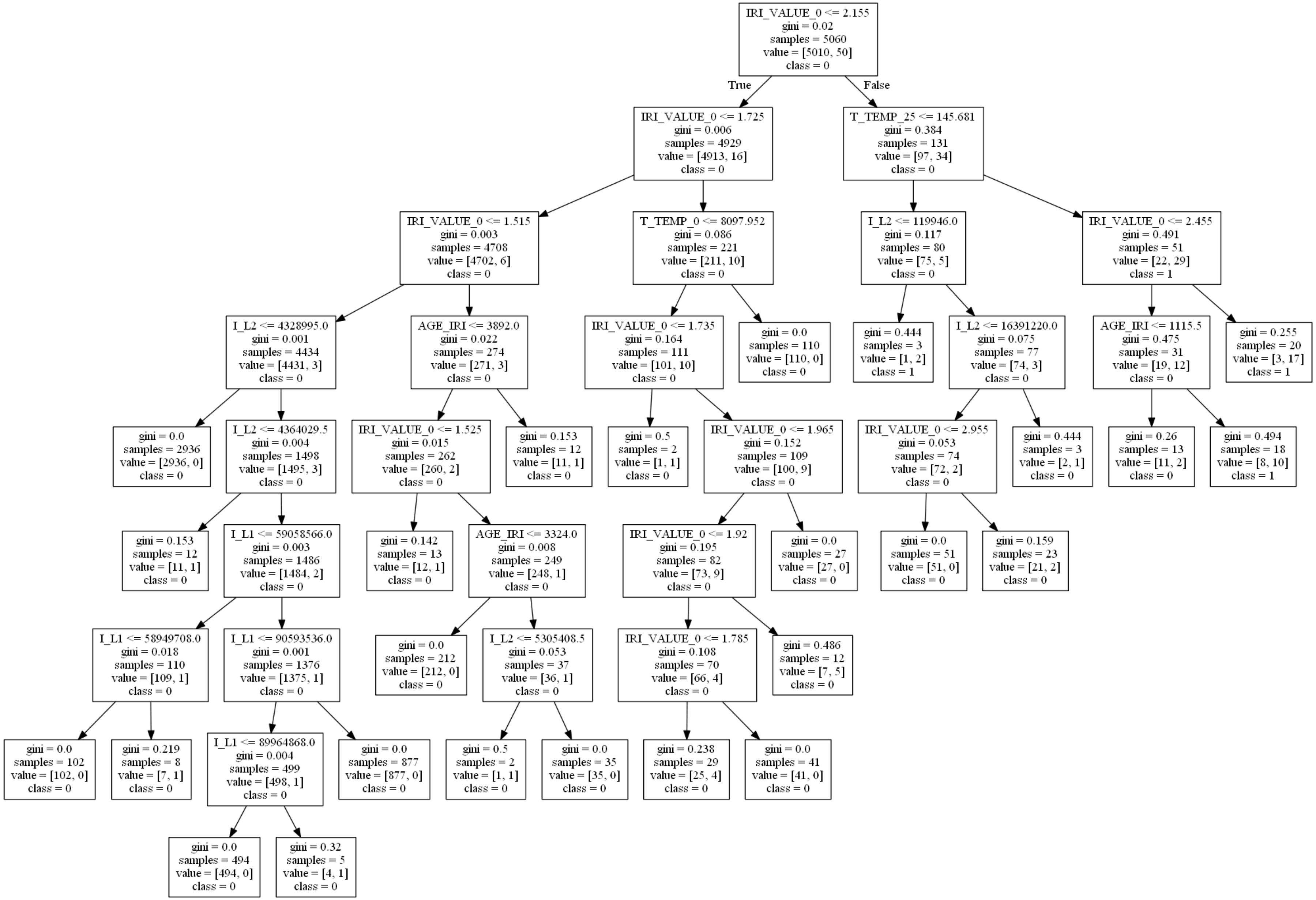


Figure 8.3: Decision tree classifier of road roughness trained by A15 data from 2015 to 2017 (Minimum sample split=30)

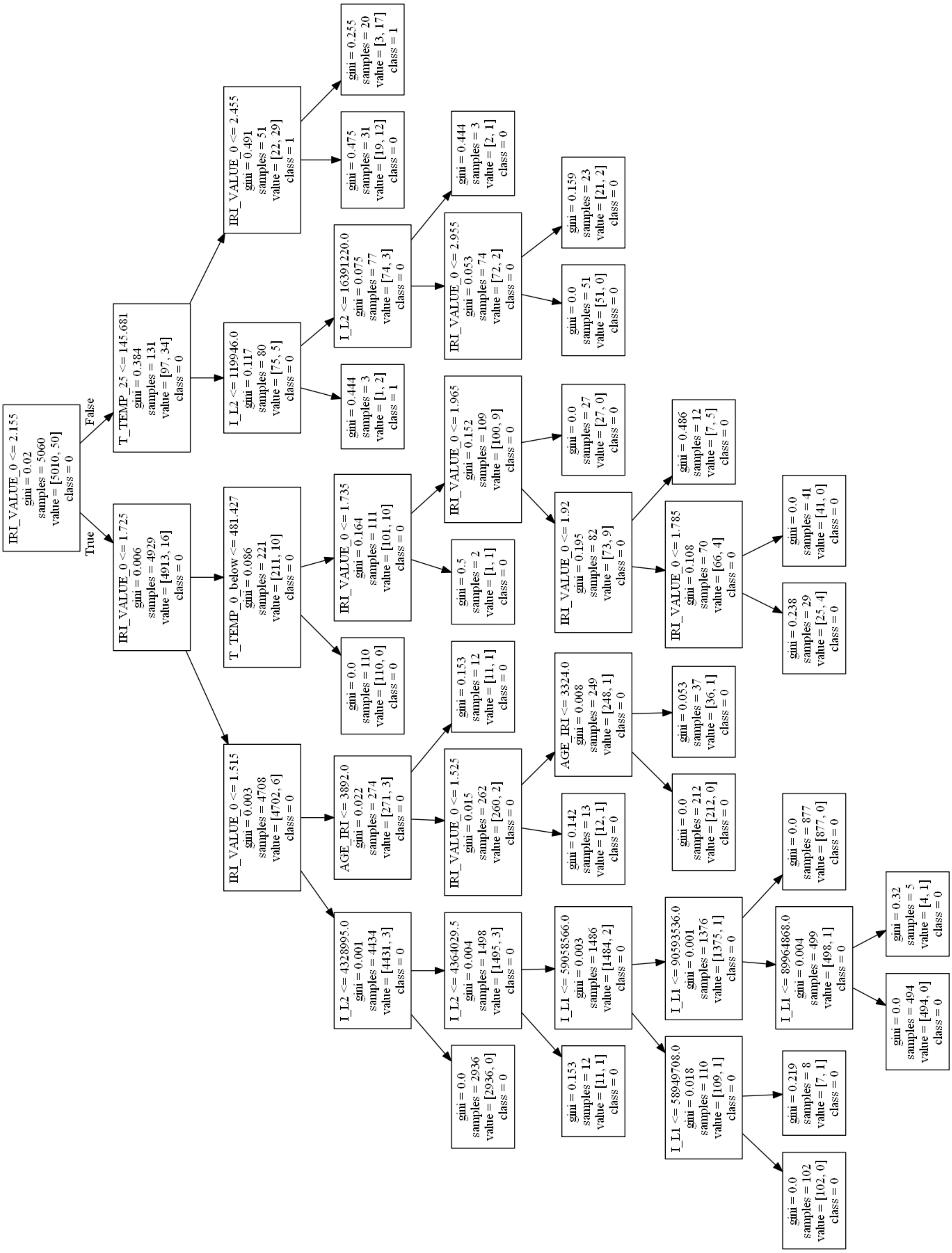


Figure 8.4: Decision tree classifier of road roughness trained by A15 data from 2015 to 2017 (Minimum sample split=40)

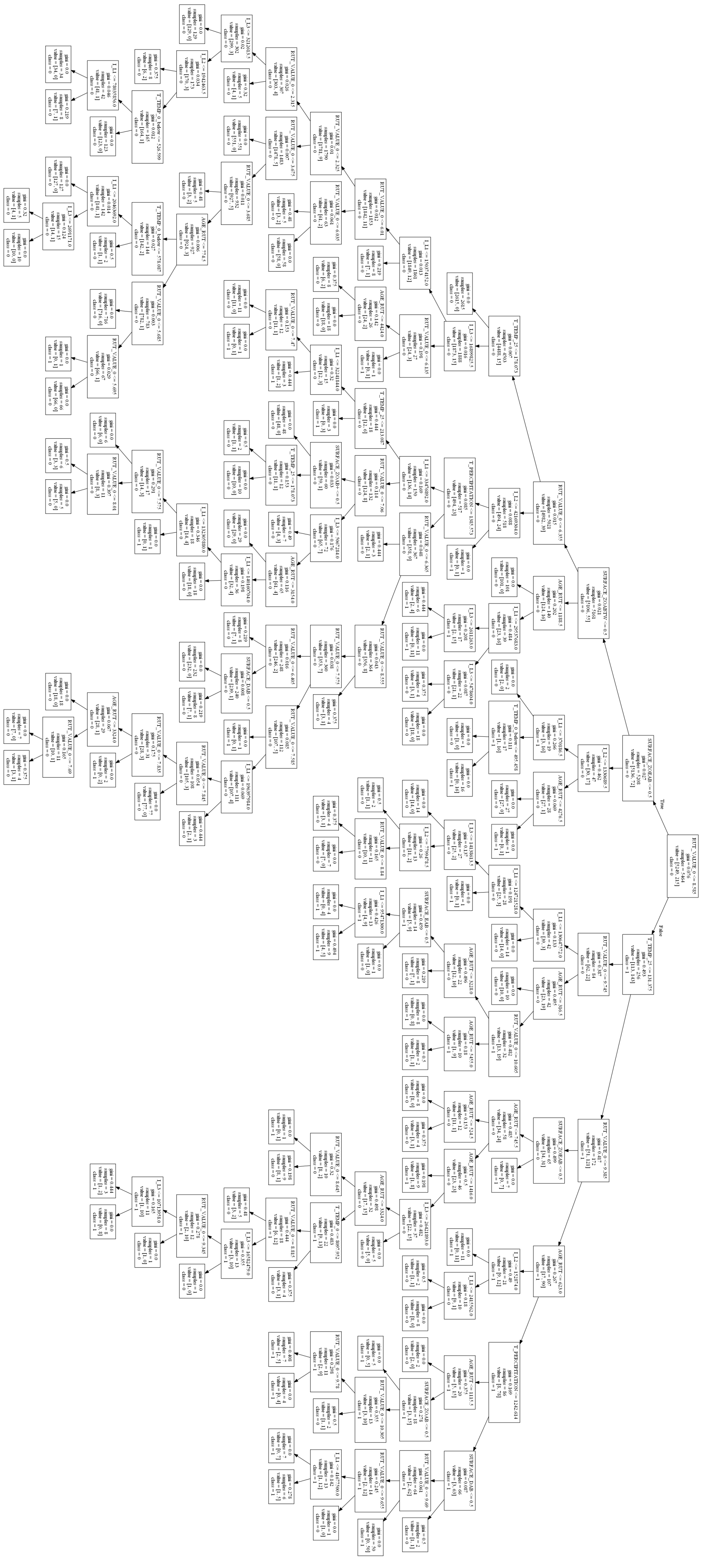


Figure 8.5: Decision tree classifier of road rutting trained by A15 data from 2015 to 2017 (Minimum sample split=10)

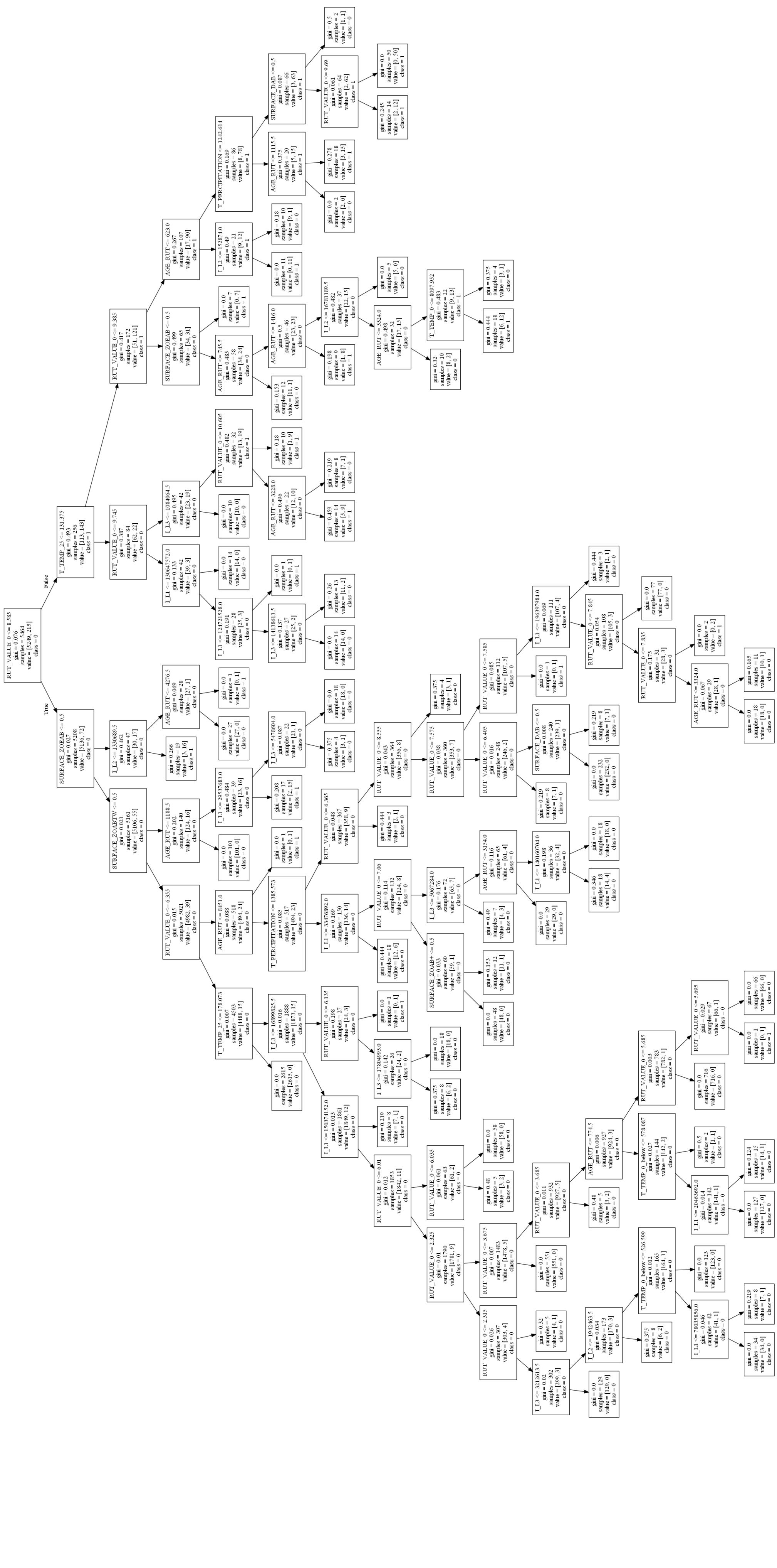


Figure 8.6: Decision tree classifier of road rutting trained by A15 data from 2015 to 2017 (Minimum sample split=20)

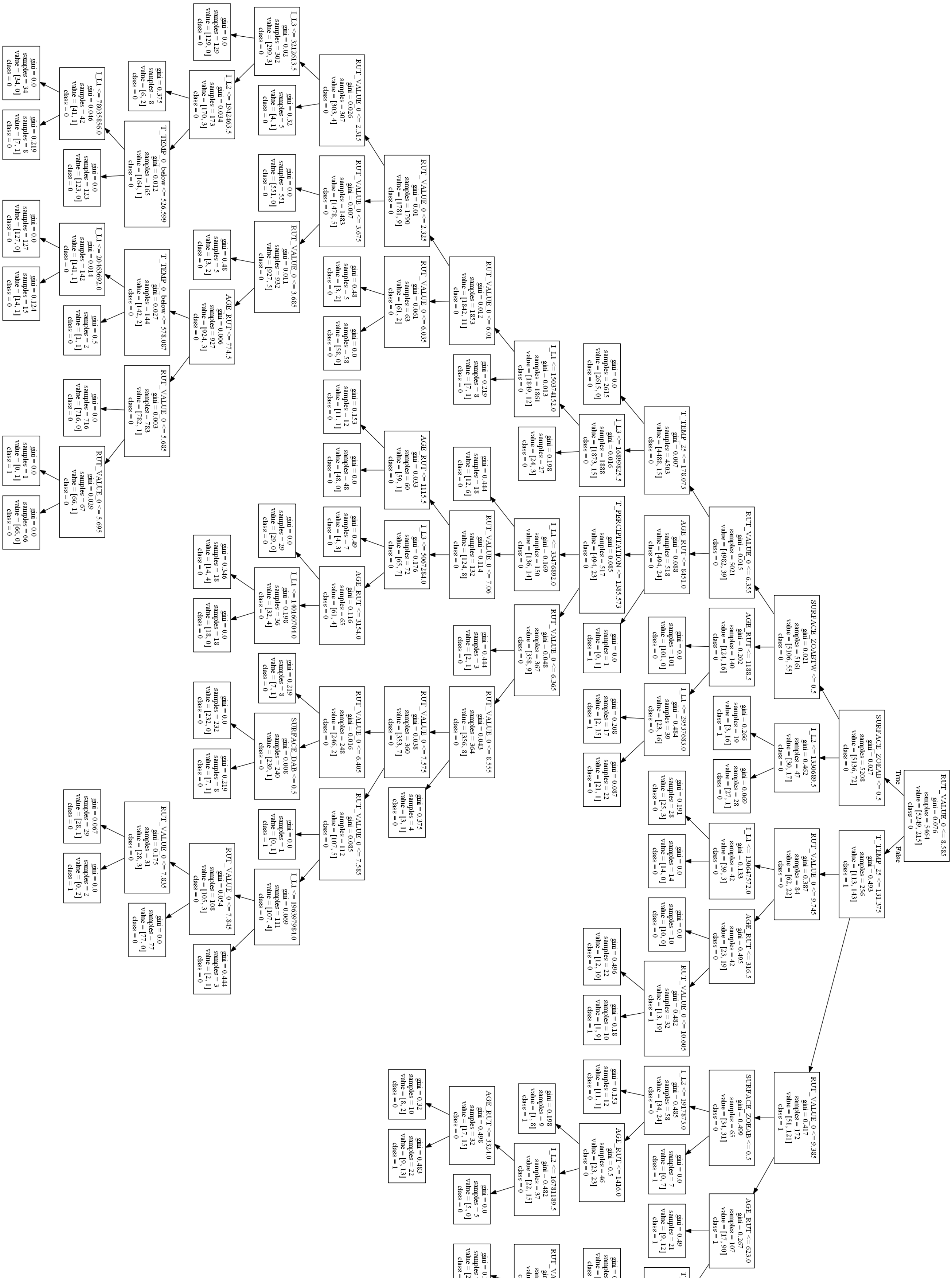


Figure 8.7: Decision tree classifier of road rutting trained by A15 data from 2015 to 2017 (Minimum sample split=30)

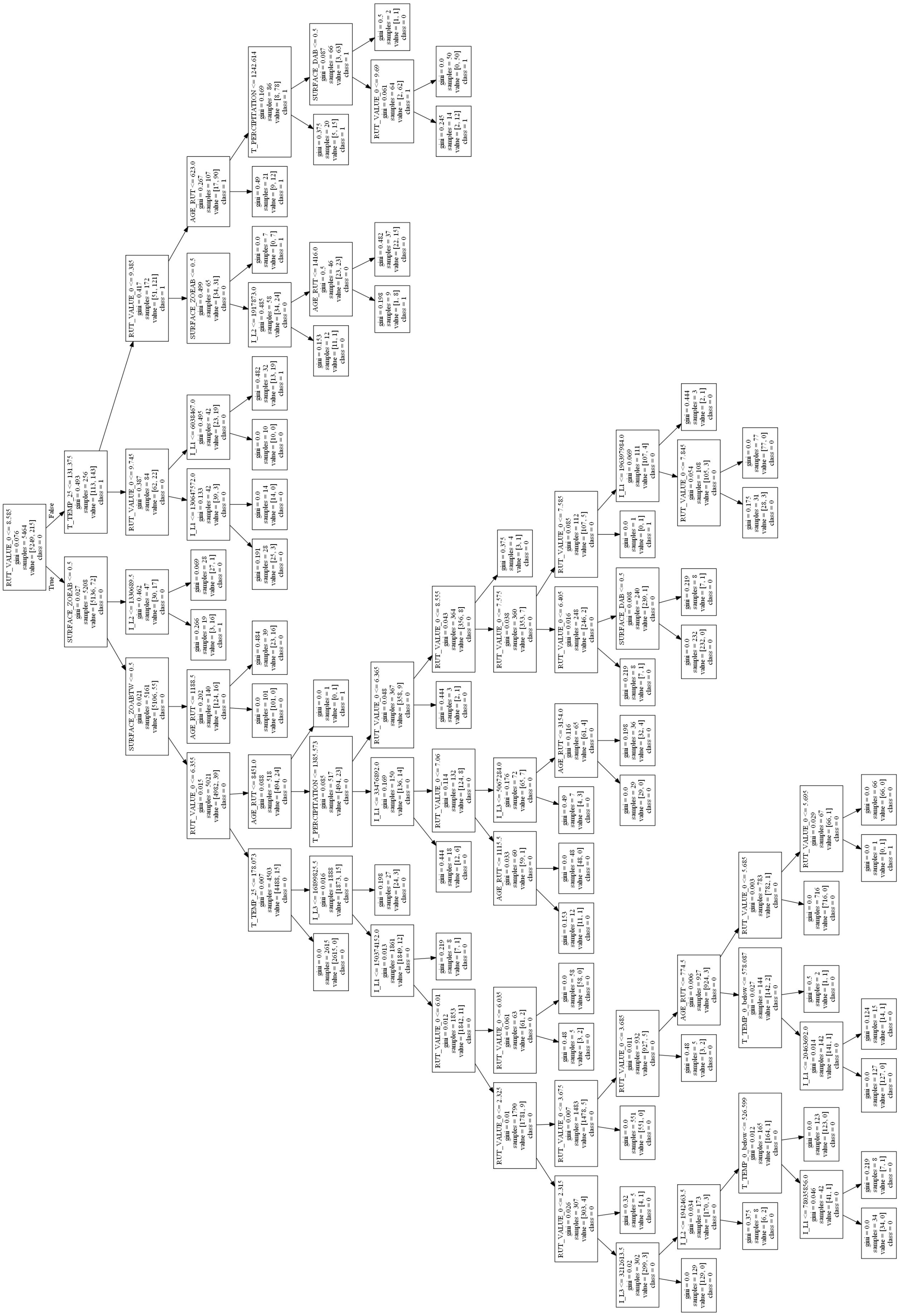


Figure 8.8: Decision tree classifier of road rutting trained by A15 data from 2015 to 2017 (Minimum sample split=40)

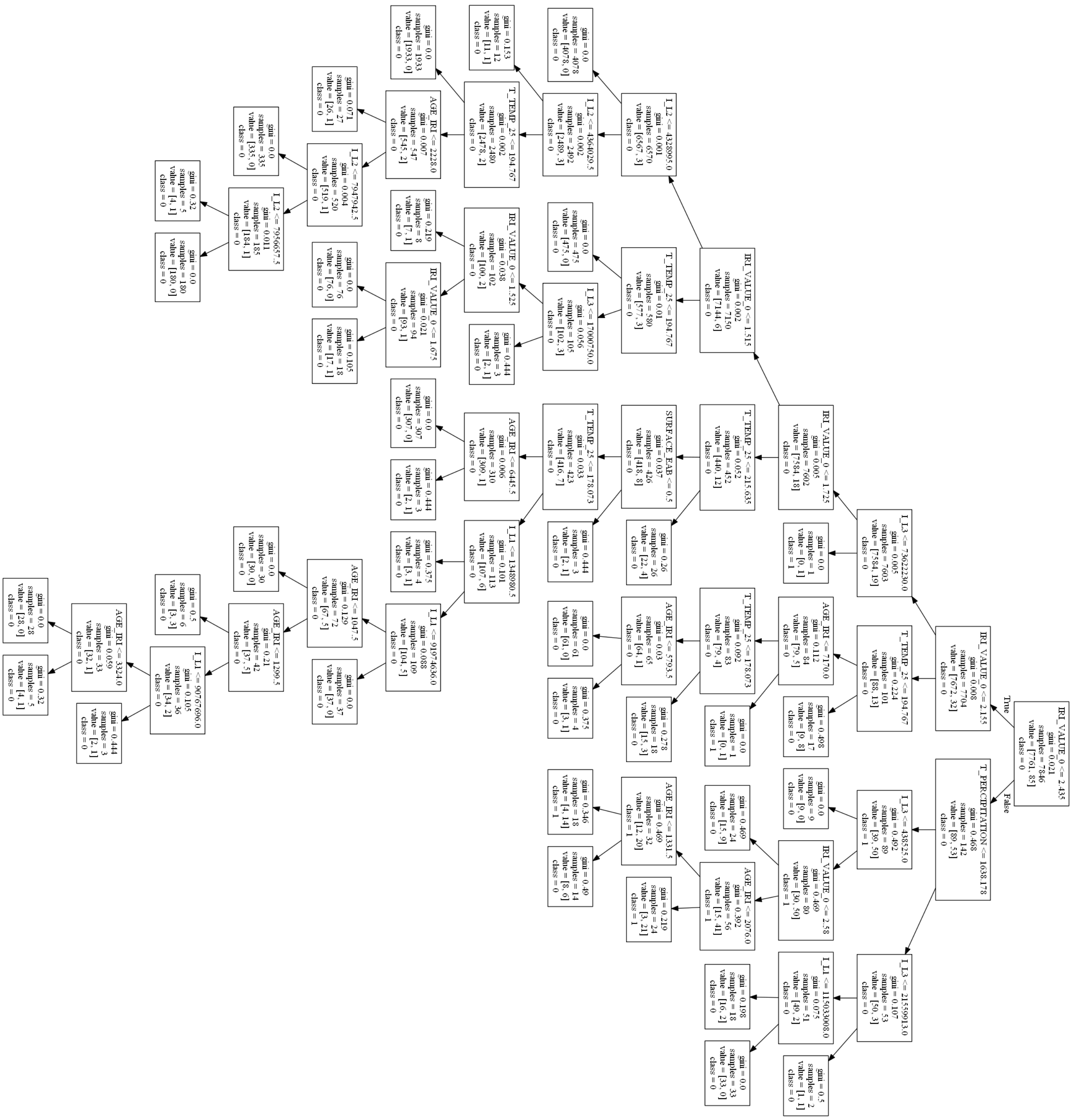


Figure 8.9: Decision tree classifier of road roughness trained by A15 data from 2015 to 2018 (Minimum sample split=30)

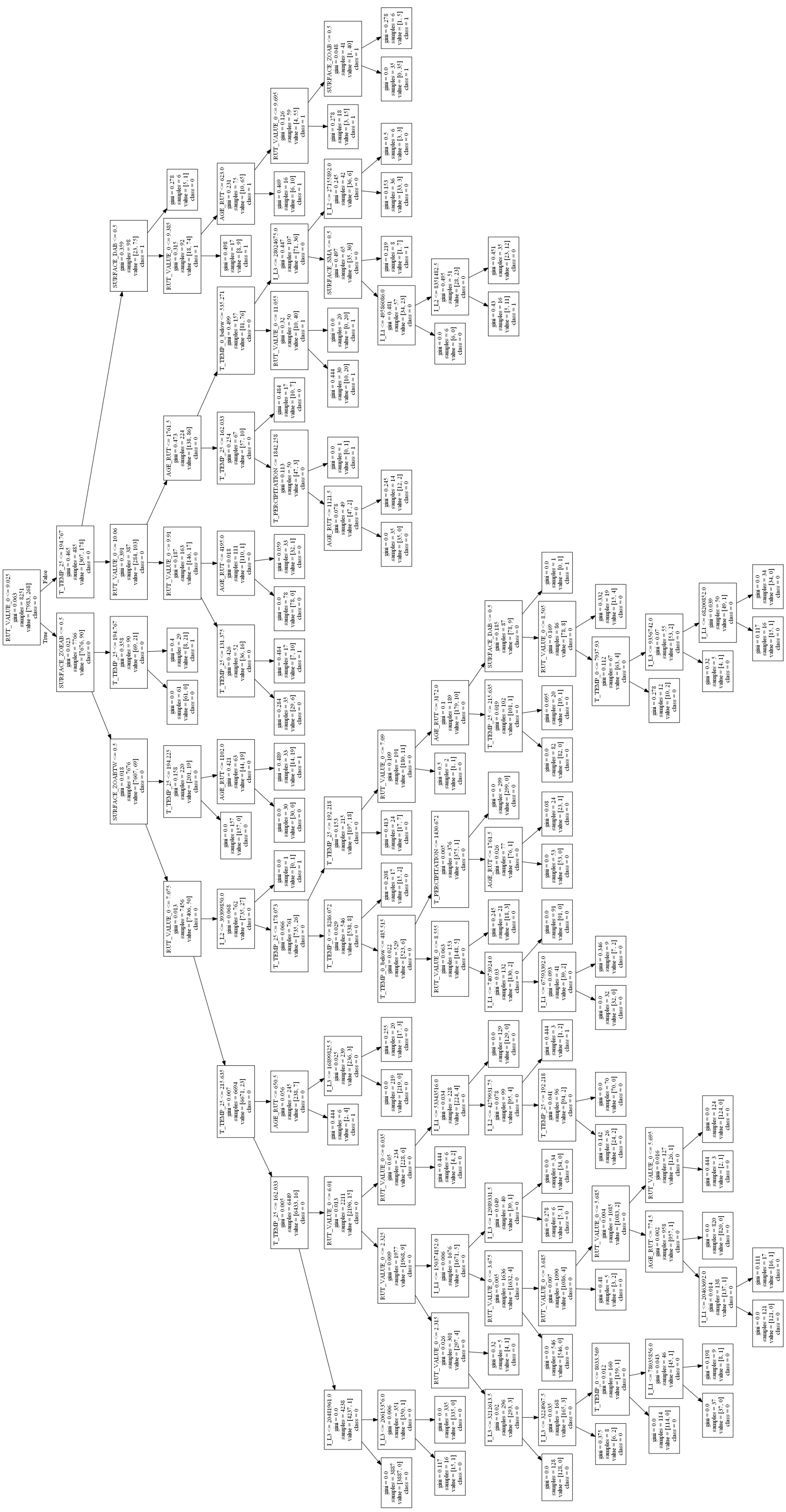


Figure 8.10: Decision tree classifier of road rutting trained by A15 data from 2015 to 2018 (Minimum sample split=40)

IV

Deliver

In Develop phase, the different models quantify the effects of traffic flows on road performance from various perspectives. All the three kinds of the performance models contribute to the road management field and help well understand the influence of the traffic characteristics on the road performance. However, it may not be enough to improve the field by quantifying these effects. It is also necessary to think about how these findings can be applied in the actual process, and how the research can improve the existing road management system in some way.

In order to enable the road management system to evaluate the road performance in the real time, the idea of establishing a simulation software is proposed in this phase. It is to predict the pavement performance based on the real-time traffic flows as well as its other characteristics of the roadway from the historical measurement data set. The design of the simulation software is elaborated in Chapter 9. The thesis ends with the conclusions and recommendations in Chapter 10.

9 Simulation Software

In Develop phase, the effects of traffic flow on road performance are qualifies by different models from various perspectives. The regression model quantifies the increment of the performance indicators caused by every vehicle. The survival model describes the decrease ratio of the expected survival time and probability by the effects of traffic flow. The decision tree classifier indicates to what extent of the traffic intensities makes the pavement approaches to the intervention level regarding its poor performance. Although the three kinds of performance models contribute to the road management field and help to well understand the influence of the traffic characteristics on the road performance, it may not be enough to improve the field by quantifying these effects only. It is important to figure out how these findings can be applied in the actual process, and how the research can improve the existing road management system. This chapter proposes the idea of establishing a traffic data-driven simulation software in order to enable the road management system to evaluate the road performance in the real time. The design contains two steps. At first, select the applicable performance model for setting up the simulation software. Then, design the development process of the software. The two parts will be elaborated on in Section 9.1 and Section 9.2 respectively. The chapter will end with the conclusion and the answer to the sub question 7 about how to apply the findings of the quantitative effects of the traffic characteristics to improve the road management field, in Section 9.3.

9.1. Model selection for simulation

The master graduation project applied three kinds of performance models to quantify the effects of the influencing factors on road performance, especially the traffic-associated factors. To set up the traffic data-driven simulation software, it is significant to select the applicable model for simulation. The selection criteria contain:

- The performance model confirms that the traffic characteristics have an apparent influence on road performance.
- The performance model can indicate the effects of the traffic characteristics in a quantitative way.
- To employ the simulation results to practical use well, the outputs of the performance model should be easily understood for the engineers in the road management field.

The candidate performance models are the regression models, the survival model, and the decision tree classifier, which were defined in Chapter 3 and applied to the case study of A15 from the observation data of 2015, 2016, 2017, and 2018 in Chapter 6, Chapter 7, and Chapter 8 respectively. The applicability of each performance model to the simulation software concerning the selection criteria will be discussed in the following paragraphs.

Regression models assume that each vehicle has an influence on road performance. The effects of traffic flow on road performance are the cumulative effects of the multiple vehicles. For a linear regression model, the effect of any vehicle on road performance is always the same. The non-linear model describes the impact of a single vehicle on road performance vary from the total amount of traffic. According to the application of the regression model in Chapter 6, the model results confirm that the traffic characteristics have the influence on road performance. But compared with the influence of the original performance state, the effects of the traffic flow are not considerable. Therefore, the model is not applicable to simulate the road performance by the traffic factors.

The survival model, defined in the research, assumes that the decrease in the survival probability of a certain surface material type of roadways under the certain weather conditions is caused by the traffic flow. The survival time of the pavement is determined by the road performance and the damage assessment criteria. According to the model application to the A15 ZOAB section in Chapter 7, the model confirms the obvious effects of the traffic flow on the expected survival lifetime of the road, and the model is able to quantify the influence. However, there are two problems when applying the model to simulate road performance. The first is that more than 90% of the roads did not fail during the observation period. In other words, there was a large number of the right-censored data. Hence, the definition of the survival time distribution is not critical. The other problem is that the research used the data analysis method to assess the condition of the road sections . The analysis results are close to the damage assessment standard, but are not identical. In terms of the road managers or the engineers, the results are a bit off the practice.

The decision tree classifier model assumes that the effects of traffic flow on road performance is transitional. In other words, if the traffic flow is within the interval, the road performance will stay rather the same. The pavement behaves rather differently only when the traffic flow are up to various levels. The model application in Chapter 8 confirms the effects of traffic flow on road performance, and it gives the specific values for the levels. The DTC figures out the common characteristics of the bad-behavior roads that meet the maintenance requirements according to the historical data. If the investigated roadway has these characteristics also, the model predicts that the road will meet maintenance requirements. The output of the model is whether or not a target road segment meets maintenance requirements, which is easily understandable and practical for the road managers or the engineers.

In summary, according to the model comparison and the selection criteria stated above, the most appropriate performance model among the three candidates for setting up the simulation software is DTC model, in this study. It should be cleared stated that the model selection results are based on the case study of A15 from 2015 to 2018. There is a possibility that the other models can be well suited for the simulation in the other case study. The chapter will continue to design the simulation software by applying the model.

9.2. Design of the simulation software

The propose of the simulation software is to provide the engineers in the road operation and maintenance field with the useful and reliable simulation results of the Dutch highway performance in the real time. To realize the aim, user journey, the well-applied design criteria in the industry design field are used (Howard, 2014). The core of the design method can be elaborated in three aspects:

- The main tasks can be achieved by the simulation software as the potential users want.
- The potential users are capable of understanding the simulation software, particularly how to use and what the output means.
- The potential users will use the function of the simulation software.

According to the design criteria of user journey, the functionality of the simulation software is defined in the study. The core functions contain:

- The software is capable to give the damage assessment (roughness, rutting) of every 100-meter section on a highway in the Netherlands from the last measurement time until a determined time point. The time point can be any moment between the last survey date until the day when using the simulation software.
- The software is capable to alarm the users when some pavement is predicted to meet the maintenance requirement. The location and the characteristics of the bad-performance roadway can be reported by the program.

To realize the core functions, the three specific functions are required as follows.

- The simulation software can read the after-processing data originating the historical performance data in IVON and real-time traffic data from DiTTlab.
- The program is able to plot the visualisation of the core algorithm of the simulation. The algorithm the software uses in the study is the DTC model.
- The software can show the simulation results which make the statement as the investigated road section has (not) met the maintenance requirements concerning the performance indicator yet, as well as a list of the road section which will be in the intervention level by the software prediction.

Accordingly, the establishment of the simulation software is designed into four phases. The first phase is building the interface of the software. The second development phase of the simulation software is proceeding the data sets for the simulation. The third development phase of the simulation software is coding the simulation model. The last one is returning the simulation results to the interface. The details of each phase will be elaborated in the following paragraphs in order.

Phase I: building the interface of the software
The interface is to set the input parameters and to plot of output clearly. The inputs of the program contain the selection of a specific roadway to predict the road performance, the determination of the prediction time point, the selection of the road performance indicator (in the study, which is either the roughness or the rut depth), the selection of the performance model (DTC model in the case). The outputs of the program have the simulation results and the visualisation of the simulation model. The simulation results are whether there are the road sections meeting the maintenance requirements on the target pavement and the list of these road sections with its predictor characteristics and the locations. The visualisation is in the shape of the decision tree. The design of the interface of the software is drawn in Fig. 9.1.

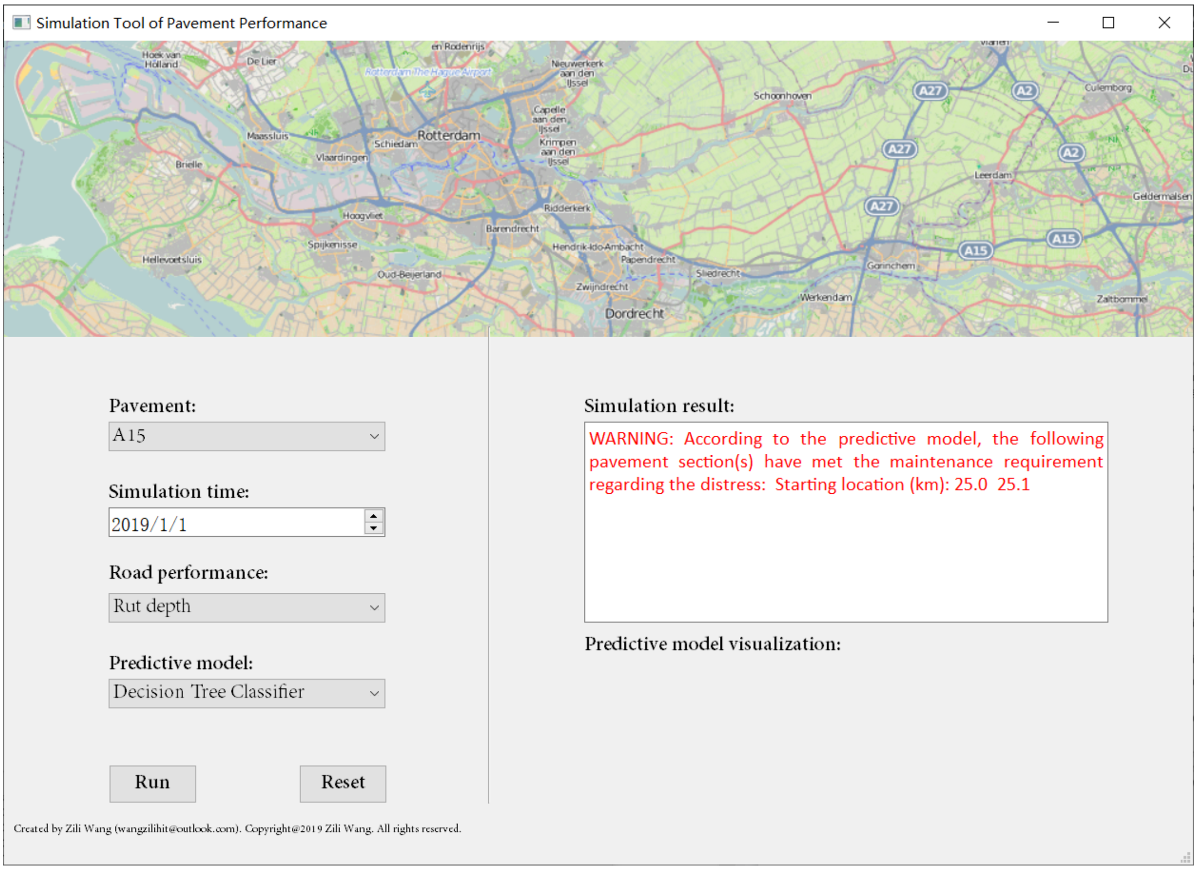


Figure 9.1: Design of the interface of the simulation software

Phase II: proceeding the data sets for the simulation
As selected in Section 9.1, DTC requires two data sets to run, the training data set and the test data set.

The training data contains the predictor variables and the response variables, but the test data only include the predictor variables, and the response variables are the prediction outcomes. The study defines the predictor variables as the service time of the pavement (AGE_IRI, AGE_RUT), the surface material of the road section (SURFACE_COMBID, SURFACE_DAB, SURFACE_EAB, SURFACE_OAB, SURFACE_SMA, SURFACE_ZOAB, SURFACE_ZOAB+, SURFACE_ZOABTW, SURFACE_ZOEAB), the measurement of the performance indicator in the last survey (IRI_VALUE_0, RUT_VALUE_0), and the cumulative traffic flow (I_L1, I_L2, and I_L3), and the response variables as the classification of IRI (R_IRI) and the rut depth (R_RUT). The raw data for proceeding the training data sets includes the measurement data of 2015, 2016, 2017 and 2018. The data for making the test data set originates the measurement data of 2018 and the minute-based traffic intensities from the last survey date until the determined prediction time provided by DiTTlab. It is necessary to compute the cumulative traffic flow of the three vehicle classes (I_L1, I_L2, and I_L3) in the test data set. The formula of the variable of I_L1 is the sum of the value of I_L1 in the training data set of the year 2018 and the product of the percentage of the passenger cars by the 2018 measurement and the aggregation of the minute-based total traffic intensities during the period the last survey date until the determined prediction time. The computation equations of I_L2, and I_L3 are similar. The difference is applying I_L2 and I_L3 in the training data set of the year 2018 and the percentages of the light trucks and the heavy trucks by the 2018 measurement instead.

Phase III: coding the simulation model
The simulation model selected in the study is DTC. The program contains five modules. The first one is calling 10 Python modules, including pandas, numpy, sklearn, matplotlib, seaborn, graphaviz, pydotplus,io, scipy, and IPython, where sklearn is the core of building up the decision tree. The library of pandas and numpy are the basic ones in Python, and the others are to plot the outcomes. The second is loading the two data tables that have been prepared in the phase II. Then the third module is naming the variables, including the predictor variables and the response variables. The names of the variables should be identical to the labels of the variables in the input excel table. The fourth is the core part of the simulation model, which is to establish DTC with the certain setting of the minimum sample split. The core algorithm originates from DTC code in sklearn Python module. In the software, the setting of the minimum sample split is free to alter by the users. At last, it is to output the road sections which would meet the maintenance requirement by the model prediction as well as its characteristics. The results are the condition sets for classifying the road sections whether would meet the maintenance requirements or not. There is one necessary step to program the condition sets to the condition statements. In the simulation model, the maintenance requirements are set as two conditions, that are IRI is more than 2.6 m/km, and the rut depth is above 10 mm. It is according to damage assessment defined by RWS (Rijkswaterstaat, 2017a). The "Run" button designed in the interface is the execute order of running the simulation model.

Phase IV: returning the simulation results to the interface
The last phase is to plot the simulation results on the interface of the simulation software. The simulation results are the image of the DTC and the selection results of the road sections whether would meet the maintenance requirements or not. The image can be read and plot directly in the window of the "Predictive model visualization" in the interface. The selection results of the sections are displayed by the IDs. Then the tool calls the rows of the selected IDs in the test data set and aggregates them to a new table. The locations in the new table are listed in the window of the "Simulation result" in the interface. The report of the characteristics of the selected road sections are shown in the Python IDLE window, which includes the roadways, the directions, the locations, the total in-service time, the surface materials, the cumulative traffic volumes of the four categories (all the vehicles, the passenger cars, the light trucks, and the heavy trucks), and the last measurements of the performance indicators.

9.3. Conclusion

The chapter provides the design of the simulation software. It applies the findings of the research to practical use. DTC model is selected as the appropriate algorithm for the simulation. The functions of the software are defined according to the industry design thinking of user journey. To realize the main and specific functions defined, the establishment of the simulation software is designed into four phases. The chapter is able to answer the sub question 7 as follows.

Sub question 7

How to apply the findings of the quantitative effects of the traffic characteristics to improve the road management field?

Based on the knowledge of the existing road management system, IVON, as introduced in Section 1.6, the current system for road health assessment is based on the annual field measurements. In order to enable the road management system to evaluate the road performance in the real time, the thesis proposes to establish a simulation software that can predict the pavement performance based on the real-time traffic flow data and the historical measurement data of the road. The DTC model is chosen as the core algorithm of the software. The implementation of setting up the program is designed in four phases: (1) building the interface of the software; (2) proceeding the data sets for the simulation; (3) coding the simulation model; (4) returning the simulation results to the interface. The simulation software can help the present road management system by predicting the road health in the real time, alarming the road managers when the road segment is in poor serviceability, and providing the assistance estimation of the target pavement for the next measurement.

10 Conclusions and recommendations

To be able to give an answer to the main research question, first the results are interpreted. This chapter will critically reflect on the research project and put the results it into context. In Section 10.1, the overall project is discussed. Section 10.2 uses the results from this research to answer the main research question. Section 10.3 describes the application of a new method in the road management process and in Section 10.4, the findings from this research are shortly compared to existing literature. The thesis ends with Section 10.5, which elaborates on the limitation of the study and offers future researches the recommendations.

10.1. Overall project

The thesis focuses on the interaction between traffic engineering and pavement engineering during road management phase of a series of in-service pavement. The proposed tool focuses on the real time road health assessment for the road managers which are remotely located away from the field. To realise the tool, the idea of taking the traffic characteristics as the predictor variables to predict the road performance is proposed, because the data collection of the traffic flow in the present system is in the real time, 24/7, automatically captured by the detector loops. The feasibility of the establishment of such a tool depends on whether the traffic characteristics have the effects on pavement performance and in what way the influence can be quantified. Thus, the explorative research presents a case study of A15 in the Netherlands, to figure out and quantify the effects of the traffic flow on the pavement performance. The time domain covered in this research is from 2015 until 2018. Three performance models are performed. According to the model results of the case study, it is confirmed that the traffic characteristics have an influence on road performance, particularly, the traffic flow contributes to the pavement which performs badly. Various performance models indicate the quantitative effects of the traffic characteristics on the pavement performance in different ways. The regression model quantifies the increment of the performance indicators caused by every vehicle. The survival model describes the decrease ratio of the expected survival time and probability by the effects of traffic flow. The decision tree classifier illustrates to what extent of the traffic intensities make the pavement approach to the intervention level regarding its poor performance. Accordingly, the thesis confirms the feasibility of setting up a simulation tool for predicting road performance by capturing the influence of the traffic flow. The specific establishment process has been designed.

10.2. Main research question

With the interpretation of the overall results of the project, it is able to provide the answers to the research questions as the thesis defined. In the previous part of the report, the sub research questions are already answered. In this section, the main research question will be discussed.

The main research question is defined from the purpose of improving the road management system. The project is a process of discovering the "user's need", proposing the potential solutions, examining the feasibility, and conducting the application. The design thinking, a four-stage project approach is perfect for conducting the process. This report shows the complete process by using the design thinking as the main project approach with the combination with a set of methods to solve the research problem.

The thesis explores the main question by the qualitative and quantitative research approaches. At the beginning of the research project, the correlation between the traffic factors and pavement performance was unclear, so a qualitative approach was used to explore the problem to come up with an approach to handle the problem afterwards. Quantitative research is a proper follow-up after qualitative research. In the qualitative research, the literature study is included, which focuses on the failure mechanism of road surfaces generated by all kinds of factors, especially the traffic-associated causes. Correlation analysis, a statistical method follows to study the strength of the relationship between two, numerically measured, continuous variables. The presence of a correlation is not sufficient to infer the presence of a causal relationship, and thus the report considers both the literature study and the correlation analysis to imply the role of transportation in the causation of road deterioration. To quantify the effects of the traffic characteristics on pavement performance, as the main research question asks, the three kinds of performance models are applied to the case study of A15. The time domain covered in this research is from 2015 until 2018. The model results direct to the answer to the main research question. The answer is elaborated in the following.

Main question

What are the effects of traffic flow characteristics on pavement response and performance in a quantitative way?

This thesis studied the effects of traffic flow characteristics on pavement performance by the regression models, survival model, and decision tree classifier. The regression models assumes that traffic flow affects the development of pavement performance. That is to say, as time goes by a road serves more and more traffic, and the performance indicators (distress, skid, roughness, and others) are assumed to change towards the bad condition accordingly. The survival model considers that traffic flow affected the expected lifetime of the roadways and the survival probability. The classification model is to find the critical value of traffic intensities that made the road ineffective. The following three paragraphs respectively will explain the results of applying the regression model, the survival model, and the classification model on A15, and the quantitative influence of three types of traffic intensities on road roughness and rutting of the three models. The results are derived from the observation data of the entire A15 road from 2015 to 2018. In addition, pavement performance is the response to the combined influence of the multiple influencing factors. The traffic-associated factors are the study focus, but the other non-traffic associated influencing factors are also considered in an appropriate manner in the application of each model in this study.

The linear regression model on A15 verifies the traffic flow' effects on the road roughness under some climate scenarios. But it can not conclude that a certain vehicle type always makes road surface more rough or smooth. According to the model results, the tendency how the vehicle class affects the road roughness is also influenced by the climate condition. The linear regression model shows the development directions influenced by the trucks (both heavy and light) of the surface roughness were always the same, which were contrary to the pavement development direction affected by the passenger cars. More data is needed to validate this finding. The linear regression model on A15 confirms that the vehicles regardless of passenger cars, light trucks, and heavy trucks have the significant influence on the rutting development in a negative way. But the quantities of the effects of the traffic flow on the rut depth vary from the different climate conditions. According to the model result, the decisive predictor variables of the models are the initial value of the road roughness and rutting. Taking these variables in the multivariate regression model makes a good fitting. In contrast, the influence of the traffic flow is not considerable.

According to the survival analysis results of A15 ZOAB data from 2015 to 2018, three types of vehicles all had a negative influence on road health determined by IRI, but the degrees of their influence were different. The intensities of the passenger cars shortened the expected survival time of the ZOAB road roughness by 4.33% and 4.05% in two regions of A15, and regarding the rut depth the expected survival time of the ZOAB pavement decreased by 3.2%, and 8.44% respectively.

(To be continued in the next page)

Main question

What are the effects of traffic flow characteristics on pavement response and performance quantitatively?

(The answer continues here)
The effects of the number of the light trucks reduced the mean of the survival time of ZOAB road roughness by 52.08% and 29.69% in the WNZZ and ONZ district, and of rutting by 30.7% and 78.32% respectively. The heavy trucks decreased the life expectancy of the ZOAB pavements by 39.85% and 37.46% about the road roughness in the two areas, and as for the rutting in the WNZZ and ONZ district, the numbers are 24.24% and 76.3% respectively. However, the analysis results significantly overestimate the impact of traffic factors on road health. Since only the data on A15 from 2015 to 2018 was used, fewer observation data indicates that the road section failed. Besides, the survival analysis in the study assumes that with the controlling variables of the climate conditions, the failure event occurring on the pavement is totally dependent on the traffic flow of any one vehicle class. The hazard ratios are pretty data sensitive, and the research does not quantify the influence of the traffic intensities on road performance correctly, but it confirms the feasibility of applying the survival model to study the transport influence instead.
The model application of the decision tree classifier confirms that the initial values of IRI and rutting are the important criteria to classify the roadways regarding road roughness and rutting respectively. The traffic flow, the climate conditions, and the service time are selected as the classification criteria regarding the pavement roughness, and the traffic flow, the climate conditions, the surface materials, and the service time are the classification criteria concerning the pavement rutting. The quantitative effects of the traffic factors found by the model are that when the measurement of the road roughness of the section in the previous year was less than 2.135 m/km, the very high cumulative intensities of the heavy trucks (above 73622230 vehicles) made it to be the bad condition class. When the unevenness value measured in the previous year was close to the intervention level (2.6 m/km), even with a slight value of the cumulative number of the heavy vehicles, the road section would be predicted to fail by the model based on the training data of A15 from 2015 to 2018. The model results did not indicate the importance of the heavy trucks on road roughness, because the variables selected as the decision conditions are sensitive to the input data. When inputting the data of A15 from 2015 to 2017 the decision trees did not select the traffic intensities of the heavy trucks as the decision condition, but the volumes of the light trucks and passenger cars played some roles instead. Thus the DTC model confirms the importance of the traffic flow on road roughness, but not specify the vehicle class. As for the effects of the traffic factors on the rutting, among 12 condition sets into which the failed roadways are categorised, 5 of them have the traffic flow as the decision criterion, and the critical values of traffic flow have the range between 4 million and 50 million, which are the considerable numbers. The critical level defined by DTC model of the traffic volumes of each vehicle class are very sample-data sensitive, and can change when inputting the different period of the measurement data and the minimum sample split.

10.3. DTC method applied in road management field

The thesis finds the DTC model is the appropriate method to improve the current road management field. The study innovatively proposes to use the DTC model to predict the road performance based on the traffic and other influencing factors, thereby providing a technical perspective for the road maintenance decision strategy. As far as the author knows, it is the first time to apply the algorithm to the field. Additionally, the thesis designs the specific implement phases of establishing the simulation software on the basis of the model. It paves the way for the practical application of the model for improving the road management field.

10.4. Findings of the research compared to the existing literature

In literature, the correlation between the traffic influence factors and pavement performance has been discussed already. But these researches focus more on the ESAL and the truck percentage. Due to the lack of measurement data the traffic flow of various vehicle types, most literature use the damage equivalent load method to generally estimate the impact of mix traffic on road performance (Paterson, 1987; Loizos and Karlaftis, 2005; Wang et al., 2005; Adlinge and Gupta, 2013). The research based on the measurement data of the number of the passenger cars, the light trucks, and the heavy trucks on A15 finds that the vehicle types have the different effects on road performance. The prediction of the pavement performance can be more realistic by taking into account the traffic flow of each vehicle type as the individual predictor variable.

The progression of road roughness and rutting by the local observations may be different from the empirical knowledge described in the literature. The regression models of road roughness proposed in the book by Paterson (1987) assumes the continuous increase in IRI, and the article by Freeme (1983) indicates the three development phases where the rut depth grows in different speeds. However the road performance progression (IRI and the rut depth) is in fluctuation, based on the observation data of A15 from 2015 to 2018. This explains why the non-linear models as defined in the previous studies that use the service time as the predictor variables does not achieve a good fitting result in this case study. Besides, the articles by Hodges et al. (1975) and Parsley and Robinson (1982) proposed the linear relationship between the changes of IRI and the rut depth and traffic volumes, but the kind of correlation has not been observed in the test data.

10.5. Limitations and suggestions for further research

In the study process, the following limitations have been found. Some of the limitations even have a significant consequence on the research results. This section elaborates on the limitations, the potential consequences caused by these, and the possible solutions to solve the problems. It is desirable that future studies can consider the limitations as well as the suggestions of the study.

Limited by the test data of A15 from 2015 to 2018, this study cannot conclude that the data correlation between the influencing factors and the pavement performance found in this research must be observed on other roadways in other countries. To truly find the general law of the relationship between the influencing factors and the road performance, a large number of the empirical data, including various types of roads (highways, distributor roads, access roads, and so on), in the different regions during a long term observation is required.

This study is based on full trust in the accuracy of the collected data. However, according to the general experience, it is almost impossible to obtain 100% accurate measurement data, especially the estimation of the pavement performance involved the human factor (Stichting Studie Centrum Wegenbouw, 1987). Rare data is incorrectly recorded in the database. In particular, there might be occasions that the records were not updated during the construction time but afterwards, resulting a lag and inaccuracy in the database. The situation can result in the records of a fresh performance with all kinds of the factors in the rather long service time, like a heavy cumulative traffic intensities. If it is the case, the actual correlations of transportation intensities and the road performance are most likely closer than the results of the data analysis.

Although the method of the control variables is a common method for studying the effects of univariate on a dependent variable, the applicability of the method is limited in the case study, due to lack of the plenty of the sample. However, the approach can be applied to studies of road performance in the laboratory, or when considering roads throughout the Netherlands.

The consideration of the climate factors in the study is limited by the way of collecting and processing weather data. The database supporting the research located one weather station in one region, and did the irregular surveys of the temperature and the precipitation in the area. To infer the weather situation during a time period at a location, the study assumes that the probability of high temperature, low temperature, and rains in the time interval of the non-measurement time is consistent with that in the measurement time. If the future research aims to confirm the effect of the weather condition on the pavement performance, it is necessary to access the weather data in the real time at a smaller region, or look into a much wider test area.

The censoring problem is a significant problem in the survival model, especially the right-censored data. In the test data, 90% of the sections do not approach the end of the lifetime, which makes it impossible to find the accurate survivor function or the distribution of the survival time in the study. The best and easiest solution is extending the observation time, and studying the larger amount of the population.

V

Appendices

A

A15 sections

Table A.1: Road sections of A15

ROAD	SECTION_ID	FROM_KM	FROM_DISTRICT	TO_KM	TO_DISTRICT
15	131278005	25.0	Knooppunt Slufter	25.7	Oostvoorne
15	132277006	25.7	Oostvoorne	26.3	Oostvoorne
15	133278005	26.3	Oostvoorne	27.0	Havens 6200-7000
15	134279018	27.0	Havens 6200-7000	27.8	Havens 6200-7000
15	135279002	27.8	Havens 6200-7000	29.5	Havens 5700-6200
15	138278005	29.5	Havens 5700-6200	30.7	Havens 5700-6200
15	139278020	30.7	Havens 5700-6200	32.2	Havens 5500-5700
15	143275040	32.2	Havens 5500-5700	32.7	Havens 5500-5700
15	146272002	34.8	Brielle	36.2	Brielle
15	148270035	36.2	Brielle	38.0	Brielle
15	152270016	38.0	Brielle	39.0	Rozenburg-Centrum
15	154267004	40.1	Rozenburg-Centrum	42.4	Havens 4100-5200
15	154269043	39.0	Rozenburg-Centrum	40.1	Rozenburg-Centrum
15	156263010	42.4	Havens 4100-5200	43.4	Havens 4100-5200
15	158263004	43.4	Havens 4100-5200	44.2	Botlekbrug
15	160263016	44.2	Botlekbrug	45.7	Spijkenisse
15	165263013	45.7	Spijkenisse	47.6	A15
15	166264048	47.6	A15	48.8	A15
15	169264011	48.8	A15	49.8	A15
15	170264046	49.8	A15	50.5	Pernis
15	171264013	50.5	Pernis	51.7	Beneluxweg
15	174264017	51.7	Beneluxweg	54.0	Rotterdam-Heijplaat
15	184261028	57.2	A15	58.2	A15
15	186262001	58.2	A15	60.5	A15
15	191261054	60.5	A15	61.0	Kp Vaanplein
15	192261052	61.0	Kp Vaanplein	61.8	A15
15	193262014	61.8	A15	62.6	Kp Ridderkerk-Noord
15	195262055	62.6	Kp Ridderkerk-Noord	63.1	Kp Ridderkerk-Noord
15	196263022	63.1	Kp Ridderkerk-Noord	64.0	Kp Ridderkerk-Noord
15	197263010	64.0	Kp Ridderkerk-Noord	64.7	Kp Ridderkerk-Noord
15	198262010	64.7	Kp Ridderkerk-Noord	65.9	Kp Ridderkerk-Zuid
15	200260030	65.9	Kp Ridderkerk-Zuid	70.6	Kp Ridderkerk-Zuid
15	202260014	70.6	Kp Ridderkerk-Zuid	71.5	Hendrik Ido Ambacht

(To be continued)

ROAD	SECTION_ID	FROM_KM	FROM_DISTRICT	TO_KM	TO_DISTRICT
15	204260038	71.5	Hendrik Ido Ambacht	72.6	Hendrik Ido Ambacht
15	208260004	72.6	Hendrik Ido Ambacht	75.4	Alblasserdam
15	211259011	75.4	Alblasserdam	76.0	Alblasserdam
15	212258039	76.0	Alblasserdam	78.0	Papendrecht
15	216256032	78.0	Papendrecht	79.0	Papendrecht
15	220254014	79.0	Papendrecht	80.4	Sliedrecht-West
15	220254012	80.4	Sliedrecht-West	80.8	Sliedrecht-West
15	222254032	80.8	Sliedrecht-West	84.6	Sliedrecht-Oost
15	228252106	84.6	Sliedrecht-Oost	85.4	Sliedrecht-Oost
15	230252024	85.4	Sliedrecht-Oost	87.7	Hardinxveld-Giessendam
15	234252045	87.7	Hardinxveld-Giessendam	88.3	Hardinxveld-Giessendam
15	247256014	94.5	Kp Gorinchem	96.9	Kp Gorinchem
15	251257009	96.9	Kp Gorinchem	99.5	Arkel
15	257256043	99.5	Arkel	99.8	Arkel
15	269256035	105.3	Leerdam	106.0	Leerdam
15	286258013	114.3	Kp Deil	116.6	Kp Deil
15	290259005	116.6	Kp Deil	118.7	Meteren
15	294260006	118.7	Meteren	119.0	Meteren
15	299260007	121.1	Geldermalsen	121.8	Geldermalsen
15	302261006	121.8	Geldermalsen	123.8	Wadenoijen
15	304263015	123.8	Wadenoijen	124.2	Wadenoijen
15	304263016	124.2	Wadenoijen	126.9	Tiel-West
15	309267058	126.9	Tiel-West	127.1	Tiel-West
15	309267067	127.1	Tiel-West	130.4	Tiel
15	315269029	130.4	Tiel	131.2	Tiel
15	316269032	131.2	Tiel	134.4	Echteld
15	322271010	134.4	Echteld	134.8	Echteld
15	323272021	134.8	Echteld	141.0	Ochten
15	335273068	141.0	Ochten	141.7	Ochten
15	344274007	145.5	Dodewaard	146.1	Dodewaard
15	356272040	151.5	Andelst	152.1	Andelst
15	357272008	152.1	Andelst	153.8	Kp Valburg
15	360272016	153.8	Kp Valburg	154.5	Kp Valburg
15	361271014	154.5	Kp Valburg	155.1	Kp Valburg
15	362270007	155.1	Kp Valburg	156.1	Kp Valburg
15	364269006	156.1	Kp Valburg	159.5	Elst
15	370267008	159.5	Elst	160.1	Elst
15	372268012	160.1	Elst	161.5	Kp Ressen
15	375268024	161.5	Kp Ressen	161.9	Kp Ressen
15	375268025	161.9	Kp Ressen	162.1	Kp Ressen
15	376268024	162.1	Kp Ressen	162.6	Kp Ressen
15	381269017	162.6	Kp Ressen	165.2	Van Elkweg

B

Pavement performance data process Matlab code

The following gives the example code of integrating the annual measurements of IRI in Matlab. With the similar idea, the integration of the data of the other performance indicators, the traffic data, and the climate data can be programmed. Limited by the length of the report, the other codes are not displaced here.

```
clear all
% Road: A15
% Surface layer: Zoab
% Direction: L
% District: WZNN

%% Data input: Location, IRI
%Input locations from excel tables
From_km_2015 = xlsread('E:\RWSA15\Zoab_WNZZ_IRI.xlsx','2015_1HRL','D2:D350');
From_km_2016 = xlsread('E:\RWSA15\Zoab_WNZZ_IRI.xlsx','2016_1HRL','D2:D350');
From_km_2017 = xlsread('E:\RWSA15\Zoab_WNZZ_IRI.xlsx','2017_1HRL','D2:D350');
From_km_2018 = xlsread('E:\RWSA15\Zoab_WNZZ_IRI.xlsx','2018_1HRL','D2:D350');
%Input IRI from excel tables
IRI_2015 = xlsread('E:\RWSA15\Zoab_WNZZ_IRI.xlsx','2015_1HRL','O2:O350');
IRI_2016 = xlsread('E:\RWSA15\Zoab_WNZZ_IRI.xlsx','2016_1HRL','O2:O350');
IRI_2017 = xlsread('E:\RWSA15\Zoab_WNZZ_IRI.xlsx','2017_1HRL','O2:O350');
IRI_2018 = xlsread('E:\RWSA15\Zoab_WNZZ_IRI.xlsx','2018_1HRL','O2:O350');

%% 4-year measuremnts integration
IRI_Diff = zeros(350,3); %IRI variance during 1 year
N_IRI_2015 = length (IRI_2015); % Number of IRI data in 2015
N_IRI_2016 = length (IRI_2016); % Number of IRI data in 2016
N_IRI_2017 = length (IRI_2017); % Number of IRI data in 2017
N_IRI_2018 = length (IRI_2018); % Number of IRI data in 2018

%Plot IRI variance in 2016
for i=1:N_IRI_2016
    n = 1; %row
    % find IRI of the same section in the different measurement year
    while From_km_2015(n) ~= From_km_2016(i) && n < N_IRI_2015
        n = n +1;
    end
    if From_km_2015(n) == From_km_2016(i)
```

```
        IRI_Diff(i,1) = IRI_2016(i)- IRI_2015(n); % IRI variance in 2016
        x_2015(i) = IRI_2015(n);
        y_2015(i) = IRI_Diff(i,1);
    end
end
subplot(1,3,1),scatter(x_2015,y_2015)
title ('IRI variance in 2016')
xlabel('IRI value in 2015')
ylabel('IRI variance during 1 year')
axis([0 4.5 -1.5 2.5])

%Plot IRI variance in 2017
for i=1:N_IRI_2017
    n = 1; %row
    % find IRI of the same section in the different measurement year
    while From_km_2016(n) ~= From_km_2017(i) && n < N_IRI_2016
            n = n +1;
    end
    if From_km_2016(n) == From_km_2017(i)
        IRI_Diff(i,2) = IRI_2017(i)- IRI_2016(n);% IRI variance in 2017
        x_2016(i) = IRI_2016(n);
        y_2016(i) = IRI_Diff(i,2);
    end
end
subplot(1,3,2),scatter(x_2016,y_2016)
title ('IRI variance in 2017')
xlabel('IRI value in 2016')
ylabel('IRI variance during 1 year')
axis([0 4.5 -1.5 2.5])

%Plot IRI variance in 2018
for i=1:N_IRI_2018
    n = 1; %row
    % find IRI of the same section in the different measurement year
    while From_km_2017(n) ~= From_km_2018(i) && n < N_IRI_2017
        n = n+1;
    end
    if From_km_2017(n) == From_km_2018(i)
        IRI_Diff(i,3) = IRI_2018(i)- IRI_2017(n);% IRI variance in 2018
        x_2017(i) = IRI_2017(n);
        y_2017(i) = IRI_Diff(i,3);
    end
end
subplot(1,3,3),scatter(x_2017,y_2017)
title ('IRI variance in 2018')
xlabel('IRI value in 2017')
ylabel('IRI variance during 1 year')
axis([0 4.5 -1.5 2.5])
```

C

Traffic data process Matlab code

The following gives the example code of calculating the variables of I_AL, I_L1, I_L2, and I_L3 in Matlab. With the similar idea, all the variables in the traffic category are computed, but the codes are not shown due to the report length limit.

```
clear all
%% Calculate Cumulative Intensities until IRI measurement time (2017)

%% Data input
I_2011_2018 = xlsread('F:\RWSA15\I_AL.xlsx','I_L_2011-2018'); % flow table on
    direction L or R
IRI_2017_ONZ = xlsread('F:\RWSA15\I_AL.xlsx','2017_L'); % IRI table on direction L
    or R
Direction = 'L'; % Input Direction is L or R

%% Calculate Cumulative Intensities from year 2010 to year 2011
% Firstly, find the same lable of section of 2 tables (flow table and IRI
% table)
TO_KM_I = IRI_2017_ONZ(:,4);
if Direction == 'L'
    TO_KM_IRI_2011 = I_2011_2018(:,58); % The unit is km
else
    TO_KM_IRI_2011 = I_2011_2018(:,56);
end
I_AL_2011 = I_2011_2018(:,59);
I_L1_2011 = I_2011_2018(:,60);
I_L2_2011 = I_2011_2018(:,61);
I_L3_2011 = I_2011_2018(:,62);

Num_TO_KM_I = length (TO_KM_I);
Num_TO_KM_IRI_2011 = length (TO_KM_IRI_2011);

I_AL_2011_v(:,Num_TO_KM_I ) = 0;
I_L1_2011_v(:,Num_TO_KM_I ) = 0;
I_L2_2011_v(:,Num_TO_KM_I ) = 0;
I_L3_2011_v(:,Num_TO_KM_I ) = 0;

for i =1:Num_TO_KM_I
    n = 1;
        while TO_KM_I(i) > TO_KM_IRI_2011(n)
            n = n + 1;
```

```
            end
            if n == 1
            I_AL_2011_v(i) = I_AL_2011(n);% if lack traffic data at the start
                section, fill in the traffic data of the first-recorded section
            I_L1_2011_v(i) = I_L1_2011(n);% if lack traffic data at the start
                section, fill in the traffic data of the first-recorded section
            I_L2_2011_v(i) = I_L2_2011(n);% if lack traffic data at the start
                section, fill in the traffic data of the first-recorded section
            I_L3_2011_v(i) = I_L3_2011(n);% if lack traffic data at the start
                section, fill in the traffic data of the first-recorded section
            else
            I_AL_2011_v(i) = I_AL_2011(n-1);
            I_L1_2011_v(i) = I_L1_2011(n-1);
            I_L2_2011_v(i) = I_L2_2011(n-1);
            I_L3_2011_v(i) = I_L3_2011(n-1);
            end
end
% fill the sections which lack of data with the data of the nearest section
for i =1:Num_TO_KM_I
    if I_AL_2011_v(i) ==0
        I_AL_2011_v(i) = I_AL_2011_v(i-1);
    end
    if I_L1_2011_v(i) ==0
        I_L1_2011_v(i) = I_L1_2011_v(i-1);
    end
    if I_L2_2011_v(i) ==0
        I_L2_2011_v(i) = I_L2_2011_v(i-1);
    end
    if I_L3_2011_v(i) ==0
        I_L3_2011_v(i) = I_L3_2011_v(i-1);
    end
end

% Secondly, cumulative intensities = days * daily flows (veh/day)
Duration_2011 = IRI_2017_ONZ(:,10);
I_AL(:,1) = Duration_2011.* I_AL_2011_v';
I_L1(:,1) = Duration_2011.* I_L1_2011_v';
I_L2(:,1) = Duration_2011.* I_L2_2011_v';
I_L3(:,1) = Duration_2011.* I_L3_2011_v';

%% Calculate Cumulative Intensities from year 2011 to year 2012
% Firstly, find the same lable of section of 2 tables (flow table and IRI
% table)
TO_KM_I = IRI_2017_ONZ(:,4);
if Direction == 'L'
    TO_KM_IRI_2012 = I_2011_2018(:,50)/1000; % Translate the unit of meter in
        original table to km
else
    TO_KM_IRI_2012 = I_2011_2018(:,49)/1000;
end
I_AL_2012 = I_2011_2018(:,51);
I_L1_2012 = I_2011_2018(:,52);
I_L2_2012 = I_2011_2018(:,53);
I_L3_2012 = I_2011_2018(:,54);

Num_TO_KM_I = length (TO_KM_I);
```

```
Num_TO_KM_IRI_2012 = length (TO_KM_IRI_2012);

I_AL_2012_v(:,Num_TO_KM_I ) = 0;
I_L1_2012_v(:,Num_TO_KM_I ) = 0;
I_L2_2012_v(:,Num_TO_KM_I ) = 0;
I_L3_2012_v(:,Num_TO_KM_I ) = 0;

for i =1:Num_TO_KM_I
    n = 1;
            while TO_KM_I(i) > TO_KM_IRI_2012(n)
                n = n + 1;
            end
            if n == 1
            I_AL_2012_v(i) = I_AL_2012(n);% if lack traffic data at the start section, fill in the traffic data of the first-recorded section
            I_L1_2012_v(i) = I_L1_2012(n);% if lack traffic data at the start section, fill in the traffic data of the first-recorded section
            I_L2_2012_v(i) = I_L2_2012(n);% if lack traffic data at the start section, fill in the traffic data of the first-recorded section
            I_L3_2012_v(i) = I_L3_2012(n);% if lack traffic data at the start section, fill in the traffic data of the first-recorded section
            else
            I_AL_2012_v(i) = I_AL_2012(n-1);
            I_L1_2012_v(i) = I_L1_2012(n-1);
            I_L2_2012_v(i) = I_L2_2012(n-1);
            I_L3_2012_v(i) = I_L3_2012(n-1);
            end
end
% fill the sections which lack of data with the data of the nearest section
for i =1:Num_TO_KM_I
    if I_AL_2012_v(i) ==0
        I_AL_2012_v(i) = I_AL_2012_v(i-1);
    end
    if I_L1_2012_v(i) ==0
        I_L1_2012_v(i) = I_L1_2012_v(i-1);
    end
    if I_L2_2012_v(i) ==0
        I_L2_2012_v(i) = I_L2_2012_v(i-1);
    end
    if I_L3_2012_v(i) ==0
        I_L3_2012_v(i) = I_L3_2012_v(i-1);
    end
end

% Secondly, cumulative intensities = days * daily flows (veh/day)
Duration_2012 = IRI_2017_ONZ(:,12);
I_AL(:,2) = Duration_2012.* I_AL_2012_v';
I_L1(:,2) = Duration_2012.* I_L1_2012_v';
I_L2(:,2) = Duration_2012.* I_L2_2012_v';
I_L3(:,2) = Duration_2012.* I_L3_2012_v';
%% Calculate Cumulative Intensities from year 2012 to year 2013
% Firstly, find the same lable of section of 2 tables (flow table and IRI
% table)
TO_KM_I = IRI_2017_ONZ(:,4);
if Direction == 'L'
```

```
    TO_KM_IRI_2013 = I_2011_2018(:,42)/1000; % Translate the unit of meter in original table to km
else
    TO_KM_IRI_2013 = I_2011_2018(:,41)/1000;
end
I_AL_2013 = I_2011_2018(:,43);
I_L1_2013 = I_2011_2018(:,44);
I_L2_2013 = I_2011_2018(:,45);
I_L3_2013 = I_2011_2018(:,46);

Num_TO_KM_I = length (TO_KM_I);
Num_TO_KM_IRI_2013 = length (TO_KM_IRI_2013);

I_AL_2013_v(:,Num_TO_KM_I ) = 0;
I_L1_2013_v(:,Num_TO_KM_I ) = 0;
I_L2_2013_v(:,Num_TO_KM_I ) = 0;
I_L3_2013_v(:,Num_TO_KM_I ) = 0;

for i =1:Num_TO_KM_I
    n = 1;
            while TO_KM_I(i) > TO_KM_IRI_2013(n)
                n = n + 1;
            end
            if n == 1
            I_AL_2013_v(i) = I_AL_2013(n);% if lack traffic data at the start section, fill in the traffic data of the first-recorded section
            I_L1_2013_v(i) = I_L1_2013(n);% if lack traffic data at the start section, fill in the traffic data of the first-recorded section
            I_L2_2013_v(i) = I_L2_2013(n);% if lack traffic data at the start section, fill in the traffic data of the first-recorded section
            I_L3_2013_v(i) = I_L3_2013(n);% if lack traffic data at the start section, fill in the traffic data of the first-recorded section
            else
            I_AL_2013_v(i) = I_AL_2013(n-1);
            I_L1_2013_v(i) = I_L1_2013(n-1);
            I_L2_2013_v(i) = I_L2_2013(n-1);
            I_L3_2013_v(i) = I_L3_2013(n-1);
            end
end
% fill the sections which lack of data with the data of the nearest section
for i =1:Num_TO_KM_I
    if I_AL_2013_v(i) ==0
        I_AL_2013_v(i) = I_AL_2013_v(i-1);
    end
    if I_L1_2013_v(i) ==0
        I_L1_2013_v(i) = I_L1_2013_v(i-1);
    end
    if I_L2_2013_v(i) ==0
        I_L2_2013_v(i) = I_L2_2013_v(i-1);
    end
    if I_L3_2013_v(i) ==0
        I_L3_2013_v(i) = I_L3_2013_v(i-1);
    end
end

% Secondly, cumulative intensities = days * daily flows (veh/day)
```

```
Duration_2013 = IRI_2017_ONZ(:,14);
I_AL(:,3) = Duration_2013.* I_AL_2013_v';
I_L1(:,3) = Duration_2013.* I_L1_2013_v';
I_L2(:,3) = Duration_2013.* I_L2_2013_v';
I_L3(:,3) = Duration_2013.* I_L3_2013_v';

%% Calculate Cumulative Intensities from year 2013 to year 2014
% Firstly, find the same lable of section of 2 tables (flow table and IRI
% table)
TO_KM_I = IRI_2017_ONZ(:,4);
if Direction == 'L'
    TO_KM_IRI_2014 = I_2011_2018(:,34)/1000; % Translate the unit of meter in
        original table to km
else
    TO_KM_IRI_2014 = I_2011_2018(:,33)/1000;
end
I_AL_2014 = I_2011_2018(:,35);
I_L1_2014 = I_2011_2018(:,36);
I_L2_2014 = I_2011_2018(:,37);
I_L3_2014 = I_2011_2018(:,38);

Num_TO_KM_I = length (TO_KM_I);
Num_TO_KM_IRI_2014 = length (TO_KM_IRI_2014);

I_AL_2014_v(:,Num_TO_KM_I ) = 0;
I_L1_2014_v(:,Num_TO_KM_I ) = 0;
I_L2_2014_v(:,Num_TO_KM_I ) = 0;
I_L3_2014_v(:,Num_TO_KM_I ) = 0;

for i =1:Num_TO_KM_I
    n = 1;
            while TO_KM_I(i) > TO_KM_IRI_2014(n)
                n = n + 1;
            end
            if n == 1
            I_AL_2014_v(i) = I_AL_2014(n);% if lack traffic data at the start
                section, fill in the traffic data of the first-recorded section
            I_L1_2014_v(i) = I_L1_2014(n);% if lack traffic data at the start
                section, fill in the traffic data of the first-recorded section
            I_L2_2014_v(i) = I_L2_2014(n);% if lack traffic data at the start
                section, fill in the traffic data of the first-recorded section
            I_L3_2014_v(i) = I_L3_2014(n);% if lack traffic data at the start
                section, fill in the traffic data of the first-recorded section
            else
            I_AL_2014_v(i) = I_AL_2014(n-1);
            I_L1_2014_v(i) = I_L1_2014(n-1);
            I_L2_2014_v(i) = I_L2_2014(n-1);
            I_L3_2014_v(i) = I_L3_2014(n-1);
            end
end

% fill the sections which lack of data with the data of the nearest section
for i =1:Num_TO_KM_I
    if I_AL_2014_v(i) ==0
        I_AL_2014_v(i) = I_AL_2014_v(i-1);
    end
```

```
    if I_L1_2014_v(i) ==0
        I_L1_2014_v(i) = I_L1_2014_v(i-1);
    end
    if I_L2_2014_v(i) ==0
        I_L2_2014_v(i) = I_L2_2014_v(i-1);
    end
    if I_L3_2014_v(i) ==0
        I_L3_2014_v(i) = I_L3_2014_v(i-1);
    end
end

% Secondly, cumulative intensities = days * daily flows (veh/day)
Duration_2014 = IRI_2017_ONZ(:,16);
I_AL(:,4) = Duration_2014.* I_AL_2014_v';
I_L1(:,4) = Duration_2014.* I_L1_2014_v';
I_L2(:,4) = Duration_2014.* I_L2_2014_v';
I_L3(:,4) = Duration_2014.* I_L3_2014_v';

%% Calculate Cumulative Intensities from year 2014 to year 2015
% Firstly, find the same lable of section of 2 tables (flow table and IRI
% table)
TO_KM_I = IRI_2017_ONZ(:,4);
if Direction == 'L'
    TO_KM_IRI_2015 = I_2011_2018(:,26)/1000; % Translate the unit of meter in original table to km
else
    TO_KM_IRI_2015 = I_2011_2018(:,25)/1000;
end
I_AL_2015 = I_2011_2018(:,27);
I_L1_2015 = I_2011_2018(:,28);
I_L2_2015 = I_2011_2018(:,29);
I_L3_2015 = I_2011_2018(:,30);

Num_TO_KM_I = length (TO_KM_I);
Num_TO_KM_IRI_2015 = length (TO_KM_IRI_2015);

I_AL_2015_v(:,Num_TO_KM_I ) = 0;
I_L1_2015_v(:,Num_TO_KM_I ) = 0;
I_L2_2015_v(:,Num_TO_KM_I ) = 0;
I_L3_2015_v(:,Num_TO_KM_I ) = 0;

for i =1:Num_TO_KM_I
    n = 1;
            while TO_KM_I(i) > TO_KM_IRI_2015(n)
                n = n + 1;
            end
            if n == 1
            I_AL_2015_v(i) = I_AL_2015(n);% if lack traffic data at the start section, fill in the traffic data of the first-recorded section
            I_L1_2015_v(i) = I_L1_2015(n);% if lack traffic data at the start section, fill in the traffic data of the first-recorded section
            I_L2_2015_v(i) = I_L2_2015(n);% if lack traffic data at the start section, fill in the traffic data of the first-recorded section
            I_L3_2015_v(i) = I_L3_2015(n);% if lack traffic data at the start section, fill in the traffic data of the first-recorded section
            else
```

```
            I_AL_2015_v(i) = I_AL_2015(n-1);
            I_L1_2015_v(i) = I_L1_2015(n-1);
            I_L2_2015_v(i) = I_L2_2015(n-1);
            I_L3_2015_v(i) = I_L3_2015(n-1);
            end
end

% fill the sections which lack of data with the data of the nearest section
for i =1:Num_TO_KM_I
    if I_AL_2015_v(i) ==0
        I_AL_2015_v(i) = I_AL_2015_v(i-1);
    end
    if I_L1_2015_v(i) ==0
        I_L1_2015_v(i) = I_L1_2015_v(i-1);
    end
    if I_L2_2015_v(i) ==0
        I_L2_2015_v(i) = I_L2_2015_v(i-1);
    end
    if I_L3_2015_v(i) ==0
        I_L3_2015_v(i) = I_L3_2015_v(i-1);
    end
end

% Secondly, cumulative intensities = days * daily flows (veh/day)
Duration_2015 = IRI_2017_ONZ(:,18);
I_AL(:,5) = Duration_2015.* I_AL_2015_v';
I_L1(:,5) = Duration_2015.* I_L1_2015_v';
I_L2(:,5) = Duration_2015.* I_L2_2015_v';
I_L3(:,5) = Duration_2015.* I_L3_2015_v';

%% Calculate Cumulative Intensities from year 2015 to year 2016
% Firstly, find the same lable of section of 2 tables (flow table and IRI
% table)
TO_KM_I = IRI_2017_ONZ(:,4);
if Direction == 'L'
    TO_KM_IRI_2016 = I_2011_2018(:,18)/1000; % Translate the unit of meter in original table to km
else
    TO_KM_IRI_2016 = I_2011_2018(:,17)/1000;
end
I_AL_2016 = I_2011_2018(:,19);
I_L1_2016 = I_2011_2018(:,20);
I_L2_2016 = I_2011_2018(:,21);
I_L3_2016 = I_2011_2018(:,22);

Num_TO_KM_I = length (TO_KM_I);
Num_TO_KM_IRI_2016 = length (TO_KM_IRI_2016);

I_AL_2016_v(:,Num_TO_KM_I ) = 0;
I_L1_2016_v(:,Num_TO_KM_I ) = 0;
I_L2_2016_v(:,Num_TO_KM_I ) = 0;
I_L3_2016_v(:,Num_TO_KM_I ) = 0;

for i =1:Num_TO_KM_I
    n = 1;
            while TO_KM_I(i) > TO_KM_IRI_2016(n)
```

```
            n = n + 1;
        end
        if n == 1
        I_AL_2016_v(i) = I_AL_2016(n);% if lack traffic data at the start section, fill in the traffic data of the first-recorded section
        I_L1_2016_v(i) = I_L1_2016(n);% if lack traffic data at the start section, fill in the traffic data of the first-recorded section
        I_L2_2016_v(i) = I_L2_2016(n);% if lack traffic data at the start section, fill in the traffic data of the first-recorded section
        I_L3_2016_v(i) = I_L3_2016(n);% if lack traffic data at the start section, fill in the traffic data of the first-recorded section
        else
        I_AL_2016_v(i) = I_AL_2016(n-1);
        I_L1_2016_v(i) = I_L1_2016(n-1);
        I_L2_2016_v(i) = I_L2_2016(n-1);
        I_L3_2016_v(i) = I_L3_2016(n-1);
        end
end

% fill the sections which lack of data with the data of the nearest section
for i =1:Num_TO_KM_I
    if I_AL_2016_v(i) ==0
        I_AL_2016_v(i) = I_AL_2016_v(i-1);
    end
    if I_L1_2016_v(i) ==0
        I_L1_2016_v(i) = I_L1_2016_v(i-1);
    end
    if I_L2_2016_v(i) ==0
        I_L2_2016_v(i) = I_L2_2016_v(i-1);
    end
    if I_L3_2016_v(i) ==0
        I_L3_2016_v(i) = I_L3_2016_v(i-1);
    end
end

% Secondly, cumulative intensities = days * daily flows (veh/day)
Duration_2016 = IRI_2017_ONZ(:,20);
I_AL(:,6) = Duration_2016.* I_AL_2016_v';
I_L1(:,6) = Duration_2016.* I_L1_2016_v';
I_L2(:,6) = Duration_2016.* I_L2_2016_v';
I_L3(:,6) = Duration_2016.* I_L3_2016_v';

%% Calculate Cumulative Intensities from year 2016 to the measurement in 2017
% Firstly, find the same lable of section of 2 tables (flow table and IRI
% table)
TO_KM_I = IRI_2017_ONZ(:,4);
TO_KM_IRI_2017 = I_2011_2018(:,9)/1000; % Translate the unit of meter in original table to km
I_AL_2017 = I_2011_2018(:,11);
I_L1_2017 = I_2011_2018(:,12);
I_L2_2017 = I_2011_2018(:,13);
I_L3_2017 = I_2011_2018(:,14);

Num_TO_KM_I = length (TO_KM_I);
Num_TO_KM_IRI_2017 = length (TO_KM_IRI_2017);

```

```
for i =1:Num_TO_KM_I
    n = 1;
            while TO_KM_I(i) > TO_KM_IRI_2017(n)
                n = n + 1;
            end
            if n == 1
            I_AL_2017_v(i) = I_AL_2017(n);% if lack traffic data at the start section, fill in the traffic data of the first-recorded section
            I_L1_2017_v(i) = I_L1_2017(n);% if lack traffic data at the start section, fill in the traffic data of the first-recorded section
            I_L2_2017_v(i) = I_L2_2017(n);% if lack traffic data at the start section, fill in the traffic data of the first-recorded section
            I_L3_2017_v(i) = I_L3_2017(n);% if lack traffic data at the start section, fill in the traffic data of the first-recorded section
            else
            I_AL_2017_v(i) = I_AL_2017(n-1);
            I_L1_2017_v(i) = I_L1_2017(n-1);
            I_L2_2017_v(i) = I_L2_2017(n-1);
            I_L3_2017_v(i) = I_L3_2017(n-1);
            end
end

% fill the sections which lack of data with the data of the nearest section
for i =1:Num_TO_KM_I
    if I_AL_2017_v(i) ==0
        I_AL_2017_v(i) = I_AL_2017_v(i-1);
    end
    if I_L1_2017_v(i) ==0
        I_L1_2017_v(i) = I_L1_2017_v(i-1);
    end
    if I_L2_2017_v(i) ==0
        I_L2_2017_v(i) = I_L2_2017_v(i-1);
    end
    if I_L3_2017_v(i) ==0
        I_L3_2017_v(i) = I_L3_2017_v(i-1);
    end
end

% Secondly, cumulative intensities = days * daily flows (veh/day)
Duration_2017 = IRI_2017_ONZ(:,22);
I_AL(:,7) = Duration_2017.* I_AL_2017_v';
I_L1(:,7) = Duration_2017.* I_L1_2017_v';
I_L2(:,7) = Duration_2017.* I_L2_2017_v';
I_L3(:,7) = Duration_2017.* I_L3_2017_v';

%% Calculate the cumulative intensities from the construction year to the measurement time in 2017
I_RESULT(:,1) = sum (I_AL,2);
I_RESULT(:,2) = sum (I_L1,2);
I_RESULT(:,3) = sum (I_L2,2);
I_RESULT(:,4) = sum (I_L3,2);
```

D

Weather stations

The chapter gives the locations of the weather stations which are near to A15, as well as the coverage.

Table D.1: Coverage of the weather stations

ID	Location_RD	Coverage_RD	Station	Coverage on A15
348	'(4.927 51.972)'	'MULTIPOLYGON(((5.07560967468475 51.9907448165024, 4.96041049114294 51.7704516409576, 4.69243473985064 51.7664718030671, 4.6822057149527 52.057096451638, 4.6826736127016 52.0584530782472, 4.97139065273025 52.1927196652515, 5.07560967468475 51.9907448165024)))'	'CABAUW'	80km-95km
356	'(5.145 51.858)'	'MULTIPOLYGON(((5.07560967468475 51.9907448165024, 5.27046965949741 51.9650842423707, 5.46655547246543 51.8720876619557, 5.42074128699955 51.744218218939, 5.19653888626825 51.5978336417625, 4.96041049114294 51.7704516409576, 5.07560967468475 51.9907448165024)))'	HERWIJNEN'	95km-165km
344	'(4.444 51.955)'	'MULTIPOLYGON(((4.29212439666521 52.0424159719175, 4.6822057149527 52.057096451638, 4.69243473985064 51.7664718030671, 4.6094382903389 51.6616003203507, 4.25479372927343 51.7280524570395, 4.29212439666521 52.0424159719175)))'	'ROTTERDAM'	45km-80km
330	'(4.124 51.993)'	'MULTIPOLYGON(((4.2921243966652 52.0424159719175, 4.25479372927344 51.7280524570395, 4.15001926586039 51.6847969446213, 3.59679253655277 51.9697205820329, 3.56046 52.0045365154265, 3.56046 52.9851426382979, 3.97522677559842 52.5871430302236, 4.28909551886793 52.0488216391509, 4.2921243966652 52.0424159719175)))'	'HOEK VAN HOLLAND'	25km-45km

E

Weather data process Matlab code

All the variables in the climate category are computed by the following codes. The results of the variables of T_TEMP_25, T_TEMP_0, T_TEMP_0_below, and T_PRECIPITATION are in Tab. E.1.

```
clear all

%% Load weather raw data
load('result.mat')

%% Input StationID
%StationID = 348,Weather station CABAUW , Coverage 95km-165km
%StationID = 356,Weather station HERWIJNEN', Coverage 80km-95km
%StationID = 344,Weather station 'ROTTERDAM', Coverage 45km-80km
%StationID = 330,Weather station 'HOEK VAN HOLLAND', Coverage 25km-45km
StationID = 330;% Input StationID
%% Find the weather data of the weather station which StationID is inputted above
f=find(table2array(result(:,1))==StationID);

%% Get the measurement year in the digital(double) form from the measurement time in
    text form
Time_Weather_ca = table2array(result(:,7)); % Transfer table to category
Time_Weather_text = char(Time_Weather_ca);% Transfer table to text
Year_Weather_text=[Time_Weather_text(:,1),Time_Weather_text(:,2),Time_Weather_text
    (:,3),Time_Weather_text(:,4)]; % Get year as text form
Year_Weather_num = str2num(Year_Weather_text); % Transfer year in text form to year
    in digital form

%% Caculate the number of data in every year
% Caculate the number of data in year 2015
Year_Weather_num_2015 = Year_Weather_num - 2014;
for n =1: length(Year_Weather_num)
    if Year_Weather_num_2015(n) >1|| Year_Weather_num_2015(n)<0
        Year_Weather_num_2015(n) = 0;
    end
end
Num_Data_2015 = sum(Year_Weather_num_2015);

% Caculate the number of data in year 2016
Year_Weather_num_2016 = Year_Weather_num -2015;
for n =1: length(Year_Weather_num)
    if Year_Weather_num_2016(n)>1 || Year_Weather_num_2016(n)<0
```

```
            Year_Weather_num_2016(n) = 0;
        end
end
Num_Data_2016 = sum(Year_Weather_num_2016);
% Caculate the number of data in year 2017
Year_Weather_num_2017 = Year_Weather_num -2016;
for n =1: length(Year_Weather_num)
     if Year_Weather_num_2017(n)>1 || Year_Weather_num_2017(n)<0
            Year_Weather_num_2017(n) = 0;
        end
end
Num_Data_2017 = sum(Year_Weather_num_2017);
% Caculate the number of data in year 2018
Year_Weather_num_2018 = Year_Weather_num -2017;
for n =1: length(Year_Weather_num)
     if Year_Weather_num_2018(n)>1 || Year_Weather_num_2018(n)<0
            Year_Weather_num_2018(n) = 0;
        end
end
Num_Data_2018 = sum(Year_Weather_num_2018);

%% Caculate the total time of temperature >25, 0-25, and <0
% in year 2015, 2016, 2017, 2018
TEMP = table2array(result(:,10));% Transfer table to matrix
T_TEMP_25_2015 = 0; % T_TEMP_25_2015 is the total time of temperature(>25) in 2015
T_TEMP_0_2015 = 0; % the total time of temperature (0-25) in 2015
T_TEMP_0_below_2015 = 0;% the total time of temperature (<0) in 2015

T_TEMP_25_2016 = 0; % T_TEMP_25_2016 is the total time of temperature(>25) in 2016
T_TEMP_0_2016 = 0; % the total time of temperature (0-25) in 2016
T_TEMP_0_below_2016 = 0;% the total time of temperature (<0) in 2016

T_TEMP_25_2017 = 0; % T_TEMP_25_2017 is the total time of temperature(>25) in 2017
T_TEMP_0_2017 = 0; % the total time of temperature (0-25) in 2017
T_TEMP_0_below_2017 = 0;% the total time of temperature (<0) in 2017

T_TEMP_25_2018 = 0; % T_TEMP_25_2018 is the total time of temperature(>25) in 2018
T_TEMP_0_2018 = 0; % the total time of temperature (0-25) in 2018
T_TEMP_0_below_2018 = 0;% the total time of temperature (<0) in 2018

%% Caculate the total time of rains
% in year 2015, 2016, 2017, 2018
PERCIPITATION = table2array(result(:,11));% Transfer table to matrix
T_PERCIPITATION_2015 = 0;
T_PERCIPITATION_2016 = 0;
T_PERCIPITATION_2017 = 0;
T_PERCIPITATION_2018 = 0;

%% Calculate the total time of the specfic weather occasions
% Calculate the total time of the specfic weather occasions in 2015
% The selected weather station
 for i = 1: length(f) % Select data of the weather station which StationID is input
       if f(i)>=1 &&f(i) <= Num_Data_2015 %weather occasions in 2015
              n = f(i);
         if TEMP(n)> 25 % Calculate the total time of TEMP>25 in 2015
              T_TEMP_25_2015 = T_TEMP_25_2015 + 1;
```

```
            end
            if TEMP(n)>= 0 && TEMP(n)<= 25  % Calculate the total time of 0<TEMP<25 in 2015
                T_TEMP_0_2015 = T_TEMP_0_2015 + 1;
            end
            if TEMP(n)<0 % Calculate the total time of TEMP<0 in 2015
                T_TEMP_0_below_2015 = T_TEMP_0_below_2015 + 1;
            end
            if PERCIPITATION(n)>0 % Calculate the total time of rains in 2015
                T_PERCIPITATION_2015 = T_PERCIPITATION_2015 +1;
            end
        end
 end

T_TEMP_25_2015 = T_TEMP_25_2015/60; % Transfer the unit of min to hours
T_TEMP_0_2015 = T_TEMP_0_2015/60;% Transfer the unit of min to hours
T_TEMP_0_below_2015 = T_TEMP_0_below_2015/60;% Transfer the unit of min to hours
T_PERCIPITATION_2015 = T_PERCIPITATION_2015/60; % Transfer the unit of min to hours
 %Some data is missing
 %Here assumes the temperature trend is the same in the meausrement time as
 %the inmeasurement time
 s = 5*365*24*60/Num_Data_2015; % result file contains 5 weather stations
 T_TEMP_25_2015 = T_TEMP_25_2015*s;
 T_TEMP_0_2015 = T_TEMP_0_2015*s;
 T_TEMP_0_below_2015 = T_TEMP_0_below_2015*s;
T_PERCIPITATION_2015 = T_PERCIPITATION_2015*s;
%% Calculate the total time of the specfic weather occasions in 2016
% The selected weather station
 for i = 1: length(f) % Select data of the weather station which StationID is input
        if f(i)>=Num_Data_2015+1 &&f(i) <= Num_Data_2015+Num_Data_2016 %weather occasions in 2016
                n = f(i);
            if TEMP(n)> 25 % Calculate the total time of TEMP>25 in 2016
                T_TEMP_25_2016 = T_TEMP_25_2016 + 1;
            end
            if TEMP(n)>= 0 && TEMP(n)<= 25  % Calculate the total time of 0<TEMP<25 in 2016
                T_TEMP_0_2016 = T_TEMP_0_2016 + 1;
            end
            if TEMP(n)<0 % Calculate the total time of TEMP<0 in 2016
                T_TEMP_0_below_2016 = T_TEMP_0_below_2016 + 1;
            end
            if PERCIPITATION(n)>0 % Calculate the total time of rains in 2016
                T_PERCIPITATION_2016 = T_PERCIPITATION_2016 +1;
            end
        end
 end

T_TEMP_25_2016 = T_TEMP_25_2016/60; % Transfer the unit of min to hours
T_TEMP_0_2016 = T_TEMP_0_2016/60;% Transfer the unit of min to hours
T_TEMP_0_below_2016 = T_TEMP_0_below_2016/60;% Transfer the unit of min to hours
T_PERCIPITATION_2016 = T_PERCIPITATION_2016/60;% Transfer the unit of min to hours
 %Some data is missing
 %Here assumes the temperature trend is the same in the meausrement time as
 %the inmeasurement time
 s = 5*365*24*60/Num_Data_2016; % result file contains 5 weather stations
```

```
T_TEMP_25_2016 = T_TEMP_25_2016*s;
T_TEMP_0_2016 = T_TEMP_0_2016*s;
T_TEMP_0_below_2016 = T_TEMP_0_below_2016*s;
T_PERCIPITATION_2016 = T_PERCIPITATION_2016*s;

%% Calculate the total time of the specfic weather occasions in 2017
% The selected weather station
for i = 1: length(f) % Select data of the weather station which StationID is input
      if f(i)>=Num_Data_2016+1 &&f(i) <= Num_Data_2016+Num_Data_2017 %weather
          occasions in 2017
            n = f(i);
        if TEMP(n)> 25 % Calculate the total time of TEMP>25 in 2017
            T_TEMP_25_2017 = T_TEMP_25_2017 + 1;
        end
        if TEMP(n)>= 0 && TEMP(n)<= 25  % Calculate the total time of 0<TEMP<25 in
            2017
            T_TEMP_0_2017 = T_TEMP_0_2017 + 1;
        end
        if TEMP(n)<0 % Calculate the total time of TEMP<0 in 2017
            T_TEMP_0_below_2017 = T_TEMP_0_below_2017 + 1;
        end
        if PERCIPITATION(n)>0 % Calculate the total time of rains in 2017
            T_PERCIPITATION_2017 = T_PERCIPITATION_2017 +1;
        end
      end
end

T_TEMP_25_2017 = T_TEMP_25_2017/60; % Transfer the unit of min to hours
T_TEMP_0_2017 = T_TEMP_0_2017/60;% Transfer the unit of min to hours
T_TEMP_0_below_2017 = T_TEMP_0_below_2017/60;% Transfer the unit of min to hours
T_PERCIPITATION_2017 = T_PERCIPITATION_2017/60;% Transfer the unit of min to hours
%Some data is missing
%Here assumes the temperature trend is the same in the meausrement time as
%the inmeasurement time
s = 5*365*24*60/Num_Data_2017; % result file contains 5 weather stations
T_TEMP_25_2017 = T_TEMP_25_2017*s;
T_TEMP_0_2017 = T_TEMP_0_2017*s;
T_TEMP_0_below_2017 = T_TEMP_0_below_2017*s;
T_PERCIPITATION_2017 = T_PERCIPITATION_2017*s;

%% Calculate the total time of the specfic weather occasions in 2018
% The selected weather station
for i = 1: length(f) % Select data of the weather station which StationID is input
      if f(i)>=Num_Data_2017+1 &&f(i) <= Num_Data_2017+Num_Data_2018 %weather
          occasions in 2018
            n = f(i);
        if TEMP(n)> 25 % Calculate the total time of TEMP>25 in 2018
            T_TEMP_25_2018 = T_TEMP_25_2018 + 1;
        end
        if TEMP(n)>= 0 && TEMP(n)<= 25  % Calculate the total time of 0<TEMP<25 in
            2018
            T_TEMP_0_2018 = T_TEMP_0_2018 + 1;
        end
        if TEMP(n)<0 % Calculate the total time of TEMP<0 in 2018
            T_TEMP_0_below_2018 = T_TEMP_0_below_2018 + 1;
        end
```

```
            if PERCIPITATION(n)>0 % Calculate the total time of rains in 2018
                T_PERCIPITATION_2018 = T_PERCIPITATION_2018 +1;
            end
        end
    end

T_TEMP_25_2018 = T_TEMP_25_2018/60; % Transfer the unit of min to hours
T_TEMP_0_2018 = T_TEMP_0_2018/60;% Transfer the unit of min to hours
T_TEMP_0_below_2018 = T_TEMP_0_below_2018/60;% Transfer the unit of min to hours
T_PERCIPITATION_2018 = T_PERCIPITATION_2018/60;% Transfer the unit of min to hours
 %Some data is missing
 %Here assumes the temperature trend is the same in the meausrement time as
 %the inmeasurement time
 s = 5*365*24*60/Num_Data_2018; % result file contains 5 weather stations
 T_TEMP_25_2018 = T_TEMP_25_2018*s;
 T_TEMP_0_2018 = T_TEMP_0_2018*s;
 T_TEMP_0_below_2018 = T_TEMP_0_below_2018*s;
 T_PERCIPITATION_2018 = T_PERCIPITATION_2018*s;

 %% Print results
T_TEMP = [T_TEMP_25_2015 T_TEMP_0_2015 T_TEMP_0_below_2015 T_PERCIPITATION_2015
T_TEMP_25_2016 T_TEMP_0_2016 T_TEMP_0_below_2016 T_PERCIPITATION_2016
T_TEMP_25_2017 T_TEMP_0_2017 T_TEMP_0_below_2017 T_PERCIPITATION_2017
T_TEMP_25_2018 T_TEMP_0_2018 T_TEMP_0_below_2018 T_PERCIPITATION_2018];
```

Table E.1: Calculation results of the weather variables

348 'CABAUW'				
Year	T_TEMP_25 (h)	T_TEMP_0 (h)	T_TEMP_0_below (h)	T_PERCIPITATION (h)
2015	40.32220944	8278.149597	441.5281933	1542.324511
2016	126.3606182	8197.39439	434.2392673	1817.186033
2017	196.773743	8000.432961	556.6759777	1415.139665
2018	159.59682	8028.54912	565.8432709	1459.170926
356 HERWIJNEN'				
Year	T_TEMP_25 (h)	T_TEMP_0 (h)	T_TEMP_0_below (h)	T_PERCIPITATION (h)
2015	42.33831991	8251.940161	465.721519	929.4269275
2016	164.4693761	8129.199771	466.3308529	1129.222667
2017	234.4972067	7921.927374	599.4972067	1080.726257
2018	192.7597956	7953.932425	609.3696763	1107.850653
344 'ROTTERDAM'				
Year	T_TEMP_25 (h)	T_TEMP_0 (h)	T_TEMP_0_below (h)	T_PERCIPITATION (h)
2015	40.32220944	8397.100115	322.5776755	1554.421174
2016	136.3892387	8288.654837	333.9530624	1867.329136
2017	191.6759777	8066.703911	496.5223464	1356.005587
2018	152.3424191	8097.9841	504.6990346	1402.172061
330 'HOEK VAN HOLLAND'				
Year	T_TEMP_25 (h)	T_TEMP_0 (h)	T_TEMP_0_below (h)	T_PERCIPITATION (h)
2015	42.33831991	8522.098964	195.5627158	792.3314154
2016	120.3434459	8475.187178	159.4550658	1245.554665
2017	154.972067	8267.555866	327.2765363	1256.089385
2018	133.6882453	8283.489495	332.6660988	1306.828507

F

Correlation coefficient Matlab code

The following gives the Matlab code of computing the correlation coefficients of Pearson's linear correlation coefficient, Kendall's tau coefficient, and Spearman's rho.

```
clear all
%% At first, load Correlation.xlsx to table
Matrix = xlsread('F:\RWSA15\Correlation.xlsx','Process'); % data input
%% Pearson's correlation coeffient
Size = size(Matrix);
for n = 1:Size(2)
    for i = 1:Size(2)
    rhop(n,i) = corr(Matrix(:,n),Matrix(:,i),'Type','Pearson','Rows','pairwise');
    end
end
%% Kendall correlation coeffient
Size = size(Matrix);
for n = 1:Size(2)
    for i = 1:Size(2)
    rhok(n,i) = corr(Matrix(:,n),Matrix(:,i),'Type','Kendall','Rows','pairwise');
    end
end
%% Spearman correlation coeffient
Size = size(Matrix);
for n = 1:Size(2)
    for i = 1:Size(2)
    rhos(n,i) = corr(Matrix(:,n),Matrix(:,i),'Type','Spearman','Rows','pairwise');
    end
end
```

G

Regression model Matlab code

The chapter gives the Matlab code to estimate all the parameters in the regression models defined in Chapter 6. Besides, it provides the code for computing R-square statistic, the F-statistic, its p-value, and the estimate of the error variance, that indicate the goodness of fitting.

```
clear all
%% Regression model for predicting roughness, rutting

%% Data input
R = xlsread('o:\RWSA15\RegressionModel.xlsx','Sheet 1'); % read the excel table

%% Model 1: IRI_t = n IRI_o + m  NE_t
R_t = R(:,15);%R_{t} =& predicted roughness at time t (mm/km Bump Intergrator trailer)
R_o = R(:,3);%R_{o} =& initial roughness at time t = 0, constant for given range of modified structural number
NE_t_AL = R(:,7)/1000000;% cumulative traffic at time t, million veh;
NE_t_L1 = R(:,8)/1000000;% cumulative traffic at time t, million veh;
NE_t_L2 = R(:,9)/1000000;% cumulative traffic at time t, million veh;
NE_t_L3 = R(:,10)/1000000;% cumulative traffic at time t, million veh;
X = [ones(size(R_t)) R_o NE_t_L1 NE_t_L2 NE_t_L3];
[b,bint,r,rint,stats] = regress(R_t,X);
%Plot the data and the model.
x1 = R_o;
x2 = NE_t_L3;
y = R_t;
i = 1:20:length(x1);
sz = 10;
scatter3(x1(i),x2(i),y(i),sz,'MarkerEdgeColor','k','MarkerFaceColor',[0 166/255 214/255])
hold on
x1fit = x1(i);
x2fit = x2(i);
[X1FIT,X2FIT] = meshgrid(x1fit,x2fit);
YFIT = b(1) + b(2)*X1FIT + b(5)*X2FIT ;
mesh(X1FIT,X2FIT,YFIT)
xlabel('Initial IRI (m/km)')
ylabel('Cumulative flows of heavy trucks (10^6) ')
zlabel('IRI (m/km)')
axis([0 2.5 0 15 0 2.5])
view(40,10)
```

```
grid on
hold off

%% Model 2: RUT_t = n RUT_o + m  NE_t
R_t = R(:,16);%R_{t} =& predicted rutting at time t (mm/km Bump Intergrator trailer)
R_o = R(:,4);%R_{o} =& initial rutting at time t = 0, constant for given range of modified structural number
NE_t_AL = R(:,7)/1000000;% cumulative traffic at time t, million veh;
NE_t_L1 = R(:,8)/1000000;% cumulative traffic at time t, million veh;
NE_t_L2 = R(:,9)/1000000;% cumulative traffic at time t, million veh;
NE_t_L3 = R(:,10)/1000000;% cumulative traffic at time t, million veh;
X = [ones(size(R_t)) R_o NE_t_L1 NE_t_L2 NE_t_L3];
[b,bint,r,rint,stats] = regress(R_t,X);
%Plot the data and the model.
x1 = R_o;
x2 = NE_t_L3;
y = R_t;
i = 1:20:length(x1);
sz = 10;
scatter3(x1(i),x2(i),y(i),sz,'MarkerEdgeColor','k','MarkerFaceColor',[0 166/255 214/255])
hold on
x1fit = x1(i);
x2fit = x2(i);
[X1FIT,X2FIT] = meshgrid(x1fit,x2fit);
YFIT = b(1) + b(2)*X1FIT + b(5)*X2FIT ;
mesh(X1FIT,X2FIT,YFIT)
xlabel('Initial RUT (mM)')
ylabel('Cumulative flows of heavy trucks (10^6) ')
zlabel('RUT (mm)')
axis([0 17 0 17 0 17])
view(40,10)
grid on
hold off

%% Model 3: Exponential model of Roughness progression: R_t = (R_o + m  NE_t)e^{nt}
t = R(:,1)/365; % t is Age of road section
y = log(R_t./R_o);
x1 = NE_t_AL./R_o;
x2 = t;
x(:,1) = x1;
x(:,2) = x2;
sz =1;
scatter(R_t,t,sz,'MarkerEdgeColor','k','MarkerFaceColor','k')
hold on
xlabel('Age (years)')
ylabel('IRI (m/km) ')
axis([0 3.5 0 30])
%view(50,10)
grid on
hold off

%% Model 4: Regression model of Roughness increment: R_var = m NE_L1_var * exp(nt} + m2 NE_L2_var * exp(nt}+ m3 NE_L3_var * exp(nt} a RUT_var  + n  R_o t_var
R_var = R(:,17);
NE_L1_var =R(:,12)/1000000;
```

```
NE_L2_var =R(:,13)/1000000;
NE_L3_var =R(:,14)/1000000;
t = R(:,1)/365;
RUT_var = R(:,18);
t_var = R(:,6)/365;
X = [ones(size(R_var)) NE_L1_var.*exp(t) NE_L2_var.*exp(t) NE_L3_var.*exp(t) RUT_var
    R_o.*t_var];
[b,bint,r,rint,stats] = regress(R_var,X);

%% Model 5: Non-linear regression model of rutting: RUT = m( NE_L1 + NE_L2+ NE_L3
    ) T^x +w
R_t = R(:,16);%R_{t} =& predicted rutting at time t (mm/km Bump Intergrator trailer)
LN_RUT = log(R_t);
LN_NE = log(NE_t_L1+NE_t_L2+NE_t_L3);
y = LN_RUT-LN_NE;
t = R(:,1)/365;
for n = 1:length(t) %% To calculate ln(t), t>=0
    if t(n) <0
        t(n) = 0.0000001;
    end
end
LN_t = log(t);
sz = 10;
%X = [ones(size(y)) LN_t];
%[b,bint,r,rint,stats] = regress(y,X);
%Plot the regreesion function and the points
x1 = -5:0.001:5;
y1 = -0.8894*x1-0.6486; % The regression function has been found
plot(LN_t,y, 'k.',x1,y1,'MarkerFaceColor',[0 0.4470 0.7410])
xlabel('Logrithm of age ')
ylabel('ln(RUT) - ln(I AL) ')
axis([-4 4 -4 4])

%% Model 6: Non-linear regression model of rutting: RUT = mln(( NE_L1 + NE_L2+
    NE_L3) T^x )+w
R_t = R(:,16);%R_{t} =& predicted rutting at time t (mm/km Bump Intergrator trailer)
LN_RUT = log(R_t);
LN_NE = log(NE_t_L1+NE_t_L2+NE_t_L3);
y = R_t- LN_NE;
t = R(:,1)/365;
for n = 1:length(t) %% To calculate ln(t), t>=0
    if t(n) <0
        t(n) = 0.0000001;
    end
end
LN_t = log(t);
X = [ones(size(y)) LN_t];
[b,bint,r,rint,stats] = regress(y,X);
%Plot the regreesion function and the points
x1 = -5:0.001:5;
y1 = -0.6326*x1+3.689; % The regression function has been found
plot(LN_t,y, 'k.',x1,y1,'MarkerFaceColor',[0 0.4470 0.7410])
xlabel('Logrithm of age ')
ylabel('RUT - ln(I AL) ')
axis([-4 4 -1 20])
```

H

Matlab code for defining the thresholds of the performance indicators

The chapter gives the sample Matlab code of computing the threshold values of the performance indicators, that is the second step of the survival model. The following code is to formulate the critical value of the road roughness on A15 in the WNZZ district region based on the data from 2015 to 2018. The concept can also use for the computation of the threshold values of the other performance indexes, like the rutting, in other roadways. In the study, The critical levels of roughness in ONZ district and rutting in WNZZ and ONZ districts are also studied. But limited by the length of the thesis, the code is not included.

```
clear all
%% Data input
%IRI_classification
%Zoab_IRI_1HRL_WNZZ
From_km_2015_1HRL = xlsread('f:\RWSA15\Zoab_WNZZ_IRI.xlsx','2015_1HRL','D2:D350');
IRI_2015_1HRL = xlsread('f:\RWSA15\Zoab_WNZZ_IRI.xlsx','2015_1HRL','O2:O350');
From_km_2016_1HRL= xlsread('f:\RWSA15\Zoab_WNZZ_IRI.xlsx','2016_1HRL','D2:D350');
IRI_2016_1HRL = xlsread('f:\RWSA15\Zoab_WNZZ_IRI.xlsx','2016_1HRL','O2:O350');
From_km_2017_1HRL= xlsread('f:\RWSA15\Zoab_WNZZ_IRI.xlsx','2017_1HRL','D2:D350');
IRI_2017_1HRL = xlsread('f:\RWSA15\Zoab_WNZZ_IRI.xlsx','2017_1HRL','O2:O350');
From_km_2018_1HRL = xlsread('f:\RWSA15\Zoab_WNZZ_IRI.xlsx','2018_1HRL','D2:D350');
IRI_2018_1HRL = xlsread('f:\RWSA15\Zoab_WNZZ_IRI.xlsx','2018_1HRL','O2:O350');
%Zoab_IRI_1HRR_WNZZ
From_km_2015_1HRR = xlsread('f:\RWSA15\Zoab_WNZZ_IRI.xlsx','2015_1HRR','D2:D350');
IRI_2015_1HRR = xlsread('f:\RWSA15\Zoab_WNZZ_IRI.xlsx','2015_1HRR','O2:O350');
From_km_2016_1HRR = xlsread('f:\RWSA15\Zoab_WNZZ_IRI.xlsx','2016_1HRR','D2:D350');
IRI_2016_1HRR = xlsread('f:\RWSA15\Zoab_WNZZ_IRI.xlsx','2016_1HRR','O2:O350');
From_km_2017_1HRR = xlsread('f:\RWSA15\Zoab_WNZZ_IRI.xlsx','2017_1HRR','D2:D350');
IRI_2017_1HRR = xlsread('f:\RWSA15\Zoab_WNZZ_IRI.xlsx','2017_1HRR','O2:O350');
From_km_2018_1HRR = xlsread('f:\RWSA15\Zoab_WNZZ_IRI.xlsx','2018_1HRR','D2:D350');
IRI_2018_1HRR = xlsread('f:\RWSA15\Zoab_WNZZ_IRI.xlsx','2018_1HRR','O2:O350');

IRI_Diff = zeros(350,3); %IRI variance during 1 year
N_IRI_2015 = length (IRI_2015_1HRL); % Number of IRI data in 2015
N_IRI_2016 = length (IRI_2016_1HRL); % Number of IRI data in 2016
N_IRI_2017 = length (IRI_2017_1HRL); % Number of IRI data in 2017
N_IRI_2018 = length (IRI_2018_1HRL); % Number of IRI data in 2018
%% IRI variance in 2016
%Zoab_IRI_1HRL_WNZZ
for i=1:N_IRI_2016
```

```
    n = 1; %row
    while From_km_2015_1HRL(n) ~= From_km_2016_1HRL(i) && n < N_IRI_2015
       n = n +1;
    end
    if From_km_2015_1HRL(n) == From_km_2016_1HRL(i)
        IRI_Diff(i,1) = IRI_2016_1HRL(i)- IRI_2015_1HRL(n);
        x_2015_1HRL(i) = IRI_2015_1HRL(n);
        y_2015_1HRL(i) = IRI_Diff(i,1);
    end
end
%% IRI variance in 2017
%Zoab_IRI_1HRL_WNZZ
for i=1:N_IRI_2017
    n = 1; %row
    while From_km_2016_1HRL(n) ~= From_km_2017_1HRL(i) && n <= N_IRI_2016
            n = n +1;
    end
    if From_km_2016_1HRL(n) == From_km_2017_1HRL(i)
        IRI_Diff(i,2) = IRI_2017_1HRL(i)- IRI_2016_1HRL(n);
        x_2016_1HRL(i) = IRI_2016_1HRL(n);
        y_2016_1HRL(i) = IRI_Diff(i,2);
    end
end
%% IRI variance in 2018
%Zoab_IRI_1HRL_WNZZ
for i=1:N_IRI_2018
    n = 1; %row
    while From_km_2017_1HRL(n) ~= From_km_2018_1HRL(i) && n < N_IRI_2017
        n = n+1;
    end
     if From_km_2017_1HRL(n) == From_km_2018_1HRL(i)
        IRI_Diff(i,3) = IRI_2018_1HRL(i)- IRI_2017_1HRL(n);
        x_2017_1HRL(i) = IRI_2017_1HRL(n);
        y_2017_1HRL(i) = IRI_Diff(i,3);
    end
end

%% Zoab_IRI_1HRR
%IRI_classification
IRI_Diff = zeros(350,3); %IRI variance during 1 year
N_IRI_2015 = length (IRI_2015_1HRR); % Number of IRI data in 2015
N_IRI_2016 = length (IRI_2016_1HRR); % Number of IRI data in 2016
N_IRI_2017 = length (IRI_2017_1HRR); % Number of IRI data in 2017
N_IRI_2018 = length (IRI_2018_1HRR); % Number of IRI data in 2018
%% IRI variance in 2016
% Zoab_IRI_1HRR
for i=1:N_IRI_2016
    n = 1; %row
    while From_km_2015_1HRR(n) ~= From_km_2016_1HRR(i) && n < N_IRI_2015
       n = n +1;
    end
     if From_km_2015_1HRR(n) == From_km_2016_1HRR(i)
        IRI_Diff(i,1) = IRI_2016_1HRR(i)- IRI_2015_1HRR(n);
        x_2015_1HRR(i) = IRI_2015_1HRR(n);
        y_2015_1HRR(i) = IRI_Diff(i,1);
    end
```

```
end
%% IRI variance in 2017
% Zoab_IRI_1HRR
for i=1:N_IRI_2017
    n = 1; %row
    while From_km_2016_1HRR(n) ~= From_km_2017_1HRR(i) && n < N_IRI_2016
            n = n +1;
    end
    if From_km_2016_1HRR(n) == From_km_2017_1HRR(i)
        IRI_Diff(i,2) = IRI_2017_1HRR(i)- IRI_2016_1HRR(n);
        x_2016_1HRR(i) = IRI_2016_1HRR(n);
        y_2016_1HRR(i) = IRI_Diff(i,2);
    end
end
%% IRI variance in 2018
% Zoab_IRI_1HRR
for i=1:N_IRI_2018
    n = 1; %row
    while From_km_2017_1HRR(n) ~= From_km_2018_1HRR(i) && n < N_IRI_2017
        n = n+1;
    end
    if From_km_2017_1HRR(n) == From_km_2018_1HRR(i)
        IRI_Diff(i,3) = IRI_2018_1HRR(i)- IRI_2017_1HRR(n);
        x_2017_1HRR(i) = IRI_2017_1HRR(n);
        y_2017_1HRR(i) = IRI_Diff(i,3);
    end
end
x=[x_2015_1HRL,x_2016_1HRL,x_2017_1HRL,x_2015_1HRR,x_2016_1HRR,x_2017_1HRR];
y=[y_2015_1HRL,y_2016_1HRL,y_2017_1HRL,y_2015_1HRR,y_2016_1HRR,y_2017_1HRR];

subplot(1,2,1),scatter(x,y)
title ('IRI variance 2015-2018')
xlabel('IRI value 2015-2018')
ylabel('IRI variance during next year')
axis([0 4 -2 1.5])

%% Median of IRI variance in next year
interval = (0:0.1:4.5); %Intervals of IRI
for i = 1:length(interval)
    a = find (x>=interval(i)&x<interval(i)+0.1);
    Median(i) = median(y(a),'omitnan');%Median of IRI variance in next year

end
subplot(1,2,2),plot(interval,Median,'k-*','MarkerIndices',1:1:length(y))
title ('Median filtering of IRI variance 2015-2018')
xlabel('IRI (m/km)')
ylabel('Median of IRI annual variance')
axis([0 4 -2 2])
hold on
grid on
hold off
```

I

Survival model Matlab code

The chapter gives the Matlab code to define the distribution of the survival time on A15. The following code is able to compute the probability density of the time to failure regarding IRI. And the model is also applied to the rut depth, with the input of the measurmenrt data of rutting and the definition of the critical value (in the study, it is 10 mm as Chapter 7.2).

```
clear all
%% Actual zoab IRI failure observation
% ONZ DISTRICT
IRI_2015_1HRL = xlsread('f:\RWSA15\Zoab_ONZ_IRI.xlsx','2015_1HRL','O2:O220');
Time_2015_1HRL = xlsread('f:\RWSA15\Zoab_ONZ_IRI.xlsx','2015_1HRL','T2:T220');
IRI_2016_1HRL = xlsread('f:\RWSA15\Zoab_ONZ_IRI.xlsx','2016_1HRL','O2:O220');
Time_2016_1HRL = xlsread('f:\RWSA15\Zoab_ONZ_IRI.xlsx','2016_1HRL','T2:T220');
IRI_2017_1HRL = xlsread('f:\RWSA15\Zoab_ONZ_IRI.xlsx','2017_1HRL','O2:O220');
Time_2017_1HRL = xlsread('f:\RWSA15\Zoab_ONZ_IRI.xlsx','2017_1HRL','T2:T220');
IRI_2018_1HRL = xlsread('f:\RWSA15\Zoab_ONZ_IRI.xlsx','2018_1HRL','O2:O220');
Time_2018_1HRL = xlsread('f:\RWSA15\Zoab_ONZ_IRI.xlsx','2018_1HRL','T2:T220');

IRI_2015_1HRR = xlsread('f:\RWSA15\Zoab_ONZ_IRI.xlsx','2015_1HRR','O2:O220');
Time_2015_1HRR = xlsread('f:\RWSA15\Zoab_ONZ_IRI.xlsx','2015_1HRR','T2:T220');
IRI_2016_1HRR = xlsread('f:\RWSA15\Zoab_ONZ_IRI.xlsx','2016_1HRR','O2:O220');
Time_2016_1HRR = xlsread('f:\RWSA15\Zoab_ONZ_IRI.xlsx','2016_1HRR','T2:T220');
IRI_2017_1HRR = xlsread('f:\RWSA15\Zoab_ONZ_IRI.xlsx','2017_1HRR','O2:O220');
Time_2017_1HRR = xlsread('f:\RWSA15\Zoab_ONZ_IRI.xlsx','2017_1HRR','T2:T220');
IRI_2018_1HRR = xlsread('f:\RWSA15\Zoab_ONZ_IRI.xlsx','2018_1HRR','O2:O220');
Time_2018_1HRR = xlsread('f:\RWSA15\Zoab_ONZ_IRI.xlsx','2018_1HRR','T2:T220');

IRI_CRI = 2.3; % Define IRI critical value

N_IRI_2015 = length (IRI_2015_1HRL); % Number of IRI data in 2015
n = 0;% No. of IRI larger than critical
for i=1:N_IRI_2015
    if IRI_2015_1HRL(i) >= IRI_CRI && Time_2015_1HRL(i)>=0
        n = n + 1;
        FT_ONZ(n) = Time_2015_1HRL(i); % FT is the failure time, which is the duration between the construction year and the observation date of the critical situation
    end
end
N_IRI_2015 = length (IRI_2015_1HRR); % Number of IRI data in 2015
for i=1:N_IRI_2015
```

```
    if IRI_2015_1HRR(i) >= IRI_CRI && Time_2015_1HRR(i)>=0
        n = n + 1 ;
         FT_ONZ(n) = Time_2015_1HRR(i); % FT is the failure time, which is the duration between the construction year and the observation date of the critical situation
    end
end
N_IRI_2016 = length (IRI_2016_1HRL); % Number of IRI data in 2016
for i=1:N_IRI_2016
    if IRI_2016_1HRL(i) >= IRI_CRI && Time_2016_1HRL(i)>=0
        n = n +1;
        FT_ONZ(n) = Time_2016_1HRL(i); % FT is the failure time, which is the duration between the construction year and the observation date of the critical situation
    end
end
N_IRI_2016 = length (IRI_2016_1HRR); % Number of IRI data in 2016
for i=1:N_IRI_2016
    if IRI_2016_1HRR(i) >= IRI_CRI && Time_2016_1HRR(i)>=0
        n = n +1;
        FT_ONZ(n) = Time_2016_1HRR(i); % FT is the failure time, which is the duration between the construction year and the observation date of the critical situation
    end
end
N_IRI_2017 = length (IRI_2017_1HRL); % Number of IRI data in 2017
for i=1:N_IRI_2017
    if IRI_2017_1HRL(i) >= IRI_CRI && Time_2017_1HRL(i)>=0
        n = n +1;
        FT_ONZ(n) = Time_2017_1HRL(i); % FT is the failure time, which is the duration between the construction year and the observation date of the critical situation
    end
end
N_IRI_2017 = length (IRI_2017_1HRR); % Number of IRI data in 2017
for i=1:N_IRI_2017
    if IRI_2017_1HRR(i) >= IRI_CRI && Time_2017_1HRR(i)>=0
        n = n +1;
        FT_ONZ(n) = Time_2017_1HRR(i); % FT is the failure time, which is the duration between the construction year and the observation date of the critical situation
    end
end
N_IRI_2018 = length (IRI_2018_1HRL); % Number of IRI data in 2018
for i=1:N_IRI_2018
    if IRI_2018_1HRL(i) >= IRI_CRI && Time_2018_1HRL(i)>=0
        n = n +1;
        FT_ONZ(n) = Time_2018_1HRL(i); % FT is the failure time, which is the duration between the construction year and the observation date of the critical situation
    end
end
N_IRI_2018 = length (IRI_2018_1HRR); % Number of IRI data in 2018
for i=1:N_IRI_2018
    if IRI_2018_1HRR(i) >= IRI_CRI && Time_2018_1HRR(i)>=0
        n = n +1;
```

```
        FT_ONZ(n) = Time_2018_1HRR(i); % FT is the failure time, which is the duration between the construction year and the observation date of the critical situation
    end
end

for i=1:n
    FT_ONZ(i) = FT_ONZ(i)/365; %Transfer the failure time from days to years
end

%% Actual zoab IRI failure observation
% WNZZ DISTRICT
IRI_2015_1HRL = xlsread('f:\RWSA15\Zoab_WNZZ_IRI.xlsx','2015_1HRL','O2:O350');
Time_2015_1HRL = xlsread('f:\RWSA15\Zoab_WNZZ_IRI.xlsx','2015_1HRL','T2:T350');
IRI_2016_1HRL = xlsread('f:\RWSA15\Zoab_WNZZ_IRI.xlsx','2016_1HRL','O2:O350');
Time_2016_1HRL = xlsread('f:\RWSA15\Zoab_WNZZ_IRI.xlsx','2016_1HRL','T2:T350');
IRI_2017_1HRL = xlsread('f:\RWSA15\Zoab_WNZZ_IRI.xlsx','2017_1HRL','O2:O350');
Time_2017_1HRL = xlsread('f:\RWSA15\Zoab_WNZZ_IRI.xlsx','2017_1HRL','T2:T350');
IRI_2018_1HRL = xlsread('f:\RWSA15\Zoab_WNZZ_IRI.xlsx','2018_1HRL','O2:O350');
Time_2018_1HRL = xlsread('f:\RWSA15\Zoab_WNZZ_IRI.xlsx','2018_1HRL','T2:T350');

IRI_2015_1HRR = xlsread('f:\RWSA15\Zoab_WNZZ_IRI.xlsx','2015_1HRR','O2:O350');
Time_2015_1HRR = xlsread('f:\RWSA15\Zoab_WNZZ_IRI.xlsx','2015_1HRR','T2:T350');
IRI_2016_1HRR = xlsread('f:\RWSA15\Zoab_WNZZ_IRI.xlsx','2016_1HRR','O2:O350');
Time_2016_1HRR = xlsread('f:\RWSA15\Zoab_WNZZ_IRI.xlsx','2016_1HRR','T2:T350');
IRI_2017_1HRR = xlsread('f:\RWSA15\Zoab_WNZZ_IRI.xlsx','2017_1HRR','O2:O350');
Time_2017_1HRR = xlsread('f:\RWSA15\Zoab_WNZZ_IRI.xlsx','2017_1HRR','T2:T350');
IRI_2018_1HRR = xlsread('f:\RWSA15\Zoab_WNZZ_IRI.xlsx','2018_1HRR','O2:O350');
Time_2018_1HRR = xlsread('f:\RWSA15\Zoab_WNZZ_IRI.xlsx','2018_1HRR','T2:T350');

IRI_CRI = 2.3; % Define IRI critical value

N_IRI_2015 = length (IRI_2015_1HRL); % Number of IRI data in 2015
n = 0;% No. of IRI larger than critical
for i=1:N_IRI_2015
    if IRI_2015_1HRL(i) >= IRI_CRI && Time_2015_1HRL(i)>=0
        n = n + 1;
        FT_WNZZ(n) = Time_2015_1HRL(i); % FT is the failure time, which is the duration between the construction year and the observation date of the critical situation
    end
end
N_IRI_2015 = length (IRI_2015_1HRR); % Number of IRI data in 2015
for i=1:N_IRI_2015
    if IRI_2015_1HRR(i) >= IRI_CRI && Time_2015_1HRR(i)>=0
        n = n + 1 ;
         FT_WNZZ(n) = Time_2015_1HRR(i); % FT is the failure time, which is the duration between the construction year and the observation date of the critical situation
    end
end
N_IRI_2016 = length (IRI_2016_1HRL); % Number of IRI data in 2016
for i=1:N_IRI_2016
    if IRI_2016_1HRL(i) >= IRI_CRI && Time_2016_1HRL(i)>=0
        n = n +1;
```

```
        FT_WNZZ(n) = Time_2016_1HRL(i); % FT is the failure time, which is the duration between the construction year and the observation date of the critical situation
    end
end
N_IRI_2016 = length (IRI_2016_1HRR); % Number of IRI data in 2016
for i=1:N_IRI_2016
    if IRI_2016_1HRR(i) >= IRI_CRI && Time_2016_1HRR(i)>=0
        n = n +1;
        FT_WNZZ(n) = Time_2016_1HRR(i); % FT is the failure time, which is the duration between the construction year and the observation date of the critical situation
    end
end
N_IRI_2017 = length (IRI_2017_1HRL); % Number of IRI data in 2017
for i=1:N_IRI_2017
    if IRI_2017_1HRL(i) >= IRI_CRI && Time_2017_1HRL(i)>=0
        n = n +1;
        FT_WNZZ(n) = Time_2017_1HRL(i); % FT is the failure time, which is the duration between the construction year and the observation date of the critical situation
    end
end
N_IRI_2017 = length (IRI_2017_1HRR); % Number of IRI data in 2017
for i=1:N_IRI_2017
    if IRI_2017_1HRR(i) >= IRI_CRI && Time_2017_1HRR(i)>=0
        n = n +1;
        FT_WNZZ(n) = Time_2017_1HRR(i); % FT is the failure time, which is the duration between the construction year and the observation date of the critical situation
    end
end
N_IRI_2018 = length (IRI_2018_1HRL); % Number of IRI data in 2018
for i=1:N_IRI_2018
    if IRI_2018_1HRL(i) >= IRI_CRI && Time_2018_1HRL(i)>=0
        n = n +1;
        FT_WNZZ(n) = Time_2018_1HRL(i); % FT is the failure time, which is the duration between the construction year and the observation date of the critical situation
    end
end
N_IRI_2018 = length (IRI_2018_1HRR); % Number of IRI data in 2018
for i=1:N_IRI_2018
    if IRI_2018_1HRR(i) >= IRI_CRI && Time_2018_1HRR(i)>=0
        n = n +1;
        FT_WNZZ(n) = Time_2018_1HRR(i); % FT is the failure time, which is the duration between the construction year and the observation date of the critical situation
    end
end

for i=1:n
    FT_WNZZ(i) = FT_WNZZ(i)/365; %Transfer the failure time from days to years
end

%% Diagram Density distribution of survival time
```

```
% T = Servival time
for i = 1: length(FT_WNZZ)
    T(1,i) = FT_WNZZ(i);
end

for i = length(FT_WNZZ)+1: length(FT_WNZZ)+length(FT_ONZ)
    T(1,i) = FT_ONZ(i-length(FT_WNZZ));
end
% Plot the distrobution
binWidth = 2;
lastVal = ceil(max(T));
binEdges = 0:binWidth:lastVal+1;
h = histogram(T,binEdges,'Normalization','pdf','FaceColor',[.9 .9 .9]);
xlabel('Time to failure(years)');
ylabel('Probability Density');
ylim([0 0.2]);
```

J

DTC model Python code

The chapter gives the Python code to model DTC defined in Section 3.5. The code is to classify the pavement regarding the road roughness. The similar program is used to categorise the road section concerning the rutting. Limited by the length of the report, the code concerning the rutting does not display.

```
import pandas as pd
import numpy as np
from sklearn.tree import DecisionTreeClassifier, export_graphviz
from sklearn import tree
from sklearn.model_selection import train_test_split
from sklearn.metrics import accuracy_score
from sklearn.metrics import classification_report
from sklearn.metrics import confusion_matrix
from matplotlib import pyplot as plt
import seaborn as sns
import graphviz
import pydotplus
import io
from scipy import misc
from IPython.display import display

# load dataset
df = pd.read_excel('DTC.xlsx',sheet_name = 'Train_IRI')
df_test = pd.read_excel('DTC.xlsx',sheet_name = 'Test_IRI')
df.shape

# target vector
df_test.head()
df_test.shape
train = df
test = df_test

features = ['AGE_IRI','SURFACE_COMBID',
            'SURFACE_DAB','SURFACE_EAB','SURFACE_OAB','SURFACE_SMA','SURFACE_ZOAB','
                SURFACE_ZOAB+','SURFACE_ZOABTW','SURFACE_ZOEAB',
            'IRI_VALUE_0',
            'I_L1','I_L2','I_L3','T_TEMP_25','T_TEMP_0','T_TEMP_0_below','
                T_PERCIPITATION']

# Decision tree
c = DecisionTreeClassifier(min_samples_split=100,random_state = 0)
```

```python

x_train = train[features]
y_train = train["IRI_CLASS"]
x_test = test[features]
y_test = test["IRI_CLASS"]

dt = c.fit(x_train, y_train)
def show_tree(tree, features, path):
    f = io.StringIO()
    export_graphviz(tree, out_file = f, feature_names = features, class_names = '01')
    pydotplus.graph_from_dot_data(f.getvalue()).write_png(path)
    img = misc.imread(path)
    plt.rcParams["figure.figsize"] = (20,20)
    plt.imshow(img)

show_tree(dt, features, 'roughness100.png')
```

K

Prediction results of decision tree classifier model

The chapter gives the prediction result of A15 by DTC model. The training data of the model is the measurement data of A15 from 2015 to 2018. And the outputs are the lists of the A15 road sections which are predicted to meet the maintenance requirement concerning road roughness or rutting on January 1st, 2019 in the following tables.

Table K.1: List of the A15 road sections which are predicted to meet the maintenance requirement concerning road roughness on January 1st, 2019 by DTC model (Part I)

ROAD	DIREC TION	FROM _KM	AGE _IRI	AGE _RUT	SURFACE _LAYER	I_AL	I_L1	I_L2	I_L3	IRI_ VALUE _0	RUT_ VALUE _0	T_ TEMP _25	T_ TEMP _0	T_ TEMP _0_ below	T_ PERCI PITA TION
15	1HRR	91.6	8507	8507	ZOAB	7.26E+08	6.13E+08	39721271	73700905	1.4	8	164.469	8129.2	466.331	1129.22
15	1HRL	123.1	3641	3641	ZOAB	######	######	######	######	3.0	14	126.361	8197.39	434.239	1817.19
15	1HRR	151.6	993	993	ZOEAB	17335000	14117417	1299010	1918604	2.96	19	196.7737	8000.433	556.676	1415.14
15	1HRR	114.5	3560	3560	ZOAB+	55438460	41836850	4309331	9292320	2.99	14	196.7737	8000.433	556.676	1415.14
15	1HRR	96.8	5385	5385	ZOAB	1.46E+08	1.2E+08	10187693	17160939	3.38	13	196.7737	8000.433	556.676	1415.14
15	1HRL	118.9	262	262	ZOEAB	527000	406782	36022	84227	2.61	15	196.7737	8000.433	556.676	1415.14
15	1HRL	118	293	293	ZOEAB	1683300	1347291	120063	215946	2.93	12	196.7737	8000.433	556.676	1415.14
15	1HRR	111.9	3560	3560	ZOAB+	86730541	68188635	6307021	12145805	2.69	12	196.7737	8000.433	556.676	1415.14
15	1HRR	151.5	993	993	ZOEAB	17335000	14117417	1299010	1918604	3.3	15	196.7737	8000.433	556.676	1415.14
15	1HRR	114.2	3560	3560	ZOAB+	55438460	41836850	4309331	9292320	3.2	16	196.7737	8000.433	556.676	1415.14
15	1HRR	117.1	3277	3277	ZOAB+	######	######	######	######	2.6	7	126.361	8197.39	434.239	1817.19
15	1HRR	115.1	1003	1003	ZOAB+	12486300	9622653	888425	1975263	3.29	8	196.7737	8000.433	556.676	1415.14
15	1HRL	94.2	7960	7960	DAB	5.87E+08	4.96E+08	35614960	56060134	2.6	7	164.469	8129.2	466.331	1129.22
15	1HRR	106.1	1125	1125	ZOEAB	23642800	18614374	1812025	3216442	3.18	9	196.7737	8000.433	556.676	1415.14
15	1HRL	94.4	3943	3943	ZOAB	2.77E+08	2.34E+08	16811383	26446810	2.8	6	164.469	8129.2	466.331	1129.22
15	1HRR	152	993	993	ZOEAB	17335000	14117417	1299010	1918604	2.64	12	196.7737	8000.433	556.676	1415.14
15	1HRR	119.6	3550	3550	ZOAB+	1.11E+08	90888344	7393024	12920234	3.43	7	196.7737	8000.433	556.676	1415.14
15	1HRR	123.2	3641	3641	ZOAB	######	######	######	######	2.6	4	126.361	8197.39	434.239	1817.19
15	1HRR	122.1	3550	3550	ZOAB+	1.12E+08	90540024	7950436	13048028	3.19	7	196.7737	8000.433	556.676	1415.14
15	1HRR	98.3	1003	1003	ZOAB+	25115400	20941065	1496965	2677411	3.6	6	196.7737	8000.433	556.676	1415.14
15	1HRR	26.6	381	381	ZOAB+	1590000	1008450	106200	475350	3.2	4	120.343	8475.19	159.455	1245.55
15	1HRL	151	715	715	ZOAB+	12228900	10209855	939487	1079223	2.86	6	196.7737	8000.433	556.676	1415.14
15	1HRL	64.5	2983	2983	ZOAB+	92649640	73156934	6915659	12577121	2.85	4	191.676	8066.704	496.5223	1356.006
15	1HRR	26.6	761	761	EAB	5601800	3553222	390458	1657900	2.59	6	154.9721	8267.556	327.2765	1256.089
15	1HRL	64.8	2983	2983	ZOAB+	1.29E+08	1.07E+08	8176794	13897372	3.19	5	191.676	8066.704	496.5223	1356.006
15	1HRR	64.1	2983	2983	ZOAB	73790080	58512525	5994401	9283154	4	4	191.676	8066.704	496.5223	1356.006
15	1HRL	81.6	1173	1173	ZOAB	64141800	53191515	4348659	6601626	3.4	2	164.469	8129.2	466.331	1129.22
15	1HRR	63.5	2192	2192	ZOAB+	37500552	29739089	3007405	5105320	3.92	5	191.676	8066.704	496.5223	1356.006
15	1HRR	26.7	381	381	ZOAB+	1590000	1008450	106200	475350	3.5	3	120.343	8475.19	159.455	1245.55
15	1HRR	64.1	2622	2622	ZOAB	1.24E+08	99280827	9370791	15435582	3.2	3	136.389	8288.65	333.953	1867.33
15	1HRL	64.5	2622	2622	ZOAB+	1.27E+08	1.1E+08	5682579	11780553	2.7	4	136.389	8288.65	333.953	1867.33
15	1HRR	26.7	761	761	EAB	5601800	3553222	390458	1657900	3.53	5	154.9721	8267.556	327.2765	1256.089
15	1HRR	127.8	2273	2273	ZOAB	66798978	54419322	4892086	7487235	3.14	5	196.7737	8000.433	556.676	1415.14

Table K.2: List of the A15 road sections which are predicted to meet the maintenance requirement concerning road roughness on January 1st, 2019 by DTC model (Part II)

ROAD	DIREC TION	FROM _KM	AGE _IRI	AGE _RUT	SURFACE _LAYER	I_AL	I_L1	I_L2	I_L3	IRI_ VALUE _0	RUT_ VALUE _0	T_ TEMP _25	T_ TEMP _0	T_ TEMP _0_ below	T_ PERCI PITA TION
15	1HRL	70.1	2496	2496	ZOAB+	51626766	43775986	2889761	4961092	3.28	4	191.676	8066.704	496.5223	1356.006
15	1HRR	130.1	2454	2454	ZOAB	72490226	59025993	5271136	8192762	3.18	5	196.7737	8000.433	556.676	1415.14
15	1HRL	162.1	4903	4903	SMA	11531100	10576072	592273	363318	2.6	10	159.597	8028.55	565.843	1459.17
15	1HRL	143.9	3925	3925	ZOAB	28775600	23491536	2176963	3107822	3.0	4	159.597	8028.55	565.843	1459.17
15	1HRL	146.8	2709	2709	ZOAB	27305200	22256330	2213990	2835439	2.9	3	159.597	8028.55	565.843	1459.17
15	1HRR	128.0	2608	2608	ZOAB	84962178	69159498	6301846	9500499	2.9	5	159.597	8028.55	565.843	1459.17
15	1HRR	160.8	2373	2373	EAB	1.04E+08	91257528	6320020	6581678	2.9	8	159.597	8028.55	565.843	1459.17
15	1HRR	118.3	2801	2801	ZOAB	1.2E+08	96394950	7955625	15286356	2.8	5	159.597	8028.55	565.843	1459.17
15	1HRR	158.9	3347	3347	DAB	1.2E+08	1.04E+08	8008849	8205821	2.6	8	159.597	8028.55	565.843	1459.17
15	1HRR	117.7	4078	4078	ZOAB+	1.25E+08	1E+08	8231809	15924452	2.8	12	159.597	8028.55	565.843	1459.17
15	1HRR	159.7	6511	6511	SMA	1.26E+08	1.08E+08	8764348	9172085	3.5	10	159.597	8028.55	565.843	1459.17
15	1HRR	159.8	6511	6511	SMA	1.26E+08	1.08E+08	8764348	9172085	2.7	11	159.597	8028.55	565.843	1459.17
15	1HRR	159.5	6842	6842	SMA	1.33E+08	1.14E+08	9252948	9682140	2.7	8	159.597	8028.55	565.843	1459.17
15	1HRR	159.9	6842	6842	SMA	1.33E+08	1.14E+08	9272313	9702135	2.8	15	159.597	8028.55	565.843	1459.17
15	1HRR	124.6	4442	4442	ZOAB	1.51E+08	1.25E+08	9970163	16228663	2.7	8	159.597	8028.55	565.843	1459.17
15	1HRR	160.1	6842	6842	SMA	1.51E+08	1.32E+08	9185926	9553352	3.8	14	159.597	8028.55	565.843	1459.17
15	1HRR	160.2	6842	6842	SMA	1.52E+08	1.33E+08	9208381	9575777	2.9	12	159.597	8028.55	565.843	1459.17
15	1HRR	158.3	6511	6511	SMA	1.65E+08	1.42E+08	11156182	11579115	2.6	7	159.597	8028.55	565.843	1459.17
15	1HRR	159.3	6842	6842	SMA	1.74E+08	1.5E+08	11767547	12234373	2.6	16	159.597	8028.55	565.843	1459.17
15	1HRL	26.7	1097	1097	EAB	11600200	6980666	923569	3696904	3.0	3	133.688	8283.49	332.666	1306.83
15	1HRL	70.1	2923	2923	ZOAB+	1.98E+08	1.72E+08	10136625	16442472	3.3	3	152.342	8097.98	504.699	1402.17
15	1HRL	83.9	731	731	ZOAB	21278500	17558706	1396845	2323490	2.8	13	192.76	7953.93	609.37	1107.85
15	1HRL	84.4	3046	3046	ZOAB	2.02E+08	1.65E+08	15101625	21716608	2.8	4	192.76	7953.93	609.37	1107.85
15	1HRL	87.6	5479	5479	ZOAB	3.8E+08	3.01E+08	33947789	45425871	2.6	7	192.76	7953.93	609.37	1107.85
15	1HRR	26.6	1188	1188	EAB	17372900	10779529	1164671	5429437	2.7	3	133.688	8283.49	332.666	1306.83
15	1HRR	26.7	1188	1188	EAB	17372900	10779529	1164671	5429437	3.4	2	133.688	8283.49	332.666	1306.83
15	1HRR	63.5	2675	2675	ZOAB+	89557752	71976297	6487661	11445056	3.6	5	152.342	8097.98	504.699	1402.17
15	1HRR	64.1	3466	3466	ZOAB	1.69E+08	1.32E+08	13894271	22527244	4.2	4	152.342	8097.98	504.699	1402.17
15	1HRR	80.8	4749	4749	ZOAB	3.5E+08	2.84E+08	26856515	38569254	2.9	8	192.76	7953.93	609.37	1107.85
15	1HRR	91.6	9316	9316	ZOAB	7.29E+08	5.78E+08	62443748	81332987	3.9	13	192.76	7953.93	609.37	1107.85
15	1HRR	96.8	5844	5844	ZOAB	2.67E+08	2.24E+08	15536882	29531991	3.4	9	159.597	8028.55	565.843	1459.17
15	1HRR	114.5	4019	4019	ZOAB+	1.63E+08	1.24E+08	12832331	25353440	2.9	11	159.597	8028.55	565.843	1459.17
15	1HRR	115.3	4019	4019	DAB	1.63E+08	1.24E+08	12832331	25353440	2.6	12	159.597	8028.55	565.843	1459.17

Table K.3: List of the A15 road sections which are predicted to meet the maintenance requirement concerning road rutting on January 1st, 2019 by DTC model (Part I)

ROAD	DIRECTION	FROM_KM	AGE_IRI	AGE_RUT	SURFACE_LAYER	I_AL	I_L1	I_L2	I_L3	IRI_VALUE_0	RUT_VALUE_0	T_TEMP_25	T_TEMP_0	T_TEMP_0_below	T_PERCIPITATION
15	1HRL	83.9	41	41	ZOAB	1578500	1314706	104345	159490	2.26	13	234.4972	7921.927	599.4972	1080.726
15	1HRR	78.9	4059	4059	ZOAB	1.1E+08	88354583	9602250	12126256	2.24	18	191.676	8066.704	496.5223	1356.006
15	1HRR	78.1	1289	1289	ZOAB	39394600	32697539	2543607	4153495	2.74	23	191.676	8066.704	496.5223	1356.006
15	1HRL	123.1	3410	3410	ZOAB	######	######	######	######	3.0	14	126.361	8197.39	434.239	1817.19
15	1HRR	151.6	762	762	ZOEAB	17335000	14117417	1299010	1918604	2.96	19	196.7737	8000.433	556.676	1415.14
15	1HRR	78.9	3709	3709	ZOAB	2.21E+08	1.83E+08	14474257	22647723	1.7	16	136.389	8288.65	333.953	1867.33
15	1HRR	79	4059	4059	ZOAB	1.85E+08	1.55E+08	13502016	15686076	1.47	15	191.676	8066.704	496.5223	1356.006
15	1HRR	82.5	57	57	ZOAB	2285700	1888524	152532	244644	1.2	13	164.469	8129.2	466.331	1129.22
15	1HRR	78.1	939	939	ZOAB	50165600	41679880	3316697	5169023	1.8	15	136.389	8288.65	333.953	1867.33
15	1HRR	114.5	3329	3329	ZOAB+	55438460	41836850	4309331	9292320	2.99	14	196.7737	8000.433	556.676	1415.14
15	1HRR	79.1	4059	4059	ZOAB	1.85E+08	1.55E+08	13502016	15686076	1.3	14	191.676	8066.704	496.5223	1356.006
15	1HRL	116.1	3177	3177	ZOAB+	55130445	41518153	4208520	9403813	2.02	16	196.7737	8000.433	556.676	1415.14
15	1HRR	116.2	772	772	ZOAB+	12486300	9622653	888425	1975263	1.9	14	196.7737	8000.433	556.676	1415.14
15	1HRR	79.0	3709	3709	ZOAB	3.38E+08	2.85E+08	20777303	31598852	1.1	14	136.389	8288.65	333.953	1867.33
15	1HRL	126.2	31	31	ZOEAB	1199700	1002726	79267	117738	1.55	15	196.7737	8000.433	556.676	1415.14
15	1HRL	125.8	31	31	ZOEAB	1199700	1002726	79267	117738	2.35	14	196.7737	8000.433	556.676	1415.14
15	1HRR	79.1	3709	3709	ZOAB	3.38E+08	2.85E+08	20777303	31598852	0.9	12	136.389	8288.65	333.953	1867.33
15	1HRR	52	926	926	ZOAB+	37341600	31786148	1997370	3558155	1.19	15	191.676	8066.704	496.5223	1356.006
15	1HRR	116.5	3046	3046	ZOAB+	######	######	######	######	2.0	12	126.361	8197.39	434.239	1817.19
15	1HRR	82.3	3329	3329	ZOAB+	1.32E+08	1.07E+08	11353158	13873560	1.42	12	234.4972	7921.927	599.4972	1080.726
15	1HRL	116.1	2827	2827	ZOAB+	89355000	70314922	5592955	13447123	1.4	12	126.361	8197.39	434.239	1817.19
15	1HRR	79.2	4059	4059	ZOAB	1.85E+08	1.55E+08	13502016	15686076	1.37	13	191.676	8066.704	496.5223	1356.006
15	1HRR	114.5	2979	2979	ZOAB+	92245000	71132683	6561647	14550670	1.7	13	126.361	8197.39	434.239	1817.19
15	1HRL	114.2	3177	3177	ZOAB+	55130445	41518153	4208520	9403813	1.63	12	196.7737	8000.433	556.676	1415.14
15	1HRL	83.9	1886	1886	ZOAB	1.38E+08	1.14E+08	9374515	14208378	1.9	21	164.469	8129.2	466.331	1129.22
15	1HRR	112.1	3329	3329	ZOAB+	86730541	68188635	6307021	12145805	1.62	13	196.7737	8000.433	556.676	1415.14
15	1HRL	156.8	5100	5100	OAB	######	######	######	######	1.1	10	126.361	8197.39	434.239	1817.19
15	1HRL	126.1	31	31	ZOEAB	1199700	1002726	79267	117738	1.54	13	196.7737	8000.433	556.676	1415.14
15	1HRR	96.8	5154	5154	ZOAB	1.46E+08	1.2E+08	10187693	17160939	3.38	13	196.7737	8000.433	556.676	1415.14
15	1HRR	159.9	5810	5810	SMA	######	######	######	######	2.2	11	126.361	8197.39	434.239	1817.19
15	1HRL	84.2	1289	1289	ZOAB+	49876800	41498150	3354981	5023344	2.18	13	234.4972	7921.927	599.4972	1080.726
15	1HRR	110.2	2979	2979	ZOAB+	1.51E+08	1.2E+08	11072686	20532991	1.2	11	126.361	8197.39	434.239	1817.19
15	1HRL	118.9	31	31	ZOEAB	527000	406782	36022	84227	2.61	15	196.7737	8000.433	556.676	1415.14
15	1HRR	82.4	3329	3329	ZOAB+	1.32E+08	1.07E+08	11353158	13873560	1.55	11	234.4972	7921.927	599.4972	1080.726
15	1HRR	110.2	3329	3329	ZOAB+	86730541	68188635	6307021	12145805	1.86	12	196.7737	8000.433	556.676	1415.14

Table K.4: List of the A15 road sections which are predicted to meet the maintenance requirement concerning road rutting on January 1st, 2019 by DTC model (Part II)

ROAD	DIREC TION	FROM _KM	AGE _IRI	AGE _RUT	SURFACE _LAYER	I_AL	I_L1	I_L2	I_L3	IRI_ VALUE _0	RUT_ VALUE _0	T_ TEMP _25	T_ TEMP _0	T_ TEMP _0_ below	T_ PERCI PITA TION
15	1HRL	156.5	5100	5100	OAB	######	######	######	######	1.5	11	126.361	8197.39	434.239	1817.19
15	1HRL	116	3177	3177	ZOAB+	55130445	41518153	4208520	9403813	1.16	13	196.7737	8000.433	556.676	1415.14
15	1HRR	115.9	3329	3329	ZOAB+	55438460	41836850	4309331	9292320	1.33	12	196.7737	8000.433	556.676	1415.14
15	1HRR	117.5	3046	3046	ZOAB+	######	######	######	######	1.1	12	126.361	8197.39	434.239	1817.19
15	1HRR	79.2	3709	3709	ZOAB	3.38E+08	2.85E+08	20777303	31598852	0.9	11	136.389	8288.65	333.953	1867.33
15	1HRR	78.8	1289	1289	ZOAB	39394600	32697539	2543607	4153495	1.73	15	191.676	8066.704	496.5223	1356.006
15	1HRR	52.3	926	926	ZOAB+	37341600	31786148	1997370	3558155	0.91	14	191.676	8066.704	496.5223	1356.006
15	1HRR	71.9	3329	3329	ZOAB+	1.3E+08	1.11E+08	8728394	10903779	1.63	12	191.676	8066.704	496.5223	1356.006
15	1HRR	116.1	3329	3329	ZOAB+	55438460	41836850	4309331	9292320	1.72	12	196.7737	8000.433	556.676	1415.14
15	1HRR	71.5	3693	3693	ZOAB	1.45E+08	1.23E+08	9840711	12255952	1.87	13	191.676	8066.704	496.5223	1356.006
15	1HRR	111.6	2979	2979	ZOAB+	1.51E+08	1.2E+08	11072686	20532991	0.9	10	126.361	8197.39	434.239	1817.19
15	1HRR	110.3	2979	2979	ZOAB+	1.51E+08	1.2E+08	11072686	20532991	1.3	10	126.361	8197.39	434.239	1817.19
15	1HRR	111.8	3329	3329	ZOAB+	86730541	68188635	6307021	12145805	1.84	12	196.7737	8000.433	556.676	1415.14
15	1HRR	113	3329	3329	ZOAB+	86673141	68146938	6301486	12135596	1.5	14	196.7737	8000.433	556.676	1415.14
15	1HRR	151.9	762	762	ZOEAB	17335000	14117417	1299010	1918604	2.02	15	196.7737	8000.433	556.676	1415.14
15	1HRR	113.6	3329	3329	ZOAB+	86751041	68213071	6301650	12147240	1.31	11	196.7737	8000.433	556.676	1415.14
15	1HRR	113.0	2979	2979	ZOAB+	1.51E+08	1.2E+08	11072686	20532991	1.0	10	126.361	8197.39	434.239	1817.19
15	1HRL	83.3	4059	4059	ZOAB	1.57E+08	1.24E+08	15776238	17538491	1.18	11	234.4972	7921.927	599.4972	1080.726
15	1HRR	96.8	4804	4804	ZOAB	1.88E+08	1.62E+08	9065339	17356912	2.3	11	126.361	8197.39	434.239	1817.19
15	1HRR	111.6	3329	3329	ZOAB+	86730541	68188635	6307021	12145805	1.43	12	196.7737	8000.433	556.676	1415.14
15	1HRR	71.2	3693	3693	ZOAB	1.75E+08	1.5E+08	11171897	13867390	1.17	11	191.676	8066.704	496.5223	1356.006
15	1HRR	80.2	3709	3709	ZOAB	3.38E+08	2.85E+08	20777303	31598852	1.1	10	164.469	8129.2	466.331	1129.22
15	1HRR	112.8	3329	3329	ZOAB+	86673141	68146938	6301486	12135596	1.8	12	196.7737	8000.433	556.676	1415.14
15	1HRR	114.9	3329	3329	ZOAB+	55438460	41836850	4309331	9292320	1.48	11	196.7737	8000.433	556.676	1415.14
15	1HRR	110.3	3329	3329	ZOAB+	86730541	68188635	6307021	12145805	1.78	12	196.7737	8000.433	556.676	1415.14
15	1HRR	82.4	2979	2979	ZOAB+	2.28E+08	1.9E+08	14337364	23921500	1.3	11	164.469	8129.2	466.331	1129.22
15	1HRR	80.2	4059	4059	ZOAB	1.85E+08	1.55E+08	13502016	15686076	1.67	13	234.4972	7921.927	599.4972	1080.726
15	1HRR	112.1	2979	2979	ZOAB+	1.51E+08	1.2E+08	11072686	20532991	1.3	11	126.361	8197.39	434.239	1817.19
15	1HRR	116.3	3329	3329	ZOAB+	55438460	41836850	4309331	9292320	1.41	12	196.7737	8000.433	556.676	1415.14
15	1HRR	116.3	2979	2979	ZOAB+	92245000	71132683	6561647	14550670	1.1	10	126.361	8197.39	434.239	1817.19
15	1HRR	115.7	3329	3329	ZOAB+	55438460	41836850	4309331	9292320	1.11	12	196.7737	8000.433	556.676	1415.14
15	1HRR	116	3329	3329	ZOAB+	55438460	41836850	4309331	9292320	1.37	12	196.7737	8000.433	556.676	1415.14
15	1HRR	111.7	3329	3329	ZOAB+	86730541	68188635	6307021	12145805	1.57	12	196.7737	8000.433	556.676	1415.14
15	1HRL	114.7	3177	3177	ZOAB+	55130445	41518153	4208520	9403813	0.97	12	196.7737	8000.433	556.676	1415.14
15	1HRR	110.1	3329	3329	ZOAB+	86730541	68188635	6307021	12145805	1.73	11	196.7737	8000.433	556.676	1415.14

Table K.5: List of the A15 road sections which are predicted to meet the maintenance requirement concerning road rutting on January 1st, 2019 by DTC model (Part III)

ROAD	DIRECTION	FROM_KM	AGE_IRI	AGE_RUT	SURFACE_LAYER	I_AL	I_L1	I_L2	I_L3	IRI_VALUE_0	RUT_VALUE_0	T_TEMP_25	T_TEMP_0	T_TEMP_0_below	T_PERCIPITATION
15	1HRR	115.6	772	772	ZOAB+	12486300	9622653	888425	1975263	1.35	12	196.7737	8000.433	556.676	1415.14
15	1HRL	113	3177	3177	ZOAB+	84102830	66557405	6004539	11419190	1.44	11	196.7737	8000.433	556.676	1415.14
15	1HRR	116.2	422	422	ZOAB+	9653900	7445484	693332	1515084	1.3	12	126.361	8197.39	434.239	1817.19
15	1HRL	118	62	62	ZOEAB	1683300	1347291	120063	215946	2.93	12	196.7737	8000.433	556.676	1415.14
15	1HRL	93.6	1289	1289	ZOAB	49471000	41738951	3011831	4719934	1.17	11	234.4972	7921.927	599.4972	1080.726
15	1HRR	115.3	3329	3329	DAB	55438460	41836850	4309331	9292320	1.99	12	196.7737	8000.433	556.676	1415.14
15	1HRL	160.0	2406	2406	SMA	######	######	######	######	2.1	10	126.361	8197.39	434.239	1817.19
15	1HRL	152.8	1310	1310	ZOABTW	36454300	30242343	2753792	3458227	0.9	11	196.7737	8000.433	556.676	1415.14
15	1HRL	154.2	1310	1310	ZOABTW	15046000	12469867	1310379	1265785	0.94	11	196.7737	8000.433	556.676	1415.14
15	1HRL	114.4	3177	3177	ZOAB+	55130445	41518153	4208520	9403813	1.47	11	196.7737	8000.433	556.676	1415.14
15	1HRL	149.5	62	62	ZOEAB	1661600	1350670	136307	174654	2.27	12	196.7737	8000.433	556.676	1415.14
15	1HRR	109.8	3329	3329	ZOAB+	86730541	68188635	6307021	12145805	1.46	11	196.7737	8000.433	556.676	1415.14
15	1HRR	109.5	3329	3329	ZOAB+	86730541	68188635	6307021	12145805	1.68	12	196.7737	8000.433	556.676	1415.14
15	1HRR	115.7	2979	2979	ZOAB+	92245000	71132683	6561647	14550670	0.8	10	126.361	8197.39	434.239	1817.19
15	1HRR	131.2	62	62	ZOEAB	2157600	1812322	150040	195269	1.5	12	196.7737	8000.433	556.676	1415.14
15	1HRR	117.6	3046	3046	ZOAB+	######	######	######	######	1.0	11	126.361	8197.39	434.239	1817.19
15	1HRL	96.9	894	894	ZOAB	27803500	22726705	1955366	3121470	1.89	14	196.7737	8000.433	556.676	1415.14
15	1HRL	149.2	62	62	ZOEAB	1661600	1350670	136307	174654	1.48	12	196.7737	8000.433	556.676	1415.14
15	1HRR	151.7	762	762	ZOEAB	17335000	14117417	1299010	1918604	2.21	16	196.7737	8000.433	556.676	1415.14
15	1HRR	82.9	1289	1289	ZOAB	50797200	42238664	3170248	5387963	1.48	12	234.4972	7921.927	599.4972	1080.726
15	1HRR	111.9	3329	3329	ZOAB+	86730541	68188635	6307021	12145805	2.69	12	196.7737	8000.433	556.676	1415.14
15	1HRR	80.5	4059	4059	ZOAB	1.53E+08	1.23E+08	14067146	16336898	1.39	11	234.4972	7921.927	599.4972	1080.726
15	1HRR	114.8	3329	3329	ZOAB+	55438460	41836850	4309331	9292320	1.53	11	196.7737	8000.433	556.676	1415.14
15	1HRL	153.6	1310	1310	ZOABTW	26111400	21885875	2054519	2170671	1.15	11	196.7737	8000.433	556.676	1415.14
15	1HRL	154.3	1310	1310	ZOABTW	15046000	12469867	1310379	1265785	1.13	11	196.7737	8000.433	556.676	1415.14
15	1HRR	97.4	3524	3524	ZOAB	1.15E+08	93523741	7711176	13303247	1.91	12	196.7737	8000.433	556.676	1415.14
15	1HRR	84.7	57	57	ZOEAB	2080500	1680873	153216	246411	1.3	10	164.469	8129.2	466.331	1129.22
15	1HRR	114.6	772	772	ZOAB+	12486300	9622653	888425	1975263	2.22	12	196.7737	8000.433	556.676	1415.14
15	1HRL	114.8	3177	3177	ZOAB+	55130445	41518153	4208520	9403813	1.14	12	196.7737	8000.433	556.676	1415.14
15	1HRR	80.1	4059	4059	ZOAB	1.85E+08	1.55E+08	13502016	15686076	1.78	13	234.4972	7921.927	599.4972	1080.726
15	1HRR	115.8	772	772	ZOAB+	12486300	9622653	888425	1975263	1.3	11	196.7737	8000.433	556.676	1415.14
15	1HRR	157.0	5479	5479	SMA	######	######	######	######	1.6	10	126.361	8197.39	434.239	1817.19
15	1HRL	83.3	3712	3712	ZOAB	2.8E+08	2.32E+08	19096139	28922286	1.3	10	164.469	8129.2	466.331	1129.22
15	1HRR	71.5	3343	3343	ZOAB	2.54E+08	2.24E+08	13711978	15514337	1.4	10	136.389	8288.65	333.953	1867.33
15	1HRL	131.3	62	62	ZOEAB	2021200	1665940	145824	209436	1.4	12	196.7737	8000.433	556.676	1415.14

Table K.6: List of the A15 road sections which are predicted to meet the maintenance requirement concerning road rutting on January 1st, 2019 by DTC model (Part IV)

ROAD	DIREC TION	FROM _KM	AGE _IRI	AGE _RUT	SURFACE _LAYER	I_AL	I_L1	I_L2	I_L3	IRI_ VALUE _0	RUT_ VALUE _0	T_ TEMP _25	T_ TEMP _0	T_ TEMP _0_ below	T_ PERCI PITA TION
15	1HRR	110.5	3329	3329	ZOAB+	86730541	68188635	6307021	12145805	1.48	11	196.7737	8000.433	556.676	1415.14
15	1HRL	153.5	1310	1310	ZOABTW	26111400	21885875	2054519	2170671	1.13	11	196.7737	8000.433	556.676	1415.14
15	1HRR	83.4	3329	3329	ZOAB+	1.32E+08	1.07E+08	11353158	13873560	1.64	11	234.4972	7921.927	599.4972	1080.726
15	1HRR	115.9	2979	2979	ZOAB+	92245000	71132683	6561647	14550670	1.0	11	126.361	8197.39	434.239	1817.19
15	1HRR	51.9	926	926	ZOAB+	37341600	31786148	1997370	3558155	1.08	13	191.676	8066.704	496.5223	1356.006
15	1HRR	80.8	4059	4059	ZOAB	1.61E+08	1.29E+08	14766347	17122308	2.69	11	234.4972	7921.927	599.4972	1080.726
15	1HRR	110.4	3329	3329	ZOAB+	86730541	68188635	6307021	12145805	1.45	11	196.7737	8000.433	556.676	1415.14
15	1HRR	52.5	926	926	ZOAB+	37341600	31786148	1997370	3558155	1.1	12	191.676	8066.704	496.5223	1356.006
15	1HRR	80.4	4059	4059	ZOAB	1.53E+08	1.23E+08	14067146	16336898	1.8	11	234.4972	7921.927	599.4972	1080.726
15	1HRR	114.3	3329	3329	ZOAB+	55438460	41836850	4309331	9292320	1.77	11	196.7737	8000.433	556.676	1415.14
15	1HRR	151.5	762	762	ZOEAB	17335000	14117417	1299010	1918604	3.3	15	196.7737	8000.433	556.676	1415.14
15	1HRR	71.7	3693	3693	ZOAB	1.45E+08	1.23E+08	9840711	12255952	1.51	12	191.676	8066.704	496.5223	1356.006
15	1HRR	82.3	2979	2979	ZOAB+	2.28E+08	1.9E+08	14337364	23921500	1.0	12	164.469	8129.2	466.331	1129.22
15	1HRR	52.0	565	565	ZOAB+	36181500	31479617	1608107	3093776	1.0	12	136.389	8288.65	333.953	1867.33
15	1HRR	78.2	1289	1289	ZOAB	39394600	32697539	2543607	4153495	1.44	11	191.676	8066.704	496.5223	1356.006
15	1HRR	111.1	3329	3329	ZOAB+	86730541	68188635	6307021	12145805	1.63	11	196.7737	8000.433	556.676	1415.14
15	1HRL	149.3	62	62	ZOEAB	1661600	1350670	136307	174654	1.77	12	196.7737	8000.433	556.676	1415.14
15	1HRR	114.2	3329	3329	ZOAB+	55438460	41836850	4309331	9292320	3.2	16	196.7737	8000.433	556.676	1415.14
15	1HRR	82.2	3329	3329	ZOAB+	1.32E+08	1.07E+08	11353158	13873560	1.69	11	234.4972	7921.927	599.4972	1080.726
15	1HRR	78.8	939	939	ZOAB	50165600	41679880	3316697	5169023	1.1	11	136.389	8288.65	333.953	1867.33
15	1HRL	131.1	62	62	ZOEAB	2021200	1634909	152737	233554	0.94	12	196.7737	8000.433	556.676	1415.14
15	1HRR	71.4	3693	3693	ZOAB	1.75E+08	1.5E+08	11171897	13867390	1.61	11	191.676	8066.704	496.5223	1356.006
15	1HRL	153.3	1310	1310	ZOABTW	26111400	21885875	2054519	2170671	1.09	11	196.7737	8000.433	556.676	1415.14
15	1HRR	153.4	1340	1340	ZOABTW	35444100	28833022	2660122	3950621	1.45	11	196.7737	8000.433	556.676	1415.14
15	1HRR	52.1	926	926	ZOAB+	37341600	31786148	1997370	3558155	1.17	17	191.676	8066.704	496.5223	1356.006
15	1HRR	112.9	3329	3329	ZOAB+	86673141	68146938	6301486	12135596	1.93	11	196.7737	8000.433	556.676	1415.14
15	1HRR	109.6	3329	3329	ZOAB+	86730541	68188635	6307021	12145805	1.69	13	196.7737	8000.433	556.676	1415.14
15	1HRR	111.8	2979	2979	ZOAB+	1.51E+08	1.2E+08	11072686	20532991	1.4	10	126.361	8197.39	434.239	1817.19
15	1HRR	152.7	4413	4413	ZOEAB	1.18E+08	96013636	8700354	13485096	1.44	13	196.7737	8000.433	556.676	1415.14
15	1HRR	71.2	3343	3343	ZOAB	3.05E+08	2.74E+08	14418415	16500369	0.8	10	136.389	8288.65	333.953	1867.33
15	1HRL	88.7	3694	3694	ZOAB	1.43E+08	1.12E+08	16243227	15066140	2.1	11	234.4972	7921.927	599.4972	1080.726
15	1HRL	131.2	62	62	ZOEAB	2021200	1665940	145824	209436	1.33	12	196.7737	8000.433	556.676	1415.14
15	1HRR	151.8	762	762	ZOEAB	17335000	14117417	1299010	1918604	2.09	13	196.7737	8000.433	556.676	1415.14
15	1HRR	131.3	62	62	ZOEAB	2157600	1812322	150040	195269	1.78	11	196.7737	8000.433	556.676	1415.14
15	1HRR	77.8	1289	1289	ZOAB	53403400	44833185	3664628	4905262	1.05	14	191.676	8066.704	496.5223	1356.006

Table K.7: List of the A15 road sections which are predicted to meet the maintenance requirement concerning road rutting on January 1st, 2019 by DTC model (Part V)

ROAD	DIRECTION	FROM_KM	AGE_IRI	AGE_RUT	SURFACE_LAYER	I_AL	I_L1	I_L2	I_L3	IRI_VALUE_0	RUT_VALUE_0	T_TEMP_25	T_TEMP_0	T_TEMP_0_below	T_PERCIPITATION
15	1HRL	105.3	4150	4150	ZOABTW	1.04E+08	78571878	8497841	16879193	1.59	11	196.7737	8000.433	556.676	1415.14
15	1HRL	116.3	3177	3177	ZOAB+	1.07E+08	86866242	7628141	12952306	1.54	12	196.7737	8000.433	556.676	1415.14
15	1HRR	113.6	2979	2979	ZOAB+	1.51E+08	1.2E+08	11072686	20532991	1.0	10	126.361	8197.39	434.239	1817.19
15	1HRR	153.5	1340	1340	ZOABTW	35444100	28833022	2660122	3950621	1.16	11	196.7737	8000.433	556.676	1415.14
15	1HRL	126.3	31	31	ZOEAB	1199700	1002726	79267	117738	1.43	13.00539	196.7737	8000.433	556.676	1415.14
15	1HRR	71.9	2979	2979	ZOAB+	2.25E+08	1.99E+08	12166434	13767501	1.1	10	136.389	8288.65	333.953	1867.33
15	1HRR	153.3	1340	1340	ZOABTW	35444100	28833022	2660122	3950621	1.41	11	196.7737	8000.433	556.676	1415.14
15	1HRL	91.6	1289	1289	ZOAB	49471000	41738951	3011831	4719934	2.64	11	234.4972	7921.927	599.4972	1080.726
15	1HRR	151.4	762	762	ZOEAB	17335000	14117417	1299010	1918604	1.74	13	196.7737	8000.433	556.676	1415.14
15	1HRR	91.6	8626	8626	ZOAB	3.43E+08	2.58E+08	42520343	36244132	2.09	11	234.4972	7921.927	599.4972	1080.726
15	1HRL	83.1	4059	4059	ZOAB	1.57E+08	1.24E+08	15776238	17538491	0.98	11	234.4972	7921.927	599.4972	1080.726
15	1HRR	152.6	4413	4413	ZOEAB	1.18E+08	96013636	8700354	13485096	1.4	12	196.7737	8000.433	556.676	1415.14
15	1HRL	114.2	2827	2827	ZOAB+	89355000	70314922	5592955	13447123	1.2	11	126.361	8197.39	434.239	1817.19
15	1HRR	114.9	2979	2979	ZOAB+	92245000	71132683	6561647	14550670	1.0	10	126.361	8197.39	434.239	1817.19
15	1HRL	71.5	3693	3693	ZOAB	1.6E+08	1.38E+08	10043948	12589227	1.5	12	191.676	8066.704	496.5223	1356.006
15	1HRL	126	31	31	ZOEAB	1199700	1002726	79267	117738	2.4	13	196.7737	8000.433	556.676	1415.14
15	1HRR	72.2	3329	3329	ZOAB+	1.3E+08	1.11E+08	8728394	10903779	1.8	11	191.676	8066.704	496.5223	1356.006
15	1HRR	111.2	3329	3329	ZOAB+	86730541	68188635	6307021	12145805	1.61	11	196.7737	8000.433	556.676	1415.14
15	1HRR	153.1	1340	1340	ZOABTW	35444100	28833022	2660122	3950621	0.95	11	196.7737	8000.433	556.676	1415.14
15	1HRL	82.5	4059	4059	ZOAB	1.57E+08	1.24E+08	15776238	17538491	1.06	14	234.4972	7921.927	599.4972	1080.726
15	1HRL	95.5	3329	3329	ZOAB	66363868	51113648	5940590	9309671	1.38	11	196.7737	8000.433	556.676	1415.14
15	1HRR	52.2	926	926	ZOAB+	37341600	31786148	1997370	3558155	1.19	13	191.676	8066.704	496.5223	1356.006
15	1HRR	153.7	1340	1340	ZOABTW	25129200	20065278	2012129	3051855	0.99	11	196.7737	8000.433	556.676	1415.14
15	1HRL	115.3	3177	3177	DAB	55130445	41518153	4208520	9403813	1.81	11	196.7737	8000.433	556.676	1415.14
15	1HRR	77.7	1289	1289	ZOAB	53403400	44833185	3664628	4905262	1.13	12	191.676	8066.704	496.5223	1356.006
15	1HRL	84	1289	1289	ZOAB+	49876800	41498150	3354981	5023344	1.62	12	234.4972	7921.927	599.4972	1080.726
15	1HRR	77.9	1289	1289	ZOAB	39394600	32697539	2543607	4153495	1.36	15	191.676	8066.704	496.5223	1356.006
15	1HRL	93.2	407	407	ZOAB	15775500	13350743	961756	1462717	1.26	12	234.4972	7921.927	599.4972	1080.726
15	1HRL	93.3	407	407	ZOAB	15775500	13350743	961756	1462717	1.38	14	234.4972	7921.927	599.4972	1080.726
15	1HRL	93.1	407	407	ZOAB	15775500	13350743	961756	1462717	1.15	12	234.4972	7921.927	599.4972	1080.726
15	1HRR	152	762	762	ZOEAB	17335000	14117417	1299010	1918604	2.64	12	196.7737	8000.433	556.676	1415.14
15	1HRR	78	1289	1289	ZOAB	39394600	32697539	2543607	4153495	1.17	11	191.676	8066.704	496.5223	1356.006
15	1HRL	84.2	942	942	ZOAB+	64141800	53191515	4348659	6601626	2.2	11	164.469	8129.2	466.331	1129.22
15	1HRL	141.2	31	31	ZOEAB	899000	752835	56141	90086	1.44	13	196.7737	8000.433	556.676	1415.14
15	1HRL	149.1	62	62	ZOEAB	1661600	1350670	136307	174654	1.88	11	196.7737	8000.433	556.676	1415.14

Table K.8: List of the A15 road sections which are predicted to meet the maintenance requirement concerning road rutting on January 1st, 2019 by DTC model (Part VI)

ROAD	DIRECTION	FROM_KM	AGE_IRI	AGE_RUT	SURFACE_LAYER	I_AL	I_L1	I_L2	I_L3	IRI_VALUE_0	RUT_VALUE_0	T_TEMP_25	T_TEMP_0	T_TEMP_0_below	T_PERCIPITATION
15	1HRL	83	4059	4059	ZOAB	1.57E+08	1.24E+08	15776238	17538491	2.28	11	234.4972	7921.927	599.4972	1080.726
15	1HRL	146.1	31	31	ZOEAB	880400	740373	56761	83297	1.21	12	196.7737	8000.433	556.676	1415.14
15	1HRR	77.6	1289	1289	ZOAB	53403400	44833185	3664628	4905262	1.11	12	191.676	8066.704	496.5223	1356.006
15	1HRR	112.7	3329	3329	ZOAB+	86673141	68146938	6301486	12135596	2.2	12	196.7737	8000.433	556.676	1415.14
15	1HRR	112.2	3329	3329	ZOAB+	86673141	68146938	6301486	12135596	2.12	12	196.7737	8000.433	556.676	1415.14
15	1HRL	71.4	3693	3693	ZOAB	1.74E+08	1.51E+08	10360479	12914778	1.6	11	191.676	8066.704	496.5223	1356.006
15	1HRR	152.8	1340	1340	ZOABTW	35444100	28833022	2660122	3950621	0.82	11	196.7737	8000.433	556.676	1415.14
15	1HRR	131.4	62	62	ZOEAB	2157600	1812322	150040	195269	1.62	11	196.7737	8000.433	556.676	1415.14
15	1HRL	117.9	62	62	ZOEAB	1683300	1347291	120063	215946	1.13	15	196.7737	8000.433	556.676	1415.14
15	1HRL	153.9	1310	1310	ZOABTW	26111400	21885875	2054519	2170671	1.04	11	196.7737	8000.433	556.676	1415.14
15	1HRL	149	62	62	ZOEAB	1661600	1350670	136307	174654	1.31	12	196.7737	8000.433	556.676	1415.14
15	1HRR	81.2	4059	4059	ZOAB	1.61E+08	1.29E+08	14766347	17122308	1.25	12	234.4972	7921.927	599.4972	1080.726
15	1HRL	153.4	1310	1310	ZOABTW	26111400	21885875	2054519	2170671	1.28	11	196.7737	8000.433	556.676	1415.14
15	1HRR	153.8	1340	1340	ZOABTW	25129200	20065278	2012129	3051855	1.01	11	196.7737	8000.433	556.676	1415.14
15	1HRL	82.6	4059	4059	ZOAB	1.57E+08	1.24E+08	15776238	17538491	1.09	13	234.4972	7921.927	599.4972	1080.726
15	1HRL	145.4	397	397	ZOEAB	10762400	8986719	754723	1020623	0.96	13	196.7737	8000.433	556.676	1415.14
15	1HRR	98	772	772	ZOAB+	25115400	20941065	1496965	2677411	1.7	11	196.7737	8000.433	556.676	1415.14
15	1HRR	152.9	1340	1340	ZOABTW	35444100	28833022	2660122	3950621	0.89	11	196.7737	8000.433	556.676	1415.14
15	1HRR	153.6	1340	1340	ZOABTW	35444100	28833022	2660122	3950621	0.96	11	196.7737	8000.433	556.676	1415.14
15	1HRL	131.4	62	62	ZOEAB	2021200	1665940	145824	209436	0.84	12	196.7737	8000.433	556.676	1415.14
15	1HRL	93	407	407	ZOAB	15775500	13350743	961756	1462717	1.54	12	234.4972	7921.927	599.4972	1080.726
15	1HRR	153.2	1340	1340	ZOABTW	35444100	28833022	2660122	3950621	1.07	11	196.7737	8000.433	556.676	1415.14
15	1HRL	152.9	1310	1310	ZOABTW	36454300	30242343	2753792	3458227	0.88	11	196.7737	8000.433	556.676	1415.14
15	1HRL	97.7	1533	1533	ZOAB	47360200	38538220	3385029	5436992	1.26	12	196.7737	8000.433	556.676	1415.14
15	1HRL	131	62	62	ZOEAB	2021200	1634909	152737	233554	2.04	12	196.7737	8000.433	556.676	1415.14
15	1HRR	66.1	2265	2265	ZOAB+	94820315	82914773	5257756	6647859	1.05	11	191.676	8066.704	496.5223	1356.006
15	1HRR	151.3	762	762	ZOEAB	19458900	15845687	1211193	2401685	2.04	13	196.7737	8000.433	556.676	1415.14
15	1HRL	125.9	31	31	ZOEAB	1199700	1002726	79267	117738	1.49	14	196.7737	8000.433	556.676	1415.14
15	1HRL	153.7	1310	1310	ZOABTW	26111400	21885875	2054519	2170671	1.43	11	196.7737	8000.433	556.676	1415.14
15	1HRR	109.3	3329	3329	ZOAB+	86730541	68188635	6307021	12145805	1.65	11	196.7737	8000.433	556.676	1415.14
15	1HRR	153	1340	1340	ZOABTW	35444100	28833022	2660122	3950621	1.35	11	196.7737	8000.433	556.676	1415.14
15	1HRL	149.4	62	62	ZOEAB	1661600	1350670	136307	174654	1.48	12	196.7737	8000.433	556.676	1415.14
15	1HRR	59.4	926	926	ZOAB+	40082400	34986362	1873091	3222654	1.98	11	191.676	8066.704	496.5223	1356.006
15	1HRR	153.9	1340	1340	ZOABTW	25129200	20065278	2012129	3051855	1.07	11	196.7737	8000.433	556.676	1415.14
15	1HRL	106.4	894	894	ZOAB	23747500	18594806	1935916	3216819	1.43	11	196.7737	8000.433	556.676	1415.14

Table K.9: List of the A15 road sections which are predicted to meet the maintenance requirement concerning road rutting on January 1st, 2019 by DTC model (Part VII)

ROAD	DIRECTION	FROM_KM	AGE_IRI	AGE_RUT	SURFACE_LAYER	I_AL	I_L1	I_L2	I_L3	IRI_VALUE_0	RUT_VALUE_0	T_TEMP_25	T_TEMP_0	T_TEMP_0_below	T_PERCIPITATION
15	1HRL	151.7	1290	1290	ZOEAB	27488100	22528623	2245759	2713911	2.2	11	159.597	8028.55	565.843	1459.17
15	1HRL	160.0	3207	3207	SMA	70658221	61748577	4579151	4330701	2.2	14	159.597	8028.55	565.843	1459.17
15	1HRR	161.0	2142	2142	EAB	72932833	63855558	4412145	4602533	2.4	14	159.597	8028.55	565.843	1459.17
15	1HRR	116.5	3847	3847	ZOAB+	90782560	70679146	6867380	13236624	2.2	12	159.597	8028.55	565.843	1459.17
15	1HRR	126.8	4211	4211	ZOAB	1.25E+08	1.01E+08	9383652	15546572	1.0	11	159.597	8028.55	565.843	1459.17
15	1HRL	116.6	3695	3695	ZOAB+	1.17E+08	93827646	8232697	14456864	1.4	12	159.597	8028.55	565.843	1459.17
15	1HRL	116.7	3695	3695	ZOAB+	1.17E+08	93827646	8232697	14456864	1.5	15	159.597	8028.55	565.843	1459.17
15	1HRL	116.8	3695	3695	ZOAB+	1.17E+08	93827646	8232697	14456864	1.2	11	159.597	8028.55	565.843	1459.17
15	1HRR	116.7	3847	3847	ZOAB+	1.25E+08	1E+08	8231809	15924452	1.4	11	159.597	8028.55	565.843	1459.17
15	1HRR	117.4	3847	3847	ZOAB+	1.25E+08	1E+08	8231809	15924452	1.4	12	159.597	8028.55	565.843	1459.17
15	1HRR	117.5	3847	3847	ZOAB+	1.25E+08	1E+08	8231809	15924452	1.3	13	159.597	8028.55	565.843	1459.17
15	1HRR	117.6	3847	3847	ZOAB+	1.25E+08	1E+08	8231809	15924452	1.4	11	159.597	8028.55	565.843	1459.17
15	1HRR	117.7	3847	3847	ZOAB+	1.25E+08	1E+08	8231809	15924452	2.8	12	159.597	8028.55	565.843	1459.17
15	1HRR	159.8	6280	6280	SMA	1.26E+08	1.08E+08	8764348	9172085	2.7	11	159.597	8028.55	565.843	1459.17
15	1HRR	127.5	4211	4211	ZOAB	1.37E+08	1.11E+08	9738428	15897019	1.1	11	159.597	8028.55	565.843	1459.17
15	1HRR	127.1	4211	4211	ZOAB	1.37E+08	1.11E+08	9775778	15948964	1.2	11	159.597	8028.55	565.843	1459.17
15	1HRR	159.9	6611	6611	SMA	1.33E+08	1.14E+08	9272313	9702135	2.8	15	159.597	8028.55	565.843	1459.17
15	1HRR	126.7	4211	4211	ZOAB	1.51E+08	1.25E+08	9970163	16228663	1.2	11	159.597	8028.55	565.843	1459.17
15	1HRR	160.1	6611	6611	SMA	1.51E+08	1.32E+08	9185926	9553352	3.8	14	159.597	8028.55	565.843	1459.17
15	1HRR	160.2	6611	6611	SMA	1.52E+08	1.33E+08	9208381	9575777	2.9	12	159.597	8028.55	565.843	1459.17
15	1HRR	159.4	6611	6611	SMA	1.62E+08	1.39E+08	11080640	11777851	2.5	13	159.597	8028.55	565.843	1459.17
15	1HRR	159.3	6611	6611	SMA	1.74E+08	1.5E+08	11767547	12234373	2.6	16	159.597	8028.55	565.843	1459.17
15	1HRL	82.5	4518	4518	ZOAB	3.35E+08	2.7E+08	27455268	37092395	1.1	12	192.76	7953.93	609.37	1107.85
15	1HRL	82.6	4518	4518	ZOAB	3.35E+08	2.7E+08	27455268	37092395	1.1	11	192.76	7953.93	609.37	1107.85
15	1HRL	83.9	500	500	ZOAB	21278500	17558706	1396845	2323490	2.8	13	192.76	7953.93	609.37	1107.85
15	1HRL	84.0	1748	1748	ZOAB+	1.19E+08	98287174	7873561	12588688	1.6	12	192.76	7953.93	609.37	1107.85
15	1HRL	97.7	1992	1992	ZOAB	89789800	73230892	5992557	10568384	1.2	12	159.597	8028.55	565.843	1459.17
15	1HRL	99.4	2814	2814	ZOAB	1.55E+08	1.24E+08	11551199	19539524	2.1	12	159.597	8028.55	565.843	1459.17
15	1HRL	114.2	3636	3636	ZOAB+	1.59E+08	1.22E+08	12905832	24304141	1.3	11	159.597	8028.55	565.843	1459.17
15	1HRL	114.4	3636	3636	ZOAB+	1.59E+08	1.22E+08	12905832	24304141	1.6	11	159.597	8028.55	565.843	1459.17
15	1HRL	115.3	3636	3636	DAB	1.59E+08	1.22E+08	12905832	24304141	2.1	11	159.597	8028.55	565.843	1459.17
15	1HRL	116.3	3636	3636	ZOAB+	2.12E+08	1.68E+08	16325453	27852634	1.3	11	159.597	8028.55	565.843	1459.17
15	1HRR	71.2	4179	4179	ZOAB	3.89E+08	3.39E+08	23123837	26341705	1.3	12	152.342	8097.98	504.699	1402.17
15	1HRR	71.4	4179	4179	ZOAB	3.89E+08	3.39E+08	23123837	26341705	1.5	11	152.342	8097.98	504.699	1402.17
15	1HRR	71.5	4179	4179	ZOAB	3.21E+08	2.76E+08	21245202	24287293	1.9	14	152.342	8097.98	504.699	1402.17

Table K.10: List of the A15 road sections which are predicted to meet the maintenance requirement concerning road rutting on January 1st, 2019 by DTC model (Part VIII)

ROAD	DIREC TION	FROM _KM	AGE _IRI	AGE _RUT	SURFACE _LAYER	I_AL	I_L1	I_L2	I_L3	IRI_ VALUE _0	RUT_ VALUE _0	T_ TEMP _25	T_ TEMP _0	T_ TEMP _0_ below	T_ PERCI PITA TION
15	1HRR	71.9	3815	3815	ZOAB+	2.92E+08	2.51E+08	19139529	21887164	1.6	13	152.342	8097.98	504.699	1402.17
15	1HRR	75.8	1775	1775	ZOAB	1.07E+08	88742316	7386943	10960091	1.6	11	152.342	8097.98	504.699	1402.17
15	1HRR	77.6	1775	1775	ZOAB	1.32E+08	1.11E+08	8650603	12129512	1.2	11	152.342	8097.98	504.699	1402.17
15	1HRR	77.7	1775	1775	ZOAB	1.32E+08	1.11E+08	8650603	12129512	1.2	12	152.342	8097.98	504.699	1402.17
15	1HRR	77.8	1775	1775	ZOAB	1.32E+08	1.11E+08	8650603	12129512	1.1	15	152.342	8097.98	504.699	1402.17
15	1HRR	77.9	1775	1775	ZOAB	97082100	80745014	5839782	10499120	1.4	15	152.342	8097.98	504.699	1402.17
15	1HRR	78.0	1775	1775	ZOAB	97082100	80745014	5839782	10499120	1.2	11	152.342	8097.98	504.699	1402.17
15	1HRR	78.1	1775	1775	ZOAB	97082100	80745014	5839782	10499120	1.8	11	152.342	8097.98	504.699	1402.17
15	1HRR	80.1	4518	4518	ZOAB	4.04E+08	3.39E+08	26527410	37137540	1.7	12	192.76	7953.93	609.37	1107.85
15	1HRR	91.6	9085	9085	ZOAB	7.29E+08	5.78E+08	62443748	81332987	3.9	13	192.76	7953.93	609.37	1107.85
15	1HRR	97.4	3983	3983	ZOAB	2.54E+08	2.07E+08	16684875	30151337	1.6	11	159.597	8028.55	565.843	1459.17
15	1HRR	112.8	3788	3788	ZOAB+	1.94E+08	1.51E+08	14824486	28196716	2.1	11	159.597	8028.55	565.843	1459.17
15	1HRR	114.5	3788	3788	ZOAB+	1.63E+08	1.24E+08	12832331	25353440	2.9	11	159.597	8028.55	565.843	1459.17
15	1HRR	115.3	3788	3788	DAB	1.63E+08	1.24E+08	12832331	25353440	2.6	12	159.597	8028.55	565.843	1459.17
15	1HRR	116.1	3788	3788	ZOAB+	1.63E+08	1.24E+08	12832331	25353440	2.1	11	159.597	8028.55	565.843	1459.17
15	1HRR	116.2	1231	1231	ZOAB+	47323600	36471994	3658175	7194703	2.1	13	159.597	8028.55	565.843	1459.17
15	1HRR	116.3	3788	3788	ZOAB+	1.63E+08	1.24E+08	12832331	25353440	1.5	11	159.597	8028.55	565.843	1459.17
15	1HRR	126.7	4211	4211	ZOAB	1.51E+08	1.25E+08	9970163	16228663	1.2	11	159.597	8028.55	565.843	1459.17
15	1HRR	160.1	6611	6611	SMA	1.51E+08	1.32E+08	9185926	9553352	3.8	14	159.597	8028.55	565.843	1459.17
15	1HRR	160.2	6611	6611	SMA	1.52E+08	1.33E+08	9208381	9575777	2.9	12	159.597	8028.55	565.843	1459.17
15	1HRR	159.4	6611	6611	SMA	1.62E+08	1.39E+08	11080640	11777851	2.5	13	159.597	8028.55	565.843	1459.17
15	1HRR	159.3	6611	6611	SMA	1.74E+08	1.5E+08	11767547	12234373	2.6	16	159.597	8028.55	565.843	1459.17

References

Abdel-Aty, M., Devarasetty, P. C., & Pande, A. (2009). Safety evaluation of multilane arterials in florida. *Accident Analysis & Prevention*, *41*(4), 777–788.

Adlinge, S. S., & Gupta, A. (2013). Pavement deterioration and its causes. *International Journal of Innovative Research and Development*, *2*(4), 437–450.

Alaswadko, N., & Hassan, R. (2018). Rutting progression models for light duty pavements. *International Journal of Pavement Engineering*, *19*(1), 37–47.

Al-Masaeid, H. R. (1997). Impact of pavement condition on rural road accidents. *Canadian Journal of Civil Engineering*, *24*(4), 523–531.

Asphalt Institute. (1996). Superpave mix design. *Superpave Series No. 2 (SP-2)*.

Asphalt Institute. (2003). *Construction of hot mix asphalt pavements: Manual Series No. 22 (Second Edition)*.

Brown, E., & Cross, S. (1989). A study of in-place rutting of asphalt pavements. *National Center for Asphalt Technology*.

Cary, W. (1960). The pavement serviceability-performance concept. *HRB Bulletin*, *250*.

Chen, D.-H. (2009). Slippage failure of a new hot-mix asphalt overlay. *Journal of Performance of Constructed Facilities*, *24*(3), 258–264.

Design Council. (2003). *Changing behavior by design.* Retrieved February 15, 2020, from `https://www.designcouncil.org.uk/sites/default/files/asset/document/Changing%20behaviour%20by%20design.pdf`

Design Council. (2005). *A study of the design process.* Retrieved February 15, 2020, from `https://www.designcouncil.org.uk/sites/default/files/asset/document/ElevenLessons_Design_Council%20(2).pdf`

Dutka, A. F., & Hanson, H. H. (1989). *Fundamentals of data normalization* (Vol. 196). Addison-Wesley New York.

Ebrahimi, B., Wallbaum, H., Svensson, K., & Gryteselv, D. (2019). Estimation of norwegian asphalt surfacing lifetimes using survival analysis coupled with road spatial data. *Journal of Transportation Engineering, Part B: Pavements*, *145*(3), 04019017.

Farrar, C. R., & Worden, K. (2006). An introduction to structural health monitoring. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *365*(1851), 303–315.

Freeme, C. (1983). Evaluation of pavement behaviour for major rehabilitation of roads. *National Institute for Transport and Road Research Technical Report RP/19/83, Pretoria, CSIR*.

Gillespie, T. D. (1993). *Effects of heavy-vehicle characteristics on pavement response and performance* (Vol. 353). Transportation Research Board.

Heath, A. C., Roesler, J. R., & Harvey, J. T. (2003). Modeling longitudinal, corner and transverse cracking in jointed concrete pavements. *International Journal of Pavement Engineering*, *4*(1), 51–58.

Hodges, J., Rolt, J., & Jones, T. (1975). The kenya road transport cost study: research on road deterioration.

Howard, T. (2014). Journey mapping: A brief overview. *Communication Design Quarterly Review*, *2*(3), 10–13.

Huang, Y. H. (1993). *Pavement analysis and design*.

Kartsonaki, C. (2016). Survival analysis. *Diagnostic Histopathology*, *22*(7), 263–270.

Kaur, M., & Garg, S. K. (2014). An approach to detect and classify bugs using data mining techniques. *International Journal of Advanced Research in Computer Science and Software Engineering*, *7*(4), 830–836.

Lee Rodgers, J., & Nicewander, W. A. (1988). Thirteen ways to look at the correlation coefficient. *The American Statistician*, *42*(1), 59–66.

Loizos, A., & Karlaftis, M. G. (2005). Prediction of pavement crack initiation from in-service pavements: A duration model approach. *Transportation research record*, *1940*(1), 38–42.

Mathworks. (2006a). *Cox proportional hazards model with time-dependent covariates.* Retrieved February 15, 2020, from `https://ww2.mathworks.cn/help/stats/cox-proportional-hazards-model-with-time-dependent-covariates.html`

Mathworks. (2006b). *Linear or rank correlation.* Retrieved February 15, 2020, from `https://ww2.mathworks.cn/help/stats/corr.html#responsive_offcanvas`

Miller, J. S., Bellinger, W. Y., et al. (2014). *Distress identification manual for the long-term pavement performance program* (Tech. Rep.). United States. Federal Highway Administration. Office of Infrastructure.

Miller, R., & Halpern, J. (1982). Regression with censored data. *Biometrika, 69*(3), 521–531.

Miradi, M. (2004). Neural network models predict ravelling and analyse material/construction properties. *Proceeding of Intelligent Systems and Control, Hawaii, USA.*

NHTSA. (2008). *National motor vehicle crash causation survey: Report to congress* (Vol. 811; Tech. Rep.).

Pais, J. C., Amorim, S. I., & Minhoto, M. J. (2013). Impact of traffic overload on road pavement performance. *Journal of transportation Engineering*, *139*(9), 873–879.

Parsley, L. L., & Robinson, R. (1982). The TRRL road investment model for developing countries (RTIM2).

Paterson, W. D. (1987). *Road deterioration and maintenance effects: Models for planning and management.*

Prozzi, J. A., & Madanat, S. M. (2000). Using duration models to analyze experimental pavement failure data. *Transportation Research Record*, *1699*(1), 87–94.

Rijkswaterstaat. (2006). *Meerjarenplanning verhardingsonderhoud ivon 2007-2011 (IR-M07-NL-1).* Rotterdam, the Netherlands.

Rijkswaterstaat. (2012). *Inweva verkeersintensteiteiten.* Retrieved February 15, 2020, from https://nis.rijkswaterstaat.nl/portalcontent/logon/p2_33.html

Rijkswaterstaat. (2017a). *Richtlijnen handelwijze bij onvoldoende stroefheid of teveel spoorvorming versie 1.2.* Rotterdam, the Netherlands.

Rijkswaterstaat. (2017b). *Schadebeoordeling- en meetmethoden bovenbouw.* Rotterdam, the Netherlands.

Rijkswaterstaat. (2019). *Lifetime of the pavement in the netherlands.* Retrieved February 15, 2020, from https://stefan-russel.shinyapps.io/Leon/

Safavian, S. R., & Landgrebe, D. (1991). A survey of decision tree classifier methodology. *IEEE transactions on systems, man, and cybernetics*, *21*(3), 660–674.

Sorum, N. G., Guite, T., & Martina, N. (2014). Pavement distress: a case study. *vol*, *3*, 274–284.

Stichting Studie Centrum Wegenbouw. (1987). *Schadcatalogus voor de visuele inspectie van wegen.* Ede, NL: Stichting Studie Centrum Wegenbouw.

Tsai, B., Harvey, J., & Monismith, C. (2003). The application of weibull theory in asphalt concrete fatigue performance prediction. In *Transportation research board annual meeting presentation, washington, dc.*

Wang, Y., Mahboub, K. C., & Hancher, D. E. (2005). Survival analysis of fatigue cracking for flexible pavements based on long-term pavement performance data. *Journal of transportation engineering*, *131*(8), 608–616.

Winfrey, R. (1969). Economic analysis for highways. (No.923).

Zhang, H., Lepech, M. D., Keoleian, G. A., Qian, S., & Li, V. C. (2009). Dynamic life-cycle modeling of pavement overlay systems: Capturing the impacts of users, construction, and roadway deterioration. *Journal of Infrastructure Systems*, *16*(4), 299–309.