A Microdata Analysis Approach to Transport Infrastructure Maintenance

Kristin Svenson

A Microdata Analysis Approach to Transport Infrastructure Maintenance

"I dwell in Possibility"

– Emily Dickinson

Doctoral Dissertation

A Microdata Analysis Approach to Transport Infrastructure Maintenance

Kristin Svenson

Abstract

Maintenance of transport infrastructure assets is widely advocated as the key in minimizing current and future costs of the transportation network. While effective maintenance decisions are often a result of engineering skills and practical knowledge, efficient decisions must also account for the net result over an asset's life-cycle. One essential aspect in the long term perspective of transport infrastructure maintenance is to proactively estimate maintenance needs. In dealing with immediate maintenance actions, support tools that can prioritize potential maintenance candidates are important to obtain an efficient maintenance strategy.

This dissertation consists of five individual research papers presenting a microdata analysis approach to transport infrastructure maintenance. Microdata analysis is a multidisciplinary field in which large quantities of data is collected, analyzed, and interpreted to improve decision-making. Increased access to transport infrastructure data enables a deeper understanding of causal effects and a possibility to make predictions of future outcomes. The microdata analysis approach covers the complete process from data collection to actual decisions and is therefore well suited for the task of improving efficiency in transport infrastructure maintenance.

Statistical modeling was the selected analysis method in this dissertation and provided solutions to the different problems presented in each of the five papers. In Paper I, a time-to-event model was used to estimate remaining road pavement lifetimes in Sweden. In Paper II, an extension of the model in Paper I assessed the impact of latent variables on road lifetimes; displaying the sections in a road network that are weaker due to e.g. subsoil conditions or undetected heavy traffic. The study in Paper III incorporated a probabilistic parametric distribution as a representation of road lifetimes into an equation for the marginal cost of road wear. Differentiated road wear marginal costs for heavy and light vehicles are an important information basis for decisions regarding vehicle miles traveled (VMT) taxation policies.

In Paper IV, a distribution based clustering method was used to distinguish between road segments that are deteriorating and road segments that have a stationary road condition. Within railway networks, temporary speed restrictions are often imposed because of maintenance and must be addressed in order to keep punctuality. The study in Paper V evaluated the empirical effect on running time of speed restrictions on a Norwegian railway line using a generalized linear mixed model.

Acknowledgements

"The Road goes ever on and on
Down from the door where it began.
Now far ahead the Road has gone,
And I must follow, if I can.
Pursuing it with eager feet,
Until it joins some larger way
Where many paths and errands meet.
And whither then? I cannot say."

– J.R.R. Tolkien

The road to a PhD is long and many people have accompanied, guided, and encouraged me along the way. First and foremost, I want to thank my supervisors Lars Rönnegård and Karin Edvardsson. You have patiently taught me how to navigate through the intricacies of research and academia. This dissertation would never have existed without your guidance and support.

I have also had the opportunity to receive input from a reference group consisting of Johan Lang at WSP Sweden, and Kenneth Natanaelsson at the Swedish Transport Administration. I am especially grateful to Johan Lang, who initiated the topic of my first ever research paper which became the foundation of the entire dissertation.

To all my colleagues at Dalarna University: you have made my PhD road ever so enjoyable! Kenneth Carling, you did a tremendous job as research leader, but even more so as leader of the medicine ball squad. To run with you to work through pathless fields and forests was a true adventure. Anders Forsman, you solved all my practical issues faster than lightning and always make your staff feel appreciated. My co-author, Moudud Alam, you have the rare ability to provide constructive criticism. Ola Nääs, your interest in new teaching techniques has motivated me to become a better teacher. Zuzana Machucova, your curiosity and eagerness to learn inspired me throughout our mandatory courses. Majbritt Felleki, you always willingly helped me out when I was a fresh and oblivious PhD candidate. Co-author Yujiao Li, your exceptional intellect and ever present smile made it delightful to work with you. Xiaoyun Zhao, your observations about life in general and Swedish culture in particular have been a pleasure to take part of. Daniel Wikström, you have pushed my limits in discussions as well as medicine ball exercises.

List of Papers

I   Svenson, K. Estimated Lifetimes of Road Pavements in Sweden Using Time-to-Event Analysis. *Journal of Transportation Engineering*, 140(11):04014056, 2014. Printed with permission from the American Society of Civil Engineers.

II   Svenson, K., Machucova, Z., Li, Y., Rönnegård, L. Evaluating Needs of Road Maintenance in Sweden with the Mixed Proportional Hazards Model. *Transportation Research Record: Journal of the Transportation Research Board*, 2589:51–58, 2016. Printed with permission from the Transportation Research Board.

III   Nilsson, J-E., Svenson, K., Haraldsson, M. Estimating the Marginal Costs of Road Wear. 2015. *Submitted*.

IV   Svenson, K., McRobbie, S., Alam, M. Detecting Road Pavement Deterioration with Finite Mixture Models. 2016. *Submitted*.

V   Svenson, K., Seim Amdahl, A., Landmark Dypvik, A. The Effect on Running Time of Temporary Speed Restrictions on a Norwegian Railway Line. 2016. *Submitted*.

Contents

Abbreviations

AADT – Annual average daily traffic

AIC – Akaike Information Criterion

ASCE – American Society of Civil Engineers

BIC – Bayesian Information Criterion

ELPV – Enhanced longitudinal profile variance

EM – Estimation-maximization

ESAL – Equivalent Single Axle Load

GLMM – Generalized Linear Mixed Model

HAPMS – Highways Agency Pavement Management Systems

HDM – Highway Design and Maintenance Model

ICAR – Intrinsic Conditional Autoregressive

IRI – International Roughness Index

LCCA – Life-Cycle Cost Analysis

NS – Near side

NVDB – Nationella Vägdatabasen (National Road Database)

OS – Off side

PMS – Pavement Management Systems

RMST – Root Mean Squared Texture

SINTEF – Stiftelsen for industriell og teknisk forskning (the Foundation for Scientific and Industrial Research)

TIOS – Train Traffic Information and Monitoring Systems

TRACS – Traffic-speed Condition Surveys

TSR – Temporary Speed Restriction

VMT – Vehicle Miles Traveled

VTI – Statens väg- och transportforskningsinstitut (Swedish National Road and Transport Research Institute)

Part I

1. Introduction

This is a doctoral dissertation in microdata analysis; a multidisciplinary field that is concerned with the gathering, summarizing, modeling, and interpretation of large quantities of data. Modern society contains complex processes which require a wide knowledge of data collection and processing techniques, as well as skills in data analysis methods and the capability of data-driven decision making.

A crucial and complex system where the microdata analysis approach is well suited is transportation. The movement of goods and people between locations involves an integrated chain of vehicles, infrastructure, and operations. As the level of domestic and international integration of the transportation system increases, so does the importance of a more effective usage of data. Global economic welfare is highly dependent on the efficiency of the transportation network. This makes questions regarding infrastructure a core issue of our time.

Infrastructure asset management is the process of reinvestment, rehabilitation, and maintenance during the life-cycle of a facility. A central part of the transportation network's efficiency is the maintenance planner's ability to allocate the right type of maintenance, at the right place, and at the right time under the constraint of a limited budget. This is not an easy task with many simultaneous considerations regarding the short and long term effects of any maintenance decision.

Unprocessed data can provide a snapshot of a road or railway network's current condition. However, in order to maximize the overall efficiency of a maintenance decision, statistical models can convey information that is more relevant from a long-term perspective. The efficiency of a maintenance activity depends on its longevity in combination with the economic cost, which can be very challenging to assess. There are also situations where the direct impact of a maintenance decision can be difficult to determine, such as the delay caused by maintenance in a railway network. Statistical modeling is an appropriate method to analyze both the short and the long-term effects of maintenance decisions. Statistical models can also ensure that alternative maintenance strategies are evaluated as objectively as possible. Opportunistic and ad hoc decisions can be avoided by using empirical results, which enables a more efficient use of resources.

1.1   Aim of the dissertation

This dissertation aims at improving the efficiency of transport infrastructure maintenance by applying a microdata analysis approach. The objective of the research is methodological validity as well as the output's practical relevance.

1.2   Dissertation outline

The dissertation consists of two parts: Part I includes an introduction, the research background, a summary of the scientific papers, a discussion with suggestions for future research, and concluding remarks. Part II consists of the five papers which are published or submitted to be published in scientific journals.

2. Research background

2.1 The importance of transport infrastructure maintenance

The transport infrastructure system is crucial to maintain and increase economic welfare (see e.g. Wachs and Taylor 1998; Geurs and van Eck 2001; Weisbrod 2008; Rietveld and Bruinsma 2012). Investments in development and extensions of transportation systems and facilities are thus important, but even more so is maintenance and preservation of existing infrastructure. Lately, many studies and technical reports have emphasized the importance of prioritizing maintenance of roads and railways in the United States, Europe, and other parts of the world. In his review essay of infrastructure investment, Gramlich (1994) concludes that maintenance often has a higher return on investment than new constructions, although federal subsidies in the United States in many cases give priority to the latter. Gramlish proposes placing priority on increasing the budget for highway maintenance. The American Society of Civil Engineers (2013) rates American railways and roads as having poor to mediocre quality and encourages the use of asset management to facilitate efficient maintenance investments. Kahn and Levinson (2011) strongly advocate that maintenance of existing roads in the United States should be given financial priority over new builds.

In a report from the Swedish National Road and Transport Research Institute, Nilsson (2013) identifies system failures in the Swedish transport sector and concludes that a large number of infrastructure investments are undertaken despite not delivering a social economic net value. Nilsson also states that there is no tradition or possibility in Sweden to compare the outcome of projects and maintenance activities, which further jeopardize an efficient use of tax money spent on infrastructure.

Gwilliam and Shalizi (1999) review studies from industrial as well as developing countries, and the results are similar all over the world: deferring road maintenance increases not only total costs in the transportation network but also the present value of future costs. This implies that the economy as a whole in any country strongly benefits from timely road maintenance. Grimes and Barkan (2006) arrive at the same conclusion for railways: increased future expenses will more than offset temporary reductions in capital spending if railroad renewal maintenance is constrained.

2.2 Models and tools for efficient infrastructure asset management

Considering the great importance of maintenance and infrastructure asset management, research within these fields is more urgent than ever. Tools that can improve the efficiency of transport infrastructure maintenance are of great interest to many professionals within the industry.

Life-cycle cost analysis (LCCA) is one of the main methods used to obtain efficient maintenance decisions. LCCA was first introduced by the United States Department of Defense, and later developed and applied in a number of different industries and businesses (see e.g. Sherif and Kolarik 1981; Fabrycky and Blanchard 1991; Woodward 1997). In 1998, the United States Federal Highway Administration provided extensive guidelines in good practice use of LCCA in the field of pavement design. They also introduced a probabilistic approach to describe the uncertainty inherited in many of the input parameters (Walls and Smith, 1998).

The World Bank's Transportation Department has since the late 1980's developed the Highway Design and Maintenance Standards Model (HDM) to simulate total life-cycle conditions and costs for single roads or entire road networks (Watanatada, 1987). A further development in recent years is HDM-4 (Highway Development and Management Tool, see Kerali et al. 2006); a tool for strategic planning of maintenance costs to obtain a suitable road standard, to identify maintenance candidates, and to rank and value competing infrastructure investments (Schutte, 2008). However, there is a constant need for further development of effect models for both roads and railways, as concluded by e.g. Andersson et al. (2011) in their report about frameworks for economic analysis of the operation, maintenance, and renewal of transport infrastructure.

Other types of interactive asset management tools are rapidly developing within the infrastructure sector. One example is the Swedish Transport Administration that publishes road condition data from their Pavement Management Systems database online in a project called PMSv3. The aim of PMSv3 is to "be a help in deciding which road sections that should be prioritized for maintenance activities" (Swedish Transport Administration, 2016). The user interface allows maintenance engineers as well as the public to freely access road condition and network data by selecting the desired road sections on a map.

The combination of a reliable modeling framework, such as LCCA and HDM-4, and a comprehensible presentation of data, similar to the interactive visualization through a map within PMSv3, provides a good foundation for efficient infrastructure asset management. The objective of microdata analysis is to

take both of these aspects into account and make sure that the input part (data and modeling) is always connected to the output part (interpretation and presentation).

2.3 The microdata analysis approach

The aim of microdata analysis is to bridge the gap between data, information, and knowledge. The scheme in Figure 1 illustrates how data is collected, stored, analyzed, and presented in order for decision makers to access relevant information and gain deeper knowledge.

To achieve the ultimate goal of increased knowledge, microdata analysis comprises a number of collaborating fields such as artificial intelligence, decision support systems, resource allocation, data modeling, experimental design, simulation, and statistical inference (Dalarna University Doctoral Programmes Board, 2013). This dissertation emphasizes the use of statistical models and inference. The focus is primarily on practical applications, but methodological soundness and model validity are also important aspects.
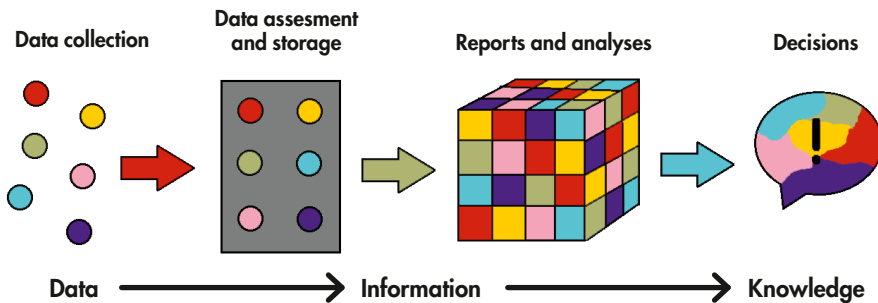
Figure 1: The microdata analysis approach.

2.4 Statistical modeling

Uncertainty and stochastic behavior occur in many settings connected to infrastructure asset management. Statistical modeling is a natural choice for handling probabilistic events. As described by Konishi and Kitagawa (2008), the purpose of a statistical model is: "to construct a model that approximates the true structure (of the underlying probability distribution) as accurately as possible through the use of available data." Statistical modeling also makes it possible to quantify uncertainty and variability in the data through statistical inference (Casella and Berger 2002; Davison 2003).

Some examples using different types of statistical models in infrastructure asset management applications are Futura et al. (2011) who use a log-normal distribution to estimate the seismic damage probability of bridges in a road network; Tighe (2001) who investigates the distributions of different LCCA input variables, such as pavement type and pavement thickness; and van Noortwijk and Frangopol (2004) who model the deterioration of transport infrastructure as a stochastic gamma process.

As the amount of and access to data grow within the transport infrastructure sector, statistical models become even more useful in LCCA of road and railway systems. A crucial stochastic input parameter in any LCCA is the expected lifetime of the object of interest; for example, the lifetime of a road pavement. Another stochastic parameter is deterioration, where the rate of deterioration is highly variable and important to consider in order to make efficient maintenance decisions. A third example within the railway sector is the correlation between maintenance and punctuality, where the latter has a very crucial role in railway LCCA (Hokstad, 1998). Stochastic elements, such as driver behavior, influence the predictability on how temporary speed restrictions caused by maintenance affect punctuality.

In this dissertation, all of the above scenarios are addressed using statistical modeling as a proposed solution to facilitate and improve maintenance decisions.

3. Summary of the papers

Paper I: Estimated lifetimes of road pavements in Sweden using time-to-event analysis

In Paper I, time-to-event analysis was used to estimate lifetimes of Swedish road pavements. Effective long term maintenance planning requires reliable estimates of roads' lifetimes, where the lifetime of a road is commonly defined as "the time interval between two maintenance activities" (Do 2011; Smith et al. 2006; Smith et al. 2005; Gharaibeh and Darter 2003, Hall et al. 1994). Major events such as resurfacing were considered as maintenance activities, but not smaller projects such as pothole fixes. The definition of a road's lifetime was extended to include road condition. The pavement life of a road was considered to have ended either when it was maintained, or when measurements of International Roughness Index (IRI) and/or rut depth had exceeded recommended levels in the maintenance standard stated by the Swedish Transport Administration. After an initial cleaning, which included removal of sections shorter than 50 meters, the data material comprised 266,614 homogeneous road sections from the Swedish Pavement Management Systems (PMS) database. The data covers maintenance activities carried out on national roads in Sweden from the 1960's until 2012. Data quality improved after 1987 when regular road condition measurements were introduced by the Swedish Transport Administration.

All roads that have not yet experienced the event of interest – being either the next maintenance activity or exceeding maintenance standard limits for IRI and/or rut depth – have unknown, i.e. censored, lifetimes. Time-to-event analysis, also known in the literature as survival analysis, is an established method to handle censored data (Klein and Moeschberger, 2005). The robust semi-parametric Cox proportional hazards model (Cox, 1972), which does not require any assumptions about the distribution of the data, was used to capture the effect on roads' lifetimes of seven covariates: pavement type, pavement stone size, road type, bearing capacity, road width, speed limit, and climate zone. Because traffic load is correlated with all other variables, the model was stratified based on eight traffic classes used by the Swedish Transport Administration in their maintenance planning. The stratification implies that each traffic class has a unique, non-parametric baseline hazard, but that the effect of the covariates is constant over all traffic classes.

The effect of the covariates was presented as hazard ratios, which in the setting of road maintenance can be interpreted as "the risk of needing maintenance in the next instant in time". A high risk of needing maintenance is equivalent to a shorter expected lifetime, and vice versa. Among the nine analyzed pavement types, stone mastic had the longest expected lifetime with a hazard ratio estimated to be 36 percent lower than asphalt concrete, the reference category. Among road types, 2+1 roads had 22 percent higher hazard ratio than ordinary roads, indicating significantly shorter lifetimes. Increased speed shortened the lifetime, while increased stone size (up to 20 mm) and increased road width extended the lifetime.

The estimated lifetimes produced by the model can be used as an input variable in different types of LCCA, for example when calculating the marginal cost of road deterioration (Paper III). The hazard ratios quantify the effect of the covariates and can help maintenance planners to evaluate different maintenance strategies (e.g. choice of pavement type or stone size) in terms of longevity.

Paper II: Evaluating needs of road maintenance in Sweden with the mixed proportional hazards model

The research in Paper II addressed the issue of latent variables in road maintenance data. National road databases often lack certain information that can be of great importance for long-term maintenance planning of paved roads. Some variables are complicated or expensive to measure, and others are difficult to quantify. In the Swedish case, latent variables of which there are no recordings in the PMS database are e.g. underlying road construction, subsoil conditions, and amount of heavy traffic in terms of Equivalent Single Axle Load (ESAL).

Certain characteristics of a road section may change between maintenance activities, such as pavement type or speed limit, but the geographical location of the road is still the same. Even when the time-to-event model includes several explanatory variables, the effect on the road's lifetime of the underlying construction design or ground materials (solid bedrock or looser sand) cannot be captured. In the Swedish PMS data, heavy traffic is distinguished from light traffic based on the length of the vehicle, but the exact weight or ESAL of each vehicle is unknown. Because national roads generally are constructed and maintained based on ESAL assumptions, this is a very important latent variable.

To capture the effect of latent variables, the Cox proportional hazards model used in Paper I was extended to a mixed proportional hazards model with random effects. The data material from the PMS database used in Paper I was also extended to include sections shorter than 50 meters, resulting in a data set of 548,091 observations. Because the spatial variation between sections was of interest, short neighboring sections were included in the analysis.

Estimation of random effects enabled identification of sections that have shorter or longer lifetimes than could be expected from the observed explanatory variables (traffic load, pavement type, road type, climate zone, road width, speed limit and bearing capacity restrictions). The random effects were assumed to be normally distributed, $N(0, \sigma_b^2)$, where σ_b^2 is the variance. A likelihood ratio test of the Cox model versus the mixed proportional hazards model showed that the variation captured by the random effects was significant.

There is also a spatial effect present in the maintenance data. In practice, road sections are seldom maintained independently of each other. If a road section needs maintenance because of e.g. ruts or cracks, neighboring sections, which are still functional, might be more likely to need maintenance in the near future. Therefore, maintenance planners can choose to maintain functional sections close to a non-functional section simply because of their spatial relationship. In order to test the spatial correlation between road sections with respect to maintenance, an Intrinsic Conditional Autoregressive (ICAR) model (Besag and Coperberg, 1995) was fitted to the random effects. The ICAR model estimated that 17 percent of the variance of the random effects were explained by spatial dependency. This fairly low percentage implied that ad hoc maintenance of still functional sections was not the main latent effect behind maintenance decisions. Other factors (subsoil conditions, ESALs, road construction, etc.) caused 83 percent of the unexplained variation.

The study in Paper II concludes that results from the mixed proportional hazards model are useful for maintenance planning because the weakest and strongest sections in a road network can be identified. The effect of the latent variables are visualized by plotting the random effect of each section in a map of the road network (see Figure 3 in Paper II). The Swedish example shows that the mixed proportional hazards and ICAR models are suitable for analyzing the effect of latent variables in national road databases.

Paper III: Estimating the marginal costs of road wear

The aim of Paper III is to estimate the marginal cost of road wear in Sweden. This was done by proposing a model for the short run marginal infrastructure cost, i.e. costs related to the impact of resurfacing for one additional vehicle using a road. Furthermore, three hypotheses were tested in the marginal cost model:

1. Roads deteriorate because of impact of heavy vehicles, expressed as Equivalent Single Axle Loads (ESALs)

2. Roads deteriorate because of impact of light vehicles

3. Roads deteriorate over time, independent of traffic load

The marginal cost model used in this study was first developed by Haraldsson (2007) in his licentiate thesis. The final equation (1) is the derivative of the present value of all future overlay costs (*PVC*), with respect to the traffic quantity (*Q*) over the period. *Q* was divided in ESALs (Q_{ESAL}) and light vehicles (Q_{CARS}), i.e. $Q = (Q_{ESAL}, Q_{CARS})'$, and marginal costs were calculated separately for each quantity. The *PVC* depends on the deterioration elasticity (ε), the cost of the pavement material (*C*), the expected lifetime of that pavement (*E(T)*), the discount rate (*r*) and the remaining lifetime of the road section, which is represented by a probability density function with parameters γ and α.

$$E\left(\frac{\delta PVC}{\delta Q}\right) = -\varepsilon \frac{C}{E(T)Q} \frac{r}{(1 - e^{-rT})} \int_0^\infty e^{-rv - \gamma v^\alpha} \, \mathrm{d}v \tag{1}$$

The Weibull distribution was chosen to represent the probability density function for the remaining lifetimes. This choice was motivated partly by the flexible nature of the Weibull hazard, and partly by the fact that the Weibull distribution is the only parametric distribution that exhibits proportional hazards. Estimated lifetimes from the more robust Cox model were compared to estimated lifetimes from the Weibull model, resulting in a difference of 8–16 percent longer lifetimes assuming a Weibull distribution.

The data for the lifetime estimation was equivalent to the PMS data set used in Paper I. Cost data differentiated between three different pavement types (Cold, Hot and Surface Dressing) and six different regions (North, South, Stockholm, Middle, East and West) was calculated from maintenance contracts in 2012 and 2013. The marginal cost function was applied to data from the

Swedish National Road Database (NVDB), a database that covers information about the current national roads in Sweden, but no historical data.

A national average marginal cost of 0.32 SEK for ESAL and 0.027 SEK for light vehicles was calculated based on equation (1). The elasticities estimated by the Weibull model provided evidence in favor of hypothesis 1) and 2): ESALs as well as light vehicles had a significant impact on road deterioration. An increase in ESAL or number of cars by 10 percent reduced the service life of pavements by about one percent for both categories. The model showed no support for hypothesis 3), i.e. there was no indication that time per se is a driving factor for deterioration. However, since a 50 percent increase in traffic did not reduce road lifetime to a similar extent, factors other than traffic seem to have an impact on road deterioration rate. As concluded in Paper II, a possible explanation is latent variables such as subsoil condition and road construction.

The marginal costs of road wear are fundamental for policy making regarding e.g. vehicle miles traveled (VMT) taxes. In countries with a freeze-thaw cycle, light traffic can cause road deterioration because of studded tire use. This is important to address in both maintenance and taxation policy decisions.

Paper IV: Detecting road pavement deterioration with finite mixture models

In Paper IV, road condition data from a part of the M4 highway in England – surveyed between January 2013 and January 2015 – was used to identify road segments with the most rapid deterioration rate. Budget restrictions limit the number of possible maintenance activities in a road network each year. Consequently, not all segments with a similar surface condition can be maintained. When making practical decisions about maintenance locations, the maintenance engineer has to look at the condition data and determine which sections should be maintained as a priority, and which can be left until later. In this decision, deterioration rate can be more important than the actual condition: maintaining a segment which shows a rapid deterioration can be a more efficient use of resources than maintaining a segment with a slightly worse, but stable, condition.

In order to identify segments with a deterioration rate that departs from a stationary condition, data was assumed to originate from two different normal distributions – a "change" distribution and an "unchanged" distribution. All segments were clustered into either of these distributions using finite mixture models. In order to estimate the parameters of the distributions and to classify the segments, eleven road condition variables from Highways England's Traffic-speed Condition Surveys (TRACS) were combined in a multivariate finite

mixture model (McLachlan and Basford, 1988). The road condition variables are left and right rutting; near side and off side enhanced longitudinal profile variance (ELPV) measured at 3, 10 and 30 meters; and near side, off side and middle Root Mean Square Texture (RMST). The Bayesian Information Criterion (BIC) implied a normal mixture with an unrestricted covariance matrix for these variables, allowing for a flexible correlation structure between the variables and distributions.

The parameter estimates obtained by the finite mixture models matched the assumption of a mixture known as an outlier distribution: two normal distributions with a possible common mean but different variances. The unchanged distribution was very narrow, showing no sign of deterioration. The estimated variances of the change distribution were 5.7 up to 300 times larger than the variances of the unchanged group, for all variables. The segments classified into the change group were compared to data from maintenance records, as some of them might have changed because of maintenance and not deterioration. Locations of sections in the maintenance records matched some of the segments classified into the change group. Not all of the road condition variables show positive change when a segment is deteriorating and negative change when a segment is maintained, and therefore it is crucial to check whether a segment has been subject to maintenance. The finite mixture models successfully identified segments that had changed. Additional data sources (e.g. maintenance records) were necessary to separate change because of maintenance and change because of deterioration.

The study in Paper IV shows that finite mixture models, in combination with maintenance records, can be helpful in finding segments with the highest deterioration. This is useful for maintenance planners when prioritizing future maintenance decisions.

Paper V: The effect of temporary speed restrictions on the running time of a Norwegian railway line

The study conducted in Paper V evaluated the effect of temporary speed restrictions (TSRs) on train running time in the Norwegian railway network. Speed restrictions are often imposed when maintenance work is carried out on the railway track. A TSR's effect on running time must be considered in order to correctly adjust timetables. The actual running time of a train may not always be the same as the theoretical in the presence of a TSR. Theoretical deterministic models of temporary speed restrictions typically include running time as a func-

tion of the speed restriction, the normal speed, the length of the TSR segment, and an acceleration and deceleration coefficient. In reality, additional influencing factors can vary substantially between individual station blocks, weeks, days, and trains. The theoretical model cannot capture effects of the train driver's behavior, which could influence the running time considerably. The driver may try to reduce the effect of a temporary speed restriction by increasing the speed on the remaining non speed-restricted part of a station block. The nature of the TSR segment (i.e. curvature or height profile) and potential seasonal effects are other variables that can influence the real effect of a TSR. Maintenance planners also may want to know whether one longer or several shorter segments with reduced speed have a greater impact on the expected running time.

A generalized linear mixed model was fitted to data from the Norwegian Train Traffic Information and Monitoring Systems database (TIOS). Trains running on the Dovre line, recorded between December 2009 and December 2015, were used in the analysis. The response variable was the running time between two stations (one station block). If the train stopped at a station, running time was defined as the time between a recorded departure from one station and the recorded arrival at the next station. If the train had no stop at the station, running time was defined as the time between recorded departures. The train travel data was combined with temporary speed restriction data during the same period.

Running time has a non-negative, skewed distribution, and a gamma distribution had the best fit to the data. To account for the heterogeneity between station blocks, a random intercept was introduced in the model. A likelihood ratio test showed a significant effect of a nested random slope for the variable Δspeed, defined as the difference between normal speed and the reduced speed limit imposed by the TSR. For the station blocks with the highest normal speed, Δspeed had minimal impact on running time. However, for station blocks with the lowest normal speed, it had a very strong effect. This result implied that the train drivers' behavior had an impact on running time. Drivers tended to increase the speed, where possible, in order to reduce the effect of a TSR.

The global model of fixed effects showed that running times were shorter in winter and on weekends. A positive height difference after a TSR segment had an effect on freight trains, indicating that acceleration up to normal speed is slowed down by an uphill. The distance from a TSR to the next station had a negative effect on running time for both freight trains and passenger trains – i.e. drivers had a greater possibility to reduce the time loss by driving faster on the remaining part of the block when the TSR was located further from the next

station. The length of the TSR segment had little or no impact on running time, implying that acceleration and deceleration are more important factors than the actual length traveled at reduced speed.

The results presented in Paper V show that a generalized linear mixed model can provide a better understanding of how empirical effects – the location of the temporary speed restricted segment, the height profile of the railway track, and driver behavior – will affect train running times in the presence of temporary speed restrictions. When maintenance is planned on the railway, and a temporary speed restriction is imposed as a consequence, the empirical results can be used as a complement to the theoretical model for more accurate rescheduling of timetables.

4. Discussion

4.1 Microdata methodology

The microdata analysis chain consists of several steps – from data collection to reporting on results – that are equally important (see Figure 1). Depending on the nature and aim of the specific research question, certain parts of the chain may be emphasized. This dissertation focuses on data analysis. The different data sets used in the five papers all originate from secondary sources: the Swedish Pavement Management Systems' database, the Swedish National Road Database, Highways England's Traffic-speed Condition Surveys, the English Highways Agency Pavement Management Systems' database, and the Norwegian Train Traffic Information and Monitoring Systems' database. Therefore, experimental designs and first-hand data collection are not discussed in detail in any of the studies.

Data assessment
The second step in the microdata analysis process deals with data assessment and storage. Storage, in terms of database systems and data managing, has not been a focus area in this dissertation. However, the assessment of data quality is an important issue for secondary data sources.

In Paper I, the PMS data was cleaned with respect to section length to avoid influence from sections shorter than 50 meters (as short sections often are junctions, slip roads, or crossings). The PMS database is managed by imputing fictitious maintenance activities for sections where the road condition (measured by IRI and rut depth) has improved, and by removing registered maintenance activities for sections where the road condition is unchanged. Because this management could have an impact on the parameter estimation in the Cox model, a sensitivity analysis was performed using the entire data set with no removals or imputations. The results showed no large differences in parameter estimates between models fitted to the cleaned versus the full data sets. The analysis in Paper II is based on the very same data set. Because the sensitivity analysis ensured that results from the Cox model are fairly robust, all road sections were included in the analysis performed in Paper II. Using all sections was appropriate in order to account for spatial correlation between sections. In Paper III, an assumption was made about the number of ESALs in the Swedish road network. Therefore, the robustness of the estimated marginal costs was evaluated using sensitivity analysis with respect to the number of ESALs.

The data used in Paper IV is aggregated road condition measurements for 10-meter road segments. To validate the quality of this data, sources of systematic errors were minimized. As a result, only surveys from one single contract period were used in the study to avoid errors caused by different equipment. Within a contract period of typically four years, one company is contracted to measure the road condition with calibrated equipment that is accredited according to defined criteria each year. To ensure that measurements from successive surveys are aligned, a two-step alignment algorithm based on GPS and profile data was implemented. This improved alignment to within a few centimeters. The random measurement errors in the data were quantified by comparing surveys taken within an interval of a few days.

The data quality in Paper V was improved by removing trains with incorrect timetables. To ensure that the calculated running times were as accurate as possible, only trains that had a registered departure time from all stations were included in the analysis.

Data analysis

Data analysis is the crucial step of turning raw data into information. Microdata analysis includes methods from e.g. computer science, geography, mathematics, and statistics, enabling a wide range of analysis methods. In an argumentative paper, Breiman (2001) advocates the use of what he calls algorithmic models (machine learning techniques such as decision trees and neural networks, where the data-generating process is unknown) over data models (statistical models assuming that data is generated from a given stochastic data model). In this dissertation, the choice of modeling framework is data models. The decision is two-folded: statistical data models deal with inference about the underlying distribution, and the techniques used for parameter estimation are validated by probability theory. In contrast, many of the algorithmic models found in the machine learning field are focused on predictive accuracy and deal with "black box" solutions.

When predictive precision is the sole interest of the analysis, algorithmic models can often outperform statistical data models (Breiman, 2001, p.214). However, for identification of potential causal factors, statistical models have a theoretical foundation that black box solutions do not adequately provide. By assessing parameters from a distribution through inference, associations among variables and estimation of beliefs and/or probabilities of past and future events can be derived. The studies in Paper I, II and V are all focused on establishing the effect explanatory variables have on the response variables (pavement life-

time and train running time). For the clustering of road segments in Paper III, a supervised algorithmic model such as random forest could have been an option. However, the data material was not suitable for such a classification model. No data source could provide sufficient information for evaluating model performance through cross-validation. The choice of a distribution-based clustering method – finite mixture models – provided parameter estimates that could be compared to the empirical measurement errors, thus allowing a possible verification of the model's validity.

Decision-making

The last step of the microdata analysis approach addresses decision-making and actions. The studies in Papers I–V produce input to different stages of LCCA, one of the main methods for efficient decision-making within the field of infrastructure maintenance (see Section 2.2). Pavement lifetime, pavement deterioration rate, and train running time are all important variables to consider in a LCCA. The first of the three cases is an issue when making long-term maintenance decisions within a time horizon of 10–30 years. The latter two are usually addressed within a shorter time perspective: how to prioritize maintenance of deteriorating road segments in next year's budget, or how to obtain punctuality when performing maintenance that requires a temporary speed restriction.

The study in Paper III assesses a topic with a more direct impact on economic policy decisions: the marginal costs of road wear. These costs can be a basis for efficient decisions regarding VMT taxes. Zhang and Lu (2013) suggest that a VMT tax in Maryland based on marginal costs would reduce the overall vehicle miles traveled, reduce pollution, and increase revenue compared to the existing revenue policy (fuel fees and tolls). The research conducted in Paper III was part of a commission from the Swedish government to the Swedish National Road and Transport Research Institute. Sweden currently has no VMT tax, but an inquiry for heavy vehicles is ongoing and a policy proposal should be finished by December 2016 (Swedish Government, 2015).

Limitations

The aim of the dissertation is to improve the efficiency of transport infrastructure maintenance by applying a microdata analysis approach. However, this objective excludes the creation of a complete decision support system and the collection of new data. Instead, efficiency is achieved by applying well-known statistical models to available data sources. The use of existing data and methods

is a very cost-effective way to increase information value. There are limitations to this approach, for example when necessary data is unavailable or when established methods are not up-to-date with the field's frontier questions. However, leveraging existing resources is an inexpensive and easily implementable way to improve decision-making, even when the choice is made to conduct additional experimental research.

The studies in this dissertation do not provide a full-scale solution to efficient infrastructure maintenance decision-making. Each of the studies contribute separate and complementary support for the use of microdata analysis in maintenance planning. The results should be implemented in a collaborative setting where professionals with different specialties (engineers, statisticians, programmers, economists, etc) can make practical use of the increased knowledge.

4.2 Future research

One recommended area of future research would be to integrate the dissertation's results with existing LCCA models, decision support systems, and visualization tools. Estimated lifetimes and random effects (Paper I and II) can be incorporated in tools such as the Swedish Transport Administration's PMSv3. By including analytical results, road engineers can access information that is not currently displayed by descriptive statistics.

The estimated marginal costs in Paper III was disaggregated to the maximum extent permitted by available data. However, the lack of ESAL measurements on Swedish roads prevented a more accurate calculation. A future research proposition is to evaluate the economic effects of a VMT tax based on marginal costs of road wear. This study should include a sensitivity analysis which quantifies the uncertainty derived from the ESAL assumptions.

The study in Paper IV used three consecutive years of road condition measurement data. With access to data from one additional survey (the last in the current contract period), the finite mixture models could be expanded to include longitudinal correlation. A longitudinal model could separate measurement errors and actual change with the possibility of increased accuracy. To further investigate if segments have changed because of maintenance or deterioration, a study which combines results from the finite mixture models with predictive models for pavement deterioration is advisable.

The model in Paper V focused on one-level-effects for individual station blocks. These results could be expanded to a hierarchical model that incorpo-

rates network level data. Thus, the effect of temporary speed restrictions on train running time from start station to end station on a railway line could be evaluated.

Microdata analysis applied to transport infrastructure maintenance is a novel research field and its potential is immense. The core of this dissertation is data analysis, but expanding the research to additional data collection and the development of tools to improve decision-making looks very promising. An example of additional useful data would be more accurate information about ESALs in the Swedish road network. The creation of decision tools which are intuitive and readily accessible to maintenance planners in the road and railway sectors would be beneficial. These could include software providing graphical presentations of research results, such as maps available both in fieldwork and for administrative purposes.

4.3 Author credits

Paper I
As the sole author of Paper I, Svenson has prepared the data, performed the statistical analysis, and written the paper. Paper I is a revision of Svenson's Master thesis at Uppsala University 2012, which was supervised by PhD Ingrid Persson, Uppsala University, and Johan Lang, WSP Sweden.

Paper II
Paper II is a joint work with Professor Lars Rönnegård and PhD candidates at Dalarna University, Yujiao Li and Zuzana Machucova. The application of the ICAR-model was designed in collaboration with Rönnegård and Li. The results were clearly displayed in a map constructed by Machucova. Svenson prepared the data, conducted the data analysis using the mixed proportional hazards model, and wrote the manuscript.

Paper III
The majority of the Paper III manuscript is written by Professor Jan-Eric Nilsson, the Swedish National Road and Transport Research Institute (VTI). Mattias Haraldsson, VTI, derived the marginal cost function. Svenson estimated the road lifetimes, wrote the parts about survival analysis (including an appendix), prepared the data, and did all the necessary calculations to obtain the final marginal costs.

Paper IV

Paper IV is written in collaboration with PhD Stuart McRobbie, Transportation Research Laboratory, UK, and PhD Moudud Alam, Dalarna University. Both have contributed to the analysis part of the paper: McRobbie with his deep knowledge of the TRACS data and road condition measurements in England, and Alam with suggestions and ideas about the finite mixture models. Svenson prepared the data, conducted the analysis, and wrote the paper.

Paper V

The manuscript of Paper V is written by Svenson, PhD Andreas Amdahl Seim, research manager at the Foundation for Scientific and Industrial Research (SINTEF), Norway, and Andreas Dypvik Landmark, research scientist at SINTEF. Svenson and Dypvik Landmark prepared the data and conducted the literature review. Svenson performed the statistical analysis and wrote the methodological parts. The discussion part was written in collaboration between all authors.

5. Conclusion

This dissertation adopted a microdata analysis approach to improve the efficiency of transport infrastructure maintenance. Microdata analysis allows great flexibility in selecting only the most useful methods for a given problem. The research produced valuable results within several different areas: estimation of road lifetimes, assessing the impact of latent variables on road lifetimes, input to a marginal cost model of road wear, detection of road deterioration, and evaluating the implications of temporary speed restrictions on train running times. The research results can be implemented in transport infrastructure asset management to facilitate decisions resulting in a more efficient use of resources.

References

American Society of Civil Engineers. 2013 Report Card for America's Infrastructure. URL: http://www.infrastructurereportcard.org/a/documents/Roads.pdf, 2013. Accessed 2016-11-21.

Andersson, M., Nyström, J., Odolinski, K., Wieweg, L., and Wikberg, Å. A strategy for developing a framework for economic analysis of operation, maintenance and renewal of road and rail infrastructures. Technical Report 707, The Swedish National Road and Transport Research Institute (VTI), 2011.

Besag, J. and Coperberg, C. On conditional and intrinsic autoregressions. *Biometrika*, 82:733–746, 1995.

Breiman, L. Statistical modeling: The two cultures (with comments and a rejoinder by the author). *Statistical Science*, 16(3):199–231, 2001.

Casella, G. and Berger, R. L. *Statistical Inference*, volume 2. Duxbury Pacific Grove, CA, 2002.

Cox, D. Regression models and life-tables. *Journal of the Royal Statistical Society*, 34(2):187–220, 1972.

Dalarna University Doctoral Programmes Board. Curriculum for Doctoral Studies in Microdata Analysis. URL: http://www.du.se/PageFiles/79500/Sep2014-Curriculum%20for%20postgraduate%20studies%20in%20Microdata%20analysis%20(2012).pdf, May 2013. Accessed 2016-11-21.

Davison, A. C. *Statistical Models*, volume 11. Cambridge University Press, 2003.

Do., M. Comparative analysis on mean life reliability with functionally classified pavement sections. *KSCE Journal of Civil Engineering*, 15(2):261–270, 2011.

Fabrycky, W. J. and Blanchard, B. S. *Life-cycle cost and economic analysis*. Prentice Hall, 1991.

Furuta, H., Frangopol, D. M., and Nakatsu, K. Life-cycle cost of civil infrastructure with emphasis on balancing structural performance and seismic risk of road network. *Structure and Infrastructure Engineering*, 7(1–2):65–74, 2011.

Geurs, K. and van Eck, J. R. Accessibility measures: review and applications. evaluation of accessibility impacts of land-use transport scenarios, and related social and economic impacts. Technical Report 408505006, National Institute of Public Health and The Environment, 2001.

Gharaibeh, N. and Darter, M. Probabilistic analysis of highway pavement life for Illinois. *Transportation Research Record: Journal of the Transportation Research Board*, 1823:111–120, 2003.

Gramlich, E. M. Infrastructure investment: A review essay. *Journal of Economic Literature*, 32(3):1176–1196, 1994.

Grimes, G. A. and Barkan, C. P. Cost-effectiveness of railway infrastructure renewal maintenance. *Journal of Transportation Engineering*, 132(8):601–608, 2006.

Gwilliam, K. and Shalizi, Z. Road funds, user charges, and taxes. *The World Bank Research Observer*, 14(2):159–185, 1999.

Hall, K., Lee, Y., Darter, M., and Lippert, D. Forecasting pavement rehabilitation needs for Illinois interstate highway system. In *Transport Research Record: 73rd Annual Meeting, Washington DC*, 1455:116–122, 1994.

Haraldsson, M. The Marginal Cost for Pavement Renewal a Duration Analysis Approach. In Essays on Transport Economics. Licentiate thesis, Uppsala University, 2007.

Hokstad, P. Life cycle cost analysis in railway systems. Technical Report STF38-A98424, SINTEF Industrial Management, 1998.

Kahn, M. E. and Levinson, D. M. Fix it first, expand it second, reward it third: New strategy for America's highways. Technical Report 2011-03, The Hamilton Project, 2011.

Kerali, H. G., Odoki, J., and Stannard, E. E. *Overview of HDM-4*. World Road Association, Highway Development and Management Series, 1st edition, 2006.

Klein, J. and Moeschberger, M. *Survival Analysis – Techniques for Censored and Truncated Data*. Springer, New York, 2nd edition, 2005.

Konishi, S. and Kitagawa, G. *Information Criteria and Statistical Modeling*. Springer Science & Business Media, 2008.

McLachlan, G. J. and Basford, K. E. *Mixture Models. Inference and applications to clustering*. Dekker, New York, 1988.

Nilsson, J.-E. System failures in Sweden's infrastructure sector. Technical Report 794, The Swedish National Road and Transport Research Institute (VTI), 2013.

Rietveld, P. and Bruinsma, F. *Is transport infrastructure effective?: Transport infrastructure and accessibility: impacts on the space economy*. Springer Science & Business Media, 2012.

Schutte, I. A User Guide to Road Management Tools. Sub-Saharan Africa Transport Policy Program, 2008. URL: https://books.google.se/books?id= pgf_nQEACAAJ. Accessed 2016-11-21.

Sherif, Y. S. and Kolarik, W. J. Life cycle costing: Concept and practice. *Omega*, 9(3):287–296, 1981.

Smith, K., Titus-Glover, L., Darter, M., Quintus, H. V., Substad, R., and Hallin, J. Evaluation of the cost benefits of continuous pavement preservation design strategies versus reconstruction: Final rep. 491. Technical Report FHWA-AZ-05-491, Arizona Department of Transportation, September 2005.

Smith, K., Titus-Glover, L., Rao, S., Quintus, H. V., and Stanley, M. Life-cycle cost analysis of SMA pavements and SMA application guidelines. Technical Report WHRP 06-11, Wisconsin Highway Research Program, August 2006.

Swedish Government. Kommittdirektiv 2015:47: Distance-based road wear tax for heavy trucks. URL: https://www.riksdagen.se/sv/dokument-lagar/dokument/kommittedirektiv/avstandsbaserad-vagslitageskatt-for-tunga_H3B147, April 2015. Accessed 2016-11-21.

Swedish Transport Administration. Pavement Management Systems PMSv3. URL: https://pmsv3.trafikverket.se/, 2016. Accessed 2016-11-21.

Tighe, S. Guidelines for probabilistic pavement life cycle cost analysis. *Transportation Research Record: Journal of the Transportation Research Board*, 1769:28–38, 2001.

van Noortwijk, J. M. and Frangopol, D. M. Two probabilistic life-cycle maintenance models for deteriorating civil infrastructures. *Probabilistic Engineering Mechanics*, 19(4):345–359, 2004.

Wachs, M. and Taylor, B. D. Can transportation strategies help meet the welfare challenge?. *Journal of the American Planning Association*, 64(1):15, 1998.

Walls, J. and Smith, M. R. Life-cycle cost analysis in pavement design: In search of better investment decisions. Technical Report FHWA-SA-98-079, US Department of Transportation Federal Highway Administration, 1998.

Watanatada, T. *The Highway Design and Maintenance Standards Model: User's manual for the HDM-III model*, volume 2. Johns Hopkins University Press, 1987.

Weisbrod, G. Models to predict the economic development impact of transportation projects: historical experience and new applications. *Annals of Regional Science*, 42(3):519–543, 2008.

Woodward, D. G. Life cycle costing: Theory, information acquisition and application. *International Journal of Project Management*, 15(6):335–344, 1997.

Zhang, L. and Lu, Y. Implementing marginal-cost vehicle mileage fees on the Maryland statewide road network. *Transportation Research Record: Journal of the Transportation Research Board*, 2345:53–62, 2013.

Part II

Paper I

Estimated Lifetimes of Road Pavements in Sweden Using Time-to-Event Analysis

Kristin Svenson[1]

Abstract: Maintenance planning of road pavement requires reliable estimates of roads' lifetimes. In determining the lifetime of a road, this study combines maintenance activities and road condition measurements. The scope of the paper is to estimate lifetimes of road pavements in Sweden with time-to-event analysis. The model is stratified according to traffic load and includes effects of pavement type, road type, bearing capacity, road width, speed limit, stone size, and climate zone. Among the nine analyzed pavement types, stone mastic had the longest expected lifetime with a hazard ratio (risk of needing maintenance) estimated to be 36% lower than asphalt concrete. Among road types, $2+1$ roads had 22% higher hazard ratio than ordinary roads indicating significantly lower lifetimes. Increased speed lowered the lifetime, while increased stone size (up to 20 mm) and increased road width lengthened the lifetime. The results are of importance for life-cycle cost analysis and road management. **DOI: 10.1061/(ASCE)TE.1943-5436.0000712.** © *2014 American Society of Civil Engineers.*

Introduction

To be able to perform long-term maintenance plans for the Swedish road network, reliable estimations of maintenance intervals are needed. The condition of the road network in Sweden is monitored by regularly performed measurements of the road surface. Roughness [in terms of international roughness index (IRI)] and rut depth are measured every 1–5 years. Since it is not economically or logistically possible to conduct surface measures on all roads every year, prediction models are implemented in the Swedish Transport Administration's application pavement management systems (PMS), which cover all state roads in Sweden. These are linear models that lack precision in the long run because the road wear is nonlinear. Normally, the rate of wear increases as the road ages. There is also a question of censored observations in the PMS database. These are roads that have only one registered maintenance activity (i.e., the next maintenance has not yet happened). If censoring is not considered, such a model is likely to be biased and misleading.

A method widely used for nonlinear and censored data to estimate lifetimes in mechanistic applications is time-to-event analysis. In medicine, this method is known as survival analysis. Time-to-event analysis has been used to model lifetimes of roads where "lifetime" is commonly defined as the interval between two maintenance activities (e.g., Do 2011; Smith et al. 2005, 2006; Gharaibeh and Darter 2003; Hall et al. 1994). The aim of this paper is to estimate lifetimes of road pavement in Sweden with time-to-event analysis as a way to provide reliable estimates for maintenance planning. In the planning process, it is also of interest to analyze the impact of different variables on the road's expected lifetime; such variables are pavement type, road type, bearing capacity, road width, speed limit, stone size, and climate zone. This will be

obtained by fitting a Cox proportional hazards model to road data from the PMS database.

Literature Review

Time-to-event analysis has been used in previous research on lifetime estimation of road pavements. Estimated expected lifetimes of four maintenance activities on flexible pavement are included in the U.S. Long-Term Pavement Performance (LTPP) program (Eltahan et al. 1999). A sample of 28 test sections were maintained in 1990 by the same contractor and then followed for 8 years. Survival curves fitted from the nonparametric Kaplan-Meier product limit estimator are used to estimate the time to failure, where failure is defined as when a section meets the criteria for poor condition (mainly cracking). The authors conclude that survival analysis is suitable for pavement maintenance data. They recommend using larger data sets and possibly parametric methods to account for factors that affect pavement performance, and also to better estimate the median lifetimes and cost-effectiveness.

The Illinois Department of Transportation (IDOT) also uses the Kaplan-Meier product estimator to find changes over time in expected pavement life and probability of failure (Gharaibeh and Darter 2003). The data consist of 1,402 homogeneous sections with a typical range of 0.8–8 km (0.5–5 mi) in length, where each section is homogeneous with respect to design, construction history, and traffic load. Termination of service life is defined as when a major rehabilitation action is carried out and results are presented as survival curves of six different pavement types. Smith et al. (2006) use a similar method to find the survival functions of stone mastic and hot mix asphalt using data from Wisconsin's pavement management systems. However, their definition of a road's lifetime differs. To determine service life, the authors fit linear regressions with IRI and PDI (overall pavement condition indicator) as response variables. They apply threshold values of IRI and PDI as stated in the guidelines of the Wisconsin Department of Transportation and estimate when each road section will pass this threshold. Roads with estimated service lives of 20 years or more are considered censored.

[1]Ph.D. Candidate, Dept. of Statistics, School of Technology and Business Studies, Dalarna Univ., Borlänge, 79188 Dalarna, Sweden. E-mail: kss@du.se

Road condition in terms of fatigue cracking is the response variable in a time-to-event analysis performed by Wang et al. (2005). Their data consist of 486 asphalt concrete road sections, and the failure of a section is defined by a threshold of 3.6% cracking. The study uses a parametric approach, and the generalized gamma distribution is found to be most appropriate for these data. Among several tested factors, thickness of the asphalt concrete surface layer, thickness of the concrete base, intensity of precipitation defined by wet days divided by total precipitation per year, and number of freeze/thaw cycles per year are found to have a significant effect on fatigue cracking. The estimated lifetimes have a large variance; therefore, the authors conclude that the survival model seems more suitable for comparing and quantifying the effect of influencing factors, rather than predicting exact failure times of a pavement.

A study by Do (2011) presents estimated mean lifetimes of highways in South Korea. A classification based on traffic volume–related variables, heavy vehicle–related variables, and directional characteristics distinguishes three types of roads: urban, rural, and recreational. The data cover 30 urban road sections, 103 rural road sections, and 47 recreation road sections. Do finds that the log-normal distribution is most suitable for urban roads and that recreation roads have the closest fit to Weibull distribution. For rural roads, no distribution is found appropriate, and therefore, a nonparametric estimation is made. Results from the analysis show that urban roads have a shorter mean life than rural and recreation roads. Do recommends that, in terms of for further research, more variables that could affect road lifetimes be added to the models.

In his doctoral dissertation, Dong (2011) evaluates the influence of different factors on the crack initiation of asphalt pavements using parametric survival analysis. The Weibull distribution is found to have the best fit to these data. Among the factors, traffic level, thickness of the pavement structure, freeze index, mixture, and mill are found to be significant. Dong uses LTPP data from 18 projects located in different states with nine test sections in each project. The author concludes that the study includes too few sections for a model for predicting failure times, and suggests that survival models at different traffic, environmental, and highway classifications should be developed for this purpose.

The literature review shows that both nonparametric and parametric methods have been used to estimate the lifetime of roads, defined as the time between maintenance activities or the time to failure according to some road condition criteria. However, the parametric approach requires specification of a distribution, and no unanimous distribution can be found in the literature. The nonparametric Kaplan-Meier estimator cannot include influencing factors that are necessary in a model for maintenance planning. This study uses the semiparametric Cox proportional hazards model in which no specific distribution is needed but where covariates can be included. Also, the definition of a road's lifetime used in this study will combine both maintenance activities and road condition information in order to provide more accurate estimates of roads' lifetimes. Because of the large data set used in this study, some of the variables suggested by other authors, such as different traffic and highway classifications, can be included in the model (Dong 2011; Do 2011).

Materials and Methods

Data Collection

The data material in this study was provided by the database in the Swedish Transport Administration's PMS. In 2012 it consisted of 390,966 observations representing homogeneous road sections. These sections are homogeneous with respect to characteristics such as pavement type, maintenance date, traffic load, road width, speed limit, and axle load restrictions. Sections vary in length, from over 1 km to only a few meters. To avoid influence from very short sections, only homogeneous sections that are 50 m or longer are used in this study.

When maintenance is carried out it is reported in the PMS database. The quality of the database varies as routine maintenance is seldom reported and periodic maintenance is sometimes not reported. The opposite scenario—that maintenance is reported but no maintenance has been performed—also occurs, often due to errors in locating the maintenance activity (Gustafsson and Lundberg 2009). In some cases, when maintenance is detected in data due to drastic changes (i.e., significantly lower values) in the measures of IRI and rut depth, a fictitious maintenance is inserted to the data. If a maintenance activity is registered but no significant changes are seen in IRI or rut depth, this registered maintenance is removed and sections are merged. However, all reported maintenance activities that do not affect the condition of the road's condition are not necessarily errors. Some road sections are maintained although their condition is good, often because adjacent sections are bad. Using maintenance activities performed on good sections in further analysis can give biased results, and estimated lifetimes will be shorter than they would have been if only the bad sections were maintained. On the issue of maintenance planning, the interest is to know the potential lifetime of the road (i.e., when the road needs maintenance due to bad condition).

There is a risk that the data management will affect the results, because IRI and rut depth do not capture, for example, edge deformations, which are also a cause for maintenance. To see the effect of the data management, a sensitivity analysis is performed by fitting the Cox proportional hazards model to the entire data set. The results from the sensitivity analysis showed that the cleaning had little effect in most parameter estimates. However, a few variables were affected, and these are mentioned and analyzed in the "Discussion" section. After cleaning, the data material used in further analysis consists of 266,614 homogeneous road sections.

Variables

Summary statistics of the qualitative variables selected in this study can be found in Tables 1 and 2. These variables are pavement type, traffic load in terms of average annual daily traffic (AADT), road type, bearing-capacity class (BCC), maximum stone size of the pavement material used on the road, and climate zone. The variables measured on a quantative scale are found in Table 3. These are road age, road width, and speed limit.

The PMS database contains six different road types: ordinary two-lane roads, motorways, undivided motorways, four-lane roads, ordinary roads with cable barriers, and undivided motorways with cable barriers. The two latter types are also known as 2 + 1 roads, which they will henceforth be called in this paper. In 1998, the Swedish Transport Administration started to rebuild some ordinary two-lane roads (13 m or wider) and undivided motorways into 2 + 1 roads. This road type consists of three lanes, with two lanes in one direction and one lane in the other altering every few kilometers. The lanes in the opposite direction are usually separated by a steel cable barrier. Research has found that 2 + 1 roads are safer than ordinary two-lane roads. The rate of fatal accidents on 2 + 1 roads has been lowered with 76% compared with an ordinary road with similar characteristics (Carlsson 2009).

Table 1. Summary of Qualitative Variables

Pavement type	Count	Proportion	Traffic class (AADT)	Count	Proportion	Road type	Count	Proportion
ABS	17,920	0.071	1, <250	54,762	0.217	Motorway	13,294	0.0527
ABT	73,103	0.29	2, 250–499	41,911	0.166	Undivided motorway	113	0.00014
Seal coat	6,919	0.027	3, 500–999	42,868	0.17	2 + 1 road	3,506	0.0139
Grouted macadam	7,217	0.029	4, 1,000–1,999	39,001	0.155	4-lane road	1,375	0.0054
Semihot mix	53,347	0.211	5, 2,000–3,999	32,132	0.127	Ordinary two-lane road	234,021	0.9275
Cold mix	2,437	0.01	6, 4,000–7,999	21,051	0.083	—	—	—
Surface dressing on gravel	28,998	0.115	7, 8,000–11,999	7,118	0.028	—	—	—
Hot mix	8,498	0.034	8, >12,000	13,466	0.053	—	—	—
Surface dressing	53,870	0.214	—	—	—	—	—	—

Note: ABS = stone mastic; ABT = asphalt concrete.

Table 2. Summary of Additional Qualitative Variables

Maximum stone size (mm)	Count	Proportion	BCC	Count	Proportion	Climate zone	Count	Proportion
<10	131,608	0.522	1	246,483	0.977	South	47,472	0.188
10–14.9	65,913	0.261	2	5,416	0.021	Central	39,547	0.157
15–20	17,552	0.07	3	407	0.002	North	165,290	0.655
>20	22,724	0.09	—	—	—	—	—	—

Table 3. Summary of Quantative Variables

Traffic class	Age (years)				Speed limit (km/h)				Road width (m)			
	Mean	Standard	Minimum	Maximum	Mean	Standard	Minimum	Maximum	Mean	Standard	Minimum	Maximum
1	17.12	10.23	3	58	68.20	10.21	30	120	5.20	1.02	2.0	17.0
2	16.54	10.01	3	60	69.07	12.99	30	100	5.85	0.93	2.5	17.2
3	16.01	9.65	3	59	69.74	14.55	30	100	6.36	1.04	2.0	17.0
4	14.28	8.71	3	51	71.27	15.46	30	110	7.05	1.25	1.5	19.5
5	12.34	7.36	3	52	73.19	15.61	30	110	7.94	1.83	3.0	23.8
6	10.62	6.16	3	52	75.49	15.96	30	120	9.18	2.34	4.0	23.5
7	9.94	5.85	3	52	82.19	20.89	30	120	10.54	2.21	3.6	20.0
8	8.84	5.78	3	51	96.09	18.13	40	120	11.42	1.71	4.0	21.7

Note: Total number of observations 266,614.

The traffic load is reported as AADT: the average number of vehicles per day on a specific road section. The heavy traffic, defined as vehicles weighing over 3,500 kg (3.5 t), is calculated separately. In general, measures of AADT are more accurate for high-traffic roads (in this study, high-traffic roads are defined as having an AADT of more than 4,000) and more uncertain for low-traffic roads (AADT less than 4,000). Heavy-traffic load estimations in the PMS database are uncertain and therefore not included in the model (J. Lang, "Assessment and analysis of road network performance using long-term surface condition data," unpublished internal report, WSP Sweden/Swedish Transport Administration, 2011).

Bearing-capacity classes (BCC) are coded from one to three, where one is the most common class. If there are axle road restrictions due to bearing-capacity limitations, a road can receive classification two or three, where the third-class roads carry the least axle load.

Sweden is a heterogeneous country in terms of climate. The impact of ground frost and studded tires varies both between and within counties. The influence of climate on road pavement lifetimes depends partly on the climate and also on how much consideration the climate has been given in pavement design. The Swedish Transport Administration uses several climate zone classifications in its database, some according to geographical coordinates and some according to a simpler administrative classification based on counties. For convenience, the latter representation

is used in this study. Sweden is divided into three climate zones: north, central, and south.

The nine most common pavement types in Sweden are asphalt concrete, stone mastic, seal coat, grouted macadam, semihot mix, cold mix, hot mix, surface dressing on gravel, and surface dressing. Asphalt concrete is used on all types of roads and in all traffic classes. Stone mastic is the most expensive material of the nine, and it is used only on high-traffic roads. Seal coat is a preventive maintenance designed to increase the remaining life of low-traffic roads by some years until the roads need more thorough maintenance. Grouted macadam is a maintenance for low-traffic roads. It is sometimes used as a preventive maintenance to cover the edges of the road but can also be used to cover the entire road surface. Semihot mix is applied on roads with an AADT of less than 2,000 vehicles. It is a softer, dynamic pavement mix that sometimes can repair small cracks simply as a result of its own movement. Cold mix is harder than semihot mix and mostly found on roads with AADT of 500–2,000 vehicles. The advantage of cold mix is that it does not have to be strongly heated before paving. Surface dressing on gravel is basically paved gravel roads. The road is prepared and covered with a thin layer of asphalt. These roads are always low-traffic roads. Hot mix is used on high-traffic roads and includes all hot mixes that are not asphalt concrete or stone mastic. Many of these materials are older and not in use any longer. Surface dressing is mainly used on low-traffic roads. It is a rough pavement that is smoothed out over time.

04014056-3

Time-to-Event Analysis

As described by Klein and Moeschberger (2005), the basic concepts of time-to-event analysis are as follows: first, consider the time to some event X as a continuous nonnegative random variable from a homogeneous population. Henceforth, capital X denotes a random variable, and lowercase x denotes fixed values.

The survival function $S(x)$ is the probability of an item surviving to time x (i.e., that the event of interest has not occured at time x). If an item has not experienced the event, it is denoted as right censored. Only right censoring is considered in this study.

Since $S(x) = Pr(X > x)$ it is also, by definition, equal to $S(x) = 1 - F(x) = \int_x^\infty f(t)dt$, where $F(x)$ is the cumulative distribution function. Consequently, the probability density function in terms of the survival function is

$$f(x) = \frac{-dS(x)}{dx}$$

The hazard rate function $h(x)$ is the chance that an item of age x experiences the event of interest in the next instant in time.

The hazard rate $h(x)$ is a limit defined as

$$h(x) = \lim_{\Delta x \to 0} \frac{Pr(x \leq X < x + \Delta x | X \geq x)}{\Delta x}$$

which can be interpreted in the way that $h(x)\Delta x$ is the approximate probability of an item of age x experiencing the event of interest in the next instant. An increasing hazard rate can be understood as natural aging (i.e., that the probability of the event increases as the item becomes older). The hazard rate can be expressed in terms of the survival function

$$h(x) = \frac{f(x)}{S(x)} = \frac{-d \ln S(x)}{dx}$$

Cox Proportional Hazards Model

The model of choice in this study is Cox proportional hazards model (Cox 1972). The Cox model is semiparametric and no assumption of a specific distribution of the data is needed. Therefore, the model is robust and the results are not biased by any misspecification.

The term *proportional hazard* refers to the fact that the ratio between the hazards of two individuals with different values of one covariate is constant in the Cox model. Data of sample size n consist of three components $[T_i, \delta_i, Z_i(t)]$, $i = 1, \ldots, n$. Then T_i is the time on study for item i, δ_i is an event indicator for item i ($\delta_i = 1$ if the event has occured, $\delta_i = 0$ if right censored), and $Z_i(t) = [Z_{i1}(t), \ldots, Z_{ip}(t)]$ is the vector of p covariates for the ith subject at time t which may affect the survival distribution of X. The values of $Z_{ik}(t)$ may be time dependent.

The hazard rate at time t is modeled directly, given the design matrix $\mathbf{Z}$ including all covariates. It consists of a baseline hazard rate $h_0(t)$ and a regression part $\mathbf{Z}\beta$

$$h(t|\mathbf{Z}) = h_0(t)\exp(\mathbf{Z}\beta)$$

A nonparametric estimate of the baseline hazard function at time t_i on the ith event is

$$\hat{h}_0(t_i) = \frac{d_i}{\sum_{k \in R(t_i)} \exp(\mathbf{Z}_k \hat{\beta})}$$

where d_i = number of events at time t_i, and $R(t_i)$ = set of individuals that could experience the event at time t_i. The $\hat{\beta}$ estimates

are obtained through maximization of a partial likelihood (Hosmer et al. 2008),

For a stratified model having s strata, the hazard rate for each stratum j is given by

$$h_j(t|\mathbf{Z}) = h_{0j}(t)\exp(\mathbf{Z}\beta), \quad j = 1, \ldots, s$$

Fitting the Cox Model

A stratified Cox proportional hazards model is fitted to data using the package "Survival" (Therneau 2013) in the software R, version 2.15.1 (R Core Team 2012). The model has eight strata according to traffic classes stated by the Swedish Transport Administration (Table 1). The stratification is motivated by the nature of road construction. When a road is built or rebuilt, it is adjusted according to the traffic load, which is the single most important factor in determining the thickness of construction needed. The construction defines the type of road, the speed limit, the road width, the pavement type, and the bearing capacity.

A model with eight strata indicates that the baseline hazard rate differs over traffic classes but that the effects of the covariates are the same over strata. In other words, the estimated parameter vector β is constant for all traffic classes, while the effect of traffic is incorporated within the differentiated baseline hazard. Hence, the dependency between traffic and all other covariates will not cause any bias in the β parameters, and the effects of the covariates can be interpreted individually (Klein and Moeschberger 2005).

As previously mentioned, the drawback of the definition of the lifetime of a road being the time between two maintenance activities is that the cause of maintenance is unknown. Gharaibeh and Darter (2003) mention that roads are usually maintained due to poor condition but that this condition varies from section to section. Some sections might not yet have been maintained even though their condition is poor. To improve the analysis and obtain more accurate estimates of expected lifetimes, information on road surface measurements is added to the model. Rut depth and IRI have been recorded since 1987 in Sweden. The Swedish Transport Administration has developed a maintenance standard stating appropriate levels of rut depth and IRI (Trafikverket 2012). The maintenance standard is a recommendation and not a necessity. All road sections in the material where a measurement of surface condition is available (about 170,000 out of 266,614 road sections in total) were compared with the limits in the maintenance standard. If the measurements of rut depth or IRI of a censored section exceeded recommended values, it was labeled as "ought to have been maintained" and considered as noncensored in the analysis.

Results

The results of the estimation are presented as hazard ratios. A hazard ratio is the relative risk that a road with a certain variable value needs maintenance compared with a chosen reference. A hazard ratio above one is equivalent to a shorter lifetime than the reference category, and a hazard ratio below one corresponds to a longer lifetime.

Large differences in estimated effects of road pavement types on roads' lifetimes were found (Table 4). Differences among road types are smaller but substantial. The effects of climate zone, bearing capacity, and maximum stone size differ significantly but not as much as the previous two variables. The continuous variables road width and speed limit have less effect on roads' lifetimes, depending on the scale on which they are measured.

Table 4. Maximum Likelihood Estimates of the Cox Proportional Hazards Model with Maintenance and Road Condition as Response Variables (Stratification: Eight Traffic Classes)

Parameter	Parameter estimate	Standard error	Z-value	P-value	Hazard ratio
Asphalt concrete[a]	0	—	—	—	1.00
Stone mastic	−0.453	0.0131	−34.6	<0.0001	0.64
Seal coat	0.676	0.0359	18.8	0.0001	1.97
Grouted macadam	−0.070	0.0599	−1.2	0.243	0.93
Semihot mix	0.397	0.0098	40.5	<0.0001	1.49
Cold mix	0.736	0.0236	31.2	<0.0001	2.09
Surface dressing on gravel	0.653	0.0118	55.5	<0.0001	1.92
Hot mix	0.231	0.0145	15.9	<0.0001	1.26
Surface dressing	0.066	0.0076	8.7	<0.0001	1.07
Ordinary two-lane road	0	—	—	—	1.00
Motorway	−0.099	0.0204	−4.9	<0.0001	0.91
Undivided motorway	0.046	0.1107	0.4	0.679	1.05
2 + 1 road	0.198	0.0268	7.4	<0.0001	1.22
Four-lane road	−0.563	0.0358	−15.7	<0.0001	0.57
Climate zone central[a]	0	—	—	—	1.00
Climate zone north	−0.122	0.0089	−13.7	<0.0001	0.89
Climate zone south	0.002	0.0073	0.3	0.796	1.00
Bearing capacity Class 1[a]	0	—	—	—	1.00
Bearing capacity Class 2	0.088	0.0183	4.8	<0.0001	1.09
Bearing capacity Class 3	−0.217	0.0707	−3.1	0.002	0.81
Maximum stone size <10 mm[a]	0	—	—	—	1.00
Maximum stone size 10–14.9 mm	−0.058	0.0114	−5.1	<0.0001	0.94
Maximum stone size 15–19.9 mm	−0.140	0.0116	−12.1	<0.0001	0.87
Maximum stone size ≥20 mm	0.046	0.0139	3.3	0.001	1.05
Road width (dm)	−0.004	0.00012	−21.1	<0.0001	0.996
Speed limit (km/h)	0.006	0.00012	29.0	<0.0001	1.006

Note: Number of observations used in analysis: 220,699.
[a]Reference category of each variable.

Effect of Including Road Condition to Define a Road's Lifetime

The definition of a road's lifetime in this model is either the time between two maintenance activities, or if the road has no second activity registered, the time between the first maintenance activity and the time when measurements of rut depth or IRI exceed the limits in the maintenance standard.

In particular, the inclusion of road condition as a response variable made a difference to 2 + 1 roads. The oldest 2 + 1 roads in the data are 13 years old, and many observations registered for this road type have censored lifetimes. When the surface condition of these roads was compared with the maintenance standard, censoring decreased from 57.3 to 52.1%. In total, 4.2% of the

2 + 1 road sections that have not been maintained had measurements of IRI and rut depth which exceeded the recommended limits in the maintenance standard.

A comparison of estimated median lifetimes between a model with and without road condition included as a response variable has been made in Table 5. With road condition included, the median lifetime of ordinary roads with cable barriers is 1–2 years shorter for all traffic classes except Class 8.

Survival Curves

The expected median lifetime is found where the survival proportion equals 0.5 (Fig. 1). Each of the eight traffic class strata has a corresponding survival curve. It is clear from the figure that low-traffic roads (traffic Classes 1–4) have longer expected lifetimes than high-traffic roads (traffic Classes 5–8).

Table 5. Comparison of Median Lifetimes for 2 + 1 Roads

	Model with maintenance activity as response variable			Model with maintenance activity and road condition as response variable		
Traffic class	Median lifetime	0.95 LCL[a]	0.95 UCL[b]	Median lifetime	0.95 LCL[a]	0.95 UCL[b]
1	28	27	30	26	25	27
2	22	21	23	20	17	21
3	18	18	19	17	16	18
4	15	14	15	14	13	14
5	12	12	12	11	11	12
6	10	9	10	9	9	9
7	9	8	9	8	8	8
8	7	7	7	7	6	7

Note: Covariate values as in Table 6.
[a]LCL = lower confidence limit.
[b]UCL = upper confidence limit.

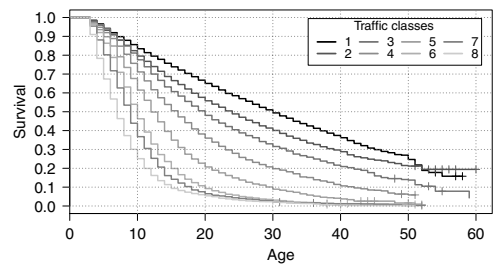**Fig. 1.** Fitted survival curves from Cox proportional hazards model

Table 6. Covariate Values in Estimated Survival Curves

Variable	Value
Pavement type	Asphalt concrete
Road type	Ordinary two-lane road
Climate zone	Central
Bearing capacity class	1
Road width (m)	7.5
Speed limit (km/h)	80
Maximum stone size (mm)	14

When fitting survival curves from the Cox model, covariate values must be specified (see Table 6 for covariate values represented in Fig. 1). For use in a maintenance planning setting, estimated lifetimes of each specific combination of covariates can be calculated and compared.

Discussion

The extensive database available in the PMS application of the Swedish Traffic Administration makes an excellent foundation for maintenance planning. The problem of censored maintenance intervals can be solved by time-to-event analysis. This study uses the Cox proportional hazards model to estimate lifetimes and analyze the effect of seven different covariates on lifetimes of Swedish road pavements. To improve the accuracy of the estimated lifetimes, road condition (IRI and rut depth) was included in the model as a response variable.

Pavement Types

Asphalt concrete is chosen as a reference category among pavement types because it is common within all traffic classes. Stone mastic has the smallest hazard ratio of 0.64, indicating that the risk of needing maintenance is 36% lower than that of asphalt concrete. Stone mastic is an expensive, high-traffic maintenance activity, and it is in line with expectations for it to have the longest lifetime.

Seal coat has the second highest probability of requiring maintenance in the next instant with a hazard ratio of 1.97 compared with the reference category. Because seal coat is a preventive maintenance to slightly increase the lifetime of a road, this estimation is credible.

Grouted macadam has a hazard ratio of 0.93, but the estimate is not significantly different from asphalt concrete. This is a low-traffic maintenance, sometimes used to cover just the edges of a road and sometimes used to cover the entire road surface. In the first scenario, grouted macadam is only a preventive maintenance to increase the lifetime of a road by some years. If only the edges of a road are remade, it will not affect measurements of IRI and rut depth which are used to determine if an activity has effect or not. In the sensitivity analysis, grouted macadam had a significant hazard ratio of 2.42. This result was expected because of the data management, indicating that maintenance activities affecting edges might have been removed.

Semihot mix has a hazard ratio that is 49% higher than asphalt concrete. Cold mix has the shortest lifetime of all pavement types with a 109% higher hazard ratio than asphalt concrete. Cold mix is mostly present on older roads and not used a lot. Surface dressing on gravel has a hazard ratio that is 1.92 times the hazard ratio of the reference category. This is a low-traffic maintenance, and it is expected to have a fairly short lifetime compared with asphalt concrete. Hot mix is a high-traffic maintenance, and its hazard ratio is 26% lower than asphalt concrete.

Surface dressing is used mainly on low-traffic roads with an AADT of up to 4,000 vehicles. It is a rougher and cheaper material than asphalt concrete, but the risk of needing maintenance is almost the same as asphalt concrete, only 7% higher.

To conclude, pavement types used mostly on high-traffic roads generally have longer expected lifetimes (i.e., smaller hazard ratios) than pavement types for low-traffic roads. Since the estimates assume a situation where all other variables are fixed, this result is natural. When maintenance activities are being planned, all factors that affect a roads' lifetime must be considered simultaneously.

Road Types

The most common road type in Sweden is the ordinary two-lane road; therefore, it is chosen as the reference category. The estimate of undivided motorway is not significant, probably due to the small number of observations belonging to this type (Table 1). Motorways are expected to have a slightly longer life than ordinary two-lane roads with a hazard ratio of 0.91. This means that if motorways carried the same amount of traffic and all other covariates were the same for both road types, motorways would have a longer expected lifetime than ordinary two-lane roads. Swedish roads are generally constructed to carry a higher amount of traffic than they do at present to compensate for an increased future traffic load. The hazard ratio of motorways indicates that they are constructed to carry a greater increase in traffic load than ordinary two-lane roads.

Studies by the Swedish National Road and Transport Research Institute (VTI) have shown that the development of ruts on $2 + 1$ roads increases at about twice the rate as ruts on ordinary two-lane roads (Carlsson 2009; Karim et al. 2011). If $2 + 1$ roads are maintained according to the maintenance standard, they should have shorter lifetimes than ordinary two-lane roads. This assumption is confirmed by the Cox model estimates. In a model that considers maintenance activities as only a response variable, the hazard ratio of $2 + 1$ roads is 1.11 (p-value 0.001). When road condition is included as a response variable, the hazard ratio of $2 + 1$ roads increased to 1.22, indicating a 22% higher risk of needing maintenance than ordinary two-lane roads. This result indicates that the safety of the $2 + 1$ roads has a cost in terms of shorter lifetimes, which in turn is likely to give higher maintenance costs over a road's life cycle.

Many aspects are similar between four-lane roads and motorways. The difference is that four-lane roads lack one or more factors needed for a motorway classification, such as restrictions for verges, junctions, speed limits, etc., that are not fulfilled. The hazard ratio is 0.57, which indicates a longer lifetime than ordinary two-lane roads. The same logic of construction as with motorways applies to four-lane roads.

Climate Zones

Climate zone central is chosen as a reference category because it lies between the northern part of Sweden with very cold winters and the southern part with milder winters. The northern climate zone has a significant hazard ratio of 0.89 which indicates longer lifetimes for roads in the north. This result is somewhat unexpected since ground frost and the use of studded tires are factors expected to have a large impact on road condition. Hence, the northern climate zone with longer winters would be expected to have shorter lifetimes than the southern and central climate zones.

Historically, the southern part of Sweden has received more funding for road maintenance due to more intense traffic (J. Lang, WSP Sweden, personal communication, 2012). However, if maintenance in the northern zone had been neglected due to budget

restrictions, it would have shown in measurements of IRI and rut depth that exceeded the maintenance standard. This is not the case, because the northern zone has a longer expected lifetime even when adjusted for road condition. It should be noted that the climate variable will not truly describe the deterioration pattern in different climate zones, but rather the ability of the road manager to adapt the road construction to the climate (i.e., freeze/thaw cycles, studded tyres, etc.) in that particular zone. The climate is a known factor accounted for when constructing the road. The explanation for the lower hazard ratio in the northern zone might actually be that the road construction is stronger in the north to compensate for the impact of ground.

Bearing-Capacity Class

BCC 1 is by far the most common bearing-capacity class in Sweden. Roads with other restrictions have some kind of weak element (e.g., bridges that do not allow heavy vehicles). Due to the restrictions, heavy traffic will not choose these kinds of roads, and the deterioration can be expected to be slower than for roads with the standard bearing capacity, which will carry heavy vehicles. BCC 2 has a hazard ratio that is 9% higher than BCC 1, indicating that the lifetime of these roads is actually shorter than the lifetime of BCC 1 roads. However, a hazard ratio of 0.81 for BCC 3 roads indicates longer expected lifetimes when this restriction is imposed. The BCC 2 restriction for heavy vehicles does not seem effective from a lifetime perspective.

Maximum Stone Size

A maximum stone size of the pavement material of less than 10 mm is chosen as a reference category. Roads with slightly larger stones have longer lifetimes, with 6% lower hazard ratio for roads with maximum 15-mm stones and 13% lower hazard ratio for roads with maximum 20-mm stones. However, for roads with a maximum stone size of 20 mm or larger, the hazard ratio is 5% higher than for roads with the smallest stone size. It seems as if stone size increases the lifetime only up to a certain level. The estimates of the sensitivity analysis were more consistent with hazard ratios below one (approximately 0.8) for all stone sizes greater than 10 mm. These results imply that stones greater than 20 mm have no effect on IRI and rut depth but might have some effect on slowing the road edge deterioration.

Continuous Variables

The continuous variables are road width and speed limit.

Road width shows an expected increased lifetime the broader the road is with a hazard ratio of 0.99 for each decimeter. This means that if the road width is increased by 1 dm (e.g., from 90 to 91 dm), the risk of needing maintenance is lowered with 1%. A wider road is expected to have a longer lifetime because drivers are known to avoid ruts if they can, which is possible on wider roads.

Speed reduces the lifetime of a road with a hazard ratio of 1.01 for each 10 km/h increase in speed limit in a range from 30 to 120 km/h. If the speed limit is, for example, increased from 50 to 90 km/h, the risk of needing maintenance is expected to increase by $1.01^4 = 1.04$ (i.e., the risk of needing maintenance on a road with speed limit 90 km/h is expected to be 4% higher).

Future Research

A variable highly desired to be included in the model is heavy traffic load or equivalent single-axle load (ESAL). The American Association of State Highway Officials' (AASHO) widely used rule of thumb implies that heavy traffic load has the greatest impact in road deterioration through the fourth power law (AASHO 1962). Heavy traffic is included in the total traffic load used in this paper, but a separation of the measures would be preferred.

In recent years, the Swedish Traffic Administration has stressed the importance of LCC analysis when procuring contracts. Combining estimated lifetimes from the Cox proportional hazards model with cost data could produce an early stage LCC analysis suitable for these types of contracts. Investigating the possibilites of such an analysis is an interesting future research topic.

All homogeneous road sections in this paper are considered as independent. This assumption is made in most time-to-event analyses concerning roads' lifetimes. However, the true relation between sections is that each maintenance can be considered as a repeated measurement—the same geographical location of a road is repeatedly maintained, except for new constructions. There is a possibility that each road section has an individual effect, for example, due to differences in soil conditions, affecting the lifetime. A possible future research prospect is to add a random effect in a frailty model to account for this heterogeneity.

Conclusions

The results from the Cox proportional hazards model showed that stone mastic, asphalt concrete, and surface dressing had the longest expected lifetimes among pavement types. Seal coat and cold mix had the shortest expected lifetimes. These results are in line with theory. Further research on life-cycle costs of the different pavement types would be interesting, especially because surface dressing, which is much cheaper than asphalt concrete, had an almost equal hazard ratio.

In terms of road types, it was shown that the inclusion of road condition made the estimated lifetimes of $2 + 1$ roads 1–3 years shorter than if road condition was not included in the model. This indicates that there are some $2 + 1$ roads in Sweden that have not been maintained despite their being in a poor condition according to the maintenance standard. Ordinary roads with cable barriers had 30% higher risk of needing maintenance than ordinary two-lane roads. Hence, if the Swedish Traffic Administration continues to rebuild two-lane roads into $2 + 1$ roads, the need for road maintenance will increase.

Climate zone north had the longest expected lifetime, which was slightly unexpected because of harder winters in the north. A more fine-tuned climate zone categorization might be preferred in future research.

A bearing-capacity limitation of Class 3 was shown to increase roads' lifetimes. Maximum stone size had a somewhat ambiguous effect: for stones up to 20 mm an increased stone size also increased lifetimes, but stones over 20 mm gave a decrease in lifetimes. Increased road width and decreased speed limit proved to increase lifetimes.

References

American Association of State Highway Officials (AASHO). (1962). "The AASHO road test." *Rep. No. 1012*, Washington, DC.

Carlsson, A. (2009). "Evaluation of $2 + 1$ roads with cable barrier." *Rep. No. 636A*, VTI Rapport, Linköping, Sweden.

Cox, D. (1972). "Regression models and life-tables." *J. R. Stat. Soc.*, 34(2), 187–220.

Do, M. (2011). "Comparative analysis on mean life reliability with functionally classified pavement sections." *KSCE J. Civ. Eng.*, 15(2), 261–270.

Dong, Q. (2011). "Enhancement of pavement maintenance decision making by evaluating the effectiveness of pavement maintenance treatments." Ph.D. thesis, Univ. of Tennessee, Knoxville, TN.

Eltahan, A., Daleiden, J., and Simpson, A. (1999). "Effectiveness of maintenance treatments of flexible pavements." *Transportation Research Record 1680*, Transportation Research Board, Washington, DC, 18–25.

Gharaibeh, N., and Darter, M. (2003). "Probabilistic analysis of highway pavement life for Illinois." *Transportation Research Record 1823*, Transportation Research Board, Washington, DC, 111–120.

Gustafsson, M., and Lundberg, T. (2009). "Prognosmodeller för tillståndsmått i trafikverkets pavement management systems." *Rep. No. 765*, VTI Rapport, Linköping, Sweden.

Hall, K., Lee, Y., Darter, M., and Lippert, D. (1994). "Forecasting pavement rehabilitation needs for Illinois interstate highway system." *Transport Research Record 1455, 73d Annual Meeting*, Transportation Research Board, Washington, DC, 116–122.

Hosmer, D., Lemeshow, S., and May, S. (2008). *Applied survival analysis: Regression modeling of time-to-event data*, 2nd Ed., Wiley, Hoboken, NJ.

Karim, H., Alam, M., and Magnusson, R. (2011). "Road barrier repair costs and influencing factors." *J. Transp. Eng.*, 10.1061/(ASCE)TE.1943 -5436.0000227, 349–359.

Klein, J., and Moeschberger, M. (2005). *Survival analysis—Techniques for censored and truncated data*, 2nd Ed., Springer, New York.

R Core Team. (2012). *R: A language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna, Austria.

Smith, K., Titus-Glover, L., Darter, M., Quintus, H. V., Substad, R., and Hallin, J. (2005). "Evaluation of the cost benefits of continuous pavement preservation design strategies versus reconstruction: Final Rep. 491." *Rep. No. FHWA-AZ-05-491*, Arizona Dept. of Transportation, Phoenix, AZ.

Smith, K., Titus-Glover, L., Rao, S., Quintus, H. V., and Stanley, M. (2006). "Life-cycle cost analysis of SMA pavements and SMA application guidelines." *Rep. No. WHRP 06-11*, Wisconsin Highway Research Program, Madison, WI.

Therneau, T. M. (2013). "A package for survival analysis." *S. R package version 2.37-4.*

Trafikverket. (2012). "Underhållsstandard belagd väg 2011." *Rep. No. 2012:074*, Trafikverket publikationer, Borlänge, Sweden.

Wang, Y., Mahboub, K., and Hancher, D. (2005). "Survival analysis of fatigue cracking for flexible pavements based on long-term pavement performance data." *J. Transp. Eng.*, 10.1061/(ASCE)0733-947X(2005) 131:8(608), 608–616.

Paper II

Evaluating Needs of Road Maintenance in Sweden with the Mixed Proportional Hazards Model

Kristin Svenson, Yujiao Li, Zuzana Macuchova, and Lars Rönnegård

National road databases often lack important information for long-term maintenance planning of paved roads. In the Swedish case, latent variables of which there are no recordings in the pavement management systems database are, for example, underlying road construction, subsoil conditions, and amount of heavy traffic measured by the equivalent single-axle load. The mixed proportional hazards model with random effects was used to capture the effect of these latent variables on a road's risk of needing maintenance. Estimation of random effects makes it possible to identify sections that have shorter or longer lifetimes than could be expected from the observed explanatory variables (traffic load, pavement type, road type, climate zone, road width, speed limit, and bearing capacity restrictions). The results indicate that the mixed proportional hazards model is useful for maintenance planning because the weakest and strongest sections in a road network can be identified. The effect of the latent variables was visualized by plotting the random effect of each section in a map of the road network. In addition, the spatial correlation between road sections was evaluated by fitting the random effects in an intrinsic conditional autoregressive model. The spatial correlation was estimated to explain 17% of the variation in lifetimes of roads that occur because of the latent variables. The Swedish example shows that the mixed proportional hazards and intrinsic conditional autoregressive models are suitable for analyzing the effect of latent variables in national road databases.

Long-term maintenance planning of paved roads requires reliable estimates of maintenance intervals. These estimates can be used in a variety of settings, such as input for life-cycle cost analyses and calculations of the marginal costs of road maintenance, but also to evaluate the future need of maintenance of a road network.

Time-to-event analysis is widely used to estimate the lifetimes of paved roads, where "lifetime" is commonly defined as the time between two maintenance treatments [see, e.g., Do (1), Smith and colleagues (2, 3), Gharaibeh and Darter (4), Eltahan et al. (5), and Hall et al. (6)]. Several authors, such as Svenson (7), Dong (8), and Wang et al. (9), have evaluated the effect on the lifetimes of roads of different covariates, such as pavement type, road type, climate or weather, road width, speed limit, and so forth. However, few previous studies have accounted for the spatial nature of the road network when modeling the lifetimes of roads. In practice, roads are seldom maintained inde-

School of Technology and Business Studies, Dalarna University, 79188 Falun, Sweden. Corresponding author: K. Svenson, kss@du.se.

pendent of each other. If a road section needs maintenance because of ruts or cracks, for example, neighboring sections that still are functional might be more likely to need maintenance in the near future. This scenario may lead to opportunistic maintenance, which could imply higher maintenance costs than necessary.

Larger data sets containing road maintenance information also exhibit repeated observations. Svenson used data from the Swedish pavement management system (PMS), which covers all national roads in Sweden and their maintenance activities from the 1960s onward (7). Certain characteristics of a road section may change between maintenance activities, such as pavement type or speed limit, but the geographical location of the road is still the same. The PMS database only contains information about the surface pavement layer. Even if the time-to-event model includes several explanatory variables, the effect on the road's lifetime of the underlying construction layers or ground materials (solid bedrock or looser sand) cannot be captured.

To account for heterogeneous construction designs, or other latent variables that cannot be captured by the explanatory variables in the model, a random effect for each road section can be added to the time-to-event analysis. Such models are known in the literature as survival frailty models or mixed proportional hazards models (10, 11). The random effect represents an extra frailty so that roads with a stronger underlying design are more likely to have longer lifetimes (i.e., less risk of needing maintenance) than roads with weaker designs. The mixed proportional hazards model has been frequently used in medical settings, such as infant mortality (12) and breast cancer (13, 14). In medical research, the aim is to identify individuals who are frailer than others, which is analogous to identifying weaker road sections. Parametric frailty models (where the hazard function is parametric rather than semiparametric as in the case of the mixed proportional hazards model) have been used to identify factors critical to spatio-temporal congestion impacts of motorway accidents (15) and to model fatigue cracking in pavements (16).

Methods for identifying spatial correlations are variograms and spatial regression models. Gagnon et al. have used variograms for estimating variability in pavement deterioration parameters (17). In recent studies, Lea and Harvey have used variograms and spatial regression models to model the spatial correlation of different pavement properties (18, 19). This paper uses the intrinsic conditional autoregressive (ICAR) model, a spatial regression model that has proved useful for large continuous spatial data sets [e.g., when modeling the occupancy of caribou (20) and the geographical variation of emergency department visits (21)].

The aim of this study was to identify sections in the Swedish paved road network whose lifetimes are shorter, because of latent variables,

than roads with similar characteristics with a mixed proportional hazards model. Possible spatial correlation between lifetimes of road sections is also investigated.

MATERIALS AND METHODS

Mixed Proportional Hazards Model

The mixed proportional hazards model is an extension of the classic Cox proportional hazards model (22). The model consists of two parts: a nonparametric baseline hazard whose distribution is unspecified, giving the model high flexibility, and a parametric part, which accounts for the effect of the covariates. The response variable of the mixed proportional hazards model is the hazard rate, which in the setting of road maintenance is interpreted as "the risk of needing maintenance in the next instant in time." The random effects (i.e., frailties) are added to the parametric part and are assumed to follow some distribution. Let $h_i(t)$ be the hazard rate for observation i at time t, X_i the design matrix of the fixed effects containing all explanatory variables, and Z_i the design matrix of the random effects, constructed by setting 1 for each unique road section and 0 otherwise:

$$h_i(t) = h_{0j}(t)\exp[X_i\beta + Z_i b] \tag{1}$$

where $h_{0j}(t)$ is the baseline hazard stratified over j strata, β is a vector of fixed effects, and the vector of random effects b follows a normal distribution $b \sim N[0, \Sigma(\theta)]$, where the variance–covariance matrix Σ depends on the parameter θ. In this paper the random effects are assumed to be identically and individually distributed, and the variance–covariance matrix $\Sigma(\theta) = \sigma_b^2 I$. The fixed and random effects in the mixed proportional hazards model are usually interpreted with respect to hazard ratios, which is the relative risk of needing maintenance in the next instant in time compared with a reference category.

Penalized Likelihood Approach

Following the work of Ripatti and Palmgren (11) and Therneau et al. (10), the penalized likelihood approach was used to estimate the mixed proportional hazards model in Equation 1. The likelihood approach does not depend on any prior assumptions about the distribution of the parameters. For fixed values of β and b, the partial log likelihood, $\log[\mathrm{PL}(\beta, b)]$, is

$$\log[\mathrm{PL}(\beta, b)] = \sum_{i=1}^{n} \int_{0}^{\infty} \left[Y_i(t)\eta_i - \log\left(\sum_j Y_j(t)e^{\eta_j}\right) \right] \partial N_i(t) \tag{2}$$

where

$\eta_i = X_i\beta + Z_i b$,

$Y_i(t)$ = risk set with $Y_i(t) = 1$ if road section i has been maintained at time t and 0 otherwise, and

$N_i(t)$ = number of observed events in $[0, t]$ for observation i.

Since the event of interest is maintenance activity and a new observation is created after maintenance occurs, $N_i(t)$ can take a maximum value of 1.

The log penalized partial likelihood function, $\log[\mathrm{PPL}(\beta, b, \theta)]$, adds a penalty to the partial log likelihood, penalizing for the extra parameter b.

$$\log[\mathrm{PPL}(\beta, b, \theta)] = \log[\mathrm{PL}(\beta, b)] - \frac{1}{2}b'\Sigma(\theta)^{-1}b \tag{3}$$

where the last term, $1/2 b'\Sigma(\theta)^{-1}b$, is the penalty. The random effects can be integrated out to obtain an integrated partial likelihood IPL(β, θ):

$$\mathrm{IPL}(\beta, \theta) = \frac{1}{(2\pi)^{\frac{q}{2}}|\Sigma(\theta)|^{\frac{1}{2}}} \int \mathrm{PPL}(\beta, b, \theta)\partial b \tag{4}$$

where q is the length of b (i.e., the number of random effects). The maximum likelihood estimates of β and b are obtained by joint maximization of the IPL over β and θ. When the variance of the random components is zero, the IPL collapses into the ordinary Cox partial likelihood. To estimate b, the following score equation is solved:

$$\frac{\partial\log[\mathrm{PPL}]}{\partial b} = \frac{\partial\log[\mathrm{PL}(\beta, b)]}{\partial b} - \frac{\partial\frac{1}{2}b'^{\Sigma(\theta)^{-1}}b}{\partial b} = 0 \tag{5}$$

The penalized likelihood approach is applied in the R-package coxme written by Therneau et al. (10). Coxme was used to estimate the mixed proportional hazards model in this paper.

Bayesian Approach

The presence of spatial correlation was evaluated with the ICAR model, developed by Besag and Coperberg (23), by using the Bayesian software INLA (see www.r-inla.org) created by Rue et al. (24). The ICAR model should ideally have been applied within the mixed proportional hazards model, but this was not possible because of the computational requirements for the large data set. The spatial correlation was therefore approximated by first fitting independent random effects in the mixed proportional hazards model, and their estimates $\hat{b}$ were subsequently modeled as

$$\hat{b} = \mu + \omega + \varepsilon \tag{6}$$

where μ is an intercept term close to 0, since $E(\hat{b}) = 0$. The spatial random effects are modeled as ICAR [i.e., ω is assumed multivariate normal $\omega \sim N(0, \tau(I - D)^{-1})$, where τ is the spatial variance component, I is the identity matrix, and D is the neighborhood matrix, constructed by setting 1 for neighboring sections and 0 otherwise]. ε is a normally distributed random error, $\varepsilon \sim N(0, \sigma_\varepsilon^2 I)$.

A prior for ε is required by INLA, since it is a Bayesian method. The given prior was $N(0, \hat{\sigma}_b^2)$ where $\hat{\sigma}_b^2$ was the estimated variance component from coxme.

Variogram

A variogram gives a visual indication of the spatial correlation (25). The empirical semivariogram $\gamma(h)$ is calculated as follows:

$$\gamma(h) = \frac{1}{2|N(h)|}\sum_{N(h)}(z_i - z_j)^2 \tag{7}$$

where

$N(h)$ = set of all pairwise Euclidean distances $i - j = h$,

$|N(h)|$ = number of distinct pairs in $N(h)$, and

z_i and z_j = data values (here, road age) at spatial locations $i - j$.

Note that variogram and semivariogram are used interchangeably in this paper. By definition, $\gamma(h)$ is the semivariogram and $2\gamma(h)$ is the variogram.

A fitted variogram can be obtained from the empirical variogram by using a weighted least squares method. The R-package gstat written by Pebesma (26) was used for estimation of the empirical and fitted variograms.

DATA COLLECTION

The data material is provided by the Swedish Transport Administration's PMS. PMS includes a database of all national roads in Sweden represented as homogeneous sections. The first recordings in the database are from the 1960s, but the quality of the database was raised when measures of surface condition were introduced in Sweden in 1987. The sections are homogeneous with respect to pavement type, maintenance date, traffic load, road width, speed limit, and axle road restrictions. Sections vary in length, from more than 1 km to only a few meters. There is also information about the coordinates of these sections, and many sections share the same coordinates but have different maintenance dates. When maintenance is carried out, it is reported in the PMS database and added as a new observation.

The variables included in the study are traffic load, pavement type (the nine most common in Sweden), road type (ordinary two-lane roads, four-lane roads, motorways, undivided motorways, and 2+1 roads), climate zone (north, central, and south), bearing capacity class (where Class 1 implies no restrictions and Class 2 and 3 are roads with axle load restrictions), road width, and speed limit. The 2+1 roads were introduced in Sweden in the 1990s. These are roads that alternate between one and two lanes in each direction, separated by a cable barrier. A more thorough description of the data material can be found in Svenson (7).

In previous work, Svenson restricted the length of each section used in the analysis to at least 50 m (7). To be able to account for any spatial correlation, all sections in the database are used in this paper. To determine the impact of including the shorter sections, a sensitivity analysis was performed with a data set consisting of only sections longer than 50 m. The results of the sensitivity analysis are presented in the next section of this paper. The full data set is preferred to the data set that only includes sections of length 50 m because of the road network's spatial structure. If short sections are removed from the data set, direct neighbors will be excluded and the spatial correlation structure is weakened.

The number of observations (road sections) in the database with coordinates available is 548,091, ranging from 1 to 24,322 m in length. The mean length is 357 m. The number of unique road sections is 229,088, with one to 10 registered maintenance activities per section (Table 1). The type of maintenance activity applied to a section may differ, but all activities include a resurfacing.

RESULTS AND DISCUSSION

Estimation of the Random Effects

A likelihood ratio test of the mixed proportional hazards model versus Cox proportional hazards model gives a χ^2-distributed test statistic of 600.34 on one degree of freedom, which is highly significant (p-value $< .0001$). This test implies that the variance of the random effects is nonzero, and hence the variation in lifetime between sections explained by the latent variables is substantial.

TABLE 1   Registered Maintenance Activities per Road Section

Number of Maintenance Activities	Number of Road Sections	Percentage of Road Sections	Length of Road Sections (km)
1	56,650	24.7	26,623
2	75,325	32.9	30,427
3	55,453	24.2	20,171
4	27,297	11.9	8,372
5	10,443	4.6	2,583
6	2,984	1.3	652
7	772	0.3	159
8	140	0.06	46
9	19	0.008	3
10	5	0.002	0.2

A comparison of the parameter estimates of the explanatory variables in Table 2 between the mixed proportional hazards model and the standard Cox model shows no large differences. All categorical variables have a reference category with a hazard ratio of 1. Having a hazard ratio of less than 1 implies less risk of needing maintenance, and hence a longer expected lifetime than the reference category. The opposite, a higher risk of needing maintenance, applies to hazard ratios above 1. The parameter estimates of the explanatory variables in Table 2 are analyzed in detail in previous published work by Svenson (7).

The model is stratified, based on eight traffic classes on which the Swedish Transport Administration has decided (Table 3). Traffic load is highly correlated with all other explanatory variables because it is the most important factor in all maintenance and reconstruction decisions. The random effects are estimated within every traffic class strata.

The idea of the mixed proportional hazards model is that the random effects should capture latent variables that cannot be accounted for by the explanatory variables. When the random effects are assumed normally distributed with mean zero, the random effect of each road section can be interpreted as the deviation from this zero mean. The estimated variance of the random effects is 0.087 (Table 2), which implies a standard deviation of $\sqrt{0.087} = 0.295$. Expressed as hazard ratios, a random effect that lies one standard deviation from the mean implies a hazard ratio of $\exp(0.295) = 1.34$ (i.e., this road section has an increased risk of needing maintenance of 34%).

A sensitivity analysis was performed by fitting the mixed proportional hazards model to a data set with sections longer than 50 m. The hazard ratios of the fixed effects changed less than 5% by excluding sections shorter than 50 m. The variance of the random effects is slightly lower than for the full data set with $\hat{\sigma}_b^2 = 0.067$. However, the variance is still highly significant according to the likelihood ratio test, which has a χ^2-distributed test statistic of 234.55 on one degree of freedom.

Figure 1 shows the distribution of random effects expressed as hazard ratios, depending on the number of observations per road section. It is clear that the "random effects" are not entirely random but rather depend on the number of observations per section. This phenomenon is known in literature as data with informative cluster sizes. In this particular case, having informative clusters makes sense; road sections that are maintained often (i.e., have

TABLE 2   Penalized and Maximum Likelihood Estimates of Mixed and Cox Proportional Hazards Models

Parameter	Mixed Proportional Hazards Model			Cox Proportional Hazards Model		
	Parameter Estimate	Standard Error	Hazard Ratio	Parameter Estimate	Standard Error	Hazard Ratio
Asphalt concrete[a]	0	—	1.00	0	—	1.00
Stone mastic	−0.447	0.010	0.64	−0.389	0.001	0.68
Surface dressing	0.098	0.006	1.10	0.103	0.006	1.11
Hot mix	0.221	0.012	1.25	0.198	0.011	1.22
Seal coat	0.395	0.037	1.48	0.390	0.013	1.48
Semi-hot mix	0.477	0.007	1.61	0.431	0.007	1.54
Grouted macadam	0.651	0.014	1.92	0.625	0.013	1.87
Cold mix	0.734	0.021	2.08	0.689	0.020	1.99
Surface dressing on gravel	0.737	0.009	2.09	0.678	0.009	1.97
Ordinary 2-lane road[a]	0		1.00	0		1.00
4-lane road	−0.491	0.027	0.61	−0.488	0.025	0.61
Motorway	−0.006	0.017	0.99	−0.027	0.016	0.98
Undivided motorway	0.049	0.088	1.05	0.041	0.081	1.04
2+1 road	0.255	0.021	1.29	0.209	0.020	1.23
Climate zone central[a]	0		1.00	0		1.00
Climate zone north	−0.106	0.007	0.90	−0.096	0.007	0.91
Climate zone south	0.0002	0.006	1.00	−0.006	0.006	0.99
Bearing capacity Class 1[a]	0	—	1.00	0	—	1.00
Bearing capacity Class 2	0.041	0.015	1.04	0.036	0.014	1.04
Bearing capacity Class 3	−0.193	0.050	0.82	−0.187	0.048	0.83
Road width	−0.003	0.0002	0.997	−0.003	0.0001	0.997
Speed limit	0.005	0.0002	1.005	0.005	0.0002	1.005
σ_b^2 (variance of random effects)	0.087	—	—	—	—	—

NOTE: Stratification of baseline hazard: eight traffic classes, defined in Table 3. — = no estimate. Number of observations used in analysis is 339,579.
[a]The reference category of each variable.

more registered maintenance activities and hence more observations per road section) tend to have shorter lifetimes. The random effects of the 25% of the road sections with only one observation lie in the lower range of the distribution. All sections with only one observation are censored, which means that they have never been maintained and the present pavement coating is still in use. Therefore, they will all have a longer final lifetime than their current age.

There is some research about how informative cluster sizes affect maximum likelihood estimation. Neuhaus and McCulloch (27) found that informative cluster sizes can yield a biased estimate of the intercept in linear mixed models and generalized linear mixed models. However, there are currently no studies on how informative cluster

TABLE 3   Definition of Traffic Classes

Class	Average Annual Daily Traffic
1	<250
2	250–499
3	500–999
4	1,000–1,999
5	2,000–3,999
6	4,000–7,999
7	8,000–12,000
8	>12,000

sizes affect the random effects of the mixed proportional hazards model. This research lies out of the scope of the current paper but certainly is a topic for future investigation.

Evaluation of Spatial Correlation

By choosing the middle point of each road section as spatial location, the correlation with respect to road lifetime (age in years of each road section) within a distance of 10 km was calculated and plotted in a semivariogram, according to Equation 4. Along with the empirical variogram represented as dots, a fitted variogram is presented as a solid line (Figure 2).

The nugget of the variogram, estimated from the fitted variogram where the distance $h = 0$, represents the variance not explained by the spatial correlation. If all variation in road age is explained by the spatial correlation, and the distance between points is measured correctly, the nugget would be zero. The nugget is estimated to 70 (i.e., the variation not explained by a spatial correlation pattern is high). The range lies around 4 km, which means that the spatial correlation of the maintenance intervals is limited to the closest neighborhood, which makes theoretical sense: roads 5 km apart can be very different in construction and subsoil conditions. Also, budget restrictions will not permit maintenance to stretch too far from the area in question.

The amount of spatial correlation between the random effects was estimated by running the ICAR model defined earlier in the paper. The response variable is the random effects estimated from the mixed proportional hazards model. The prior given for the random
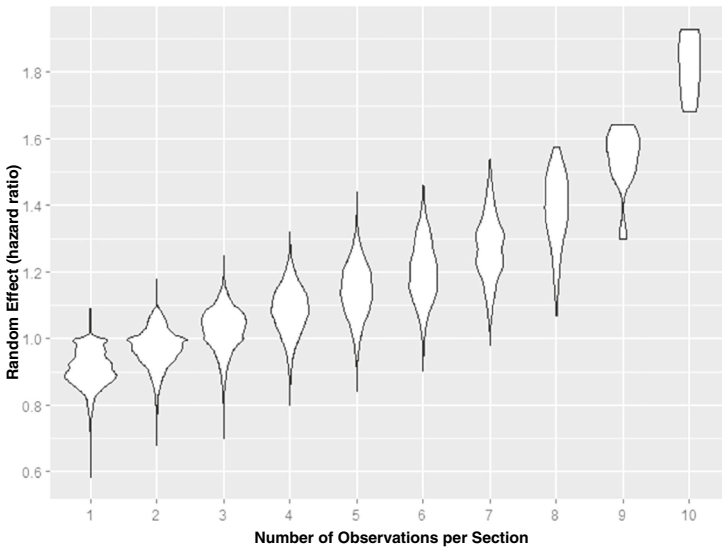
FIGURE 1   Violin plot of distribution of random effects with respect to number of maintenance activities per road section.
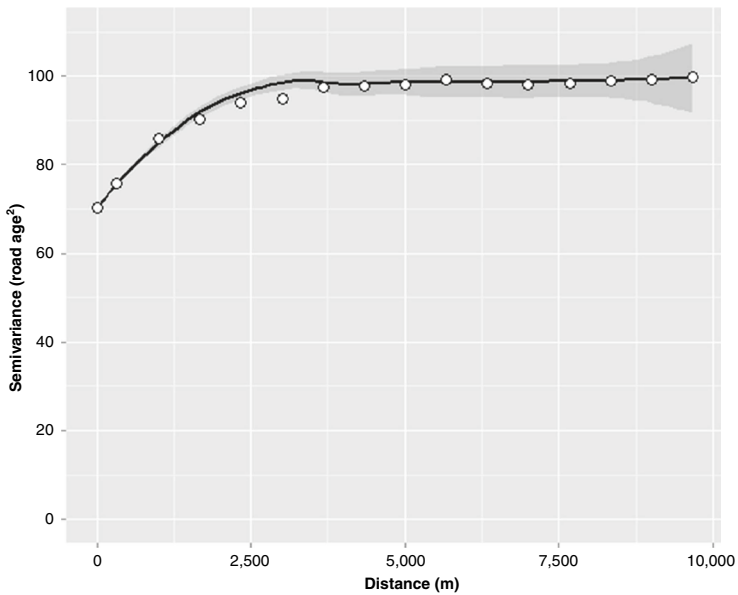

FIGURE 2   Variogram of road age in Swedish pavement management systems database (cutoff 10 km).

error is $N(0, 0.087)$, where the mean is zero by definition and the variance is estimated from the mixed proportional hazards model (Table 2). From the ICAR model, the intraclass correlation of the spatial component can be calculated to indicate how much of the total variation between the random effects exists because of their spatial correlation.

Only a subset of Swedish roads could be used to fit the linear mixed model because of the data size. Sweden has six maintenance regions and, out of these regions, Mitt was chosen, consisting of 30,683 unique road sections in the middle of the country. The intraclass correlation was estimated to 17% (i.e., spatial correlation is present but does not seem to be the crucial source of unexplained variation in lifetime between road sections). Factors such as road construction and heavy traffic are likely to hold more explanatory power.

Map of Random Effects

All road construction and maintenance work is performed to manage the present traffic load but also expected future increases in traffic and future changes in climatic factors. However, the estimated average lifetime of a Swedish road is 7 to 26 years, depending on traffic load (7). This expectation covers a long time span, and it is impossible to be completely accurate about future traffic increase or climate changes. If a road is maintained according to the actual traffic increase, the hazard ratio of the random effect will lie close to 1. If the traffic increase deviates from expectation, the hazard ratio of the random effect will be less than 1 if a road has a lower increase in traffic than expected and more than 1 if it has a higher increase in traffic than expected.

Another possible latent variable in the model is subsoil conditions. Some road sections lie on solid bedrock, while other roads are built on more flexible ground. If road construction work is done right, according to subsoil conditions, this might not be an issue, but if it is not, subsoil conditions could affect the lifetime of a particular road section.

Apart from road construction and subsoil conditions, the random effect could also represent the effect on the lifetime of roads for roads with a large share of heavy traffic. Heavy traffic is only registered in the Swedish PMS database as the share of annual average daily traffic (AADT) weighing more than 3,500 kg, but the figures on heavy traffic are very uncertain for low-traffic roads (AADT less than 4,000 vehicles). There is no information about actual vehicle weight or equivalent single-axle load (ESAL). There is a possibility that roads with low AADT have a large share of heavy vehicles (mainly roads in counties with a large logging industry), which in theory should affect the risk of maintenance for these roads. That is, a higher hazard ratio is expected for roads where a greater proportion of the total traffic consists of heavy vehicles, compared with roads with similar characteristics but a lesser proportion of heavy vehicles.

The variables that should be regarded as latent may vary depending on the data available for inclusion as explanatory variables. Geological data could possibly contribute to information about subsoil condition but do not guarantee that the road construction is built accordingly. ESAL can be estimated from the share of heavy traffic but will never be completely accurate since the data collection about heavy traffic in Sweden is incomplete.

Figure 3 shows the random effects for each road section in the Swedish county Dalarna, which is part of region Mitt. Light gray lines represent roads that are not part of the national road network or roads where information is missing in the database. The green lines represent roads with random effects, where the hazard ratio lies between 0.72 (the lowest estimated random effect in Dalarna) and 0.95, indicating that these roads have a longer lifetime than the average road with the same explanatory variables. Yellow lines represent roads with an average expected lifetime (i.e., a hazard ratio close to one), and the red lines are the roads with a random effect hazard ratio of 1.05 up to 1.36 (the highest estimated random effect in Dalarna). These red roads have 5% to 36% higher risk of needing maintenance than the average road with corresponding explanatory variables. The division of the random effects is arbitrarily chosen so that the yellow lines represent 35% of all road sections, green lines represent 32%, and red lines represent 33%.

Two patterns are clearly visible in the map. First, there are longer routes of red lines where several connected road sections have a hazard ratio of 1.05 or more. One example is the southern road along Lake Siljan, County Road 938, between Leksand and Mora, which is a rather small road (width 5.8 to 6.5 m) on which both passenger traffic and heavy logging trucks travel. The proportion of AADT consisting of heavy traffic is around 6% on Road 938, but the actual ESAL is unknown. A likely explanation for the increased risk of the need for maintenance of this road could be that the logging trucks have an ESAL that is higher than the amount for which the road was intended. It is also possible that the road construction is too weak to match the current amount of traffic (AADT is about 2,000 to 3,000 vehicles).

The second pattern on the map is the presence of small red "dots" (i.e., one or just a few sections with a higher risk of needing maintenance than neighboring sections). These sections could be where a local deterioration in subsoil conditions is not accounted for by the road construction. In general, the road is good but a few sections need extra maintenance.

CONCLUSION

The mixed proportional hazards model works well for identifying road sections with shorter lifetimes than corresponding roads with equivalent explanatory variables. In the example county of Dalarna, both adjacent and single sections with a higher risk of needing maintenance could be identified.

A spatial correlation between Swedish roads with respect to lifetime exists according to the variogram up to a distance of about 4 km. The spatial correlation explained about 17% of the variation between the random effects. Most of the variation in lifetime between roads that is not explained by the explanatory variables seems to originate from other sources than the spatial correlation between road sections. Subsoil conditions, road construction, and actual ESAL are identified as possible latent variables accounting for this variation.

The fixed and the random effects estimated from the mixed proportional hazards model can be used for different purposes: the fixed part of the model is useful for maintenance planning in general or as an input in, for example, life-cycle cost analyses. The random effects are useful for identifying existing sections that perform worse (or better) than the expected average section. The random effects can give the maintenance planning manager a guideline on which sections require a closer inspection. If maintenance activities are
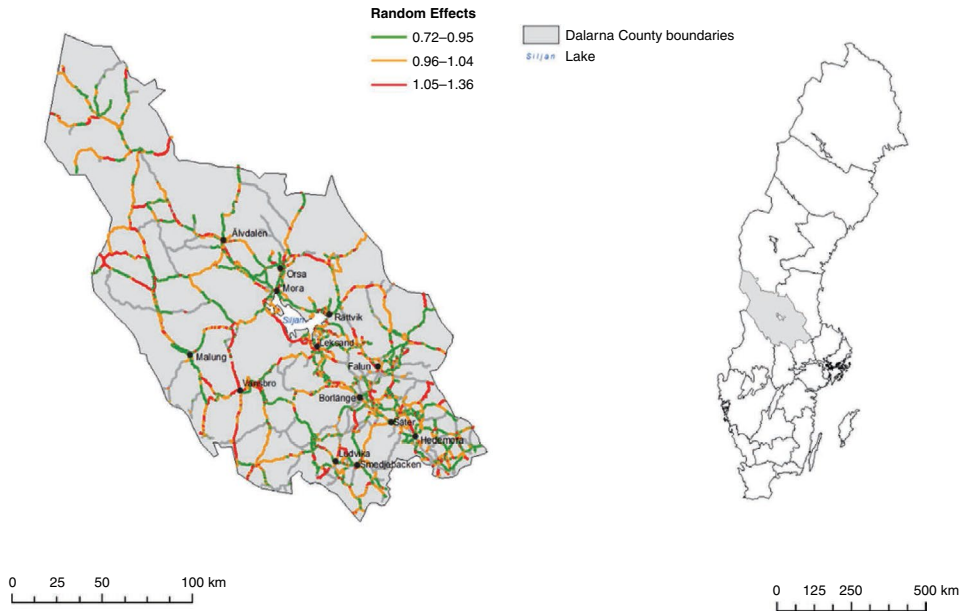
FIGURE 3 Fitted random effects on hazard ratio scale in road network of Dalarna County, Sweden.

planned perfectly according to the future traffic load, the individual variation between sections would cancel out, and the fixed effects could possibly capture all major sources of variation in road lifetime. As long as the variance of the random effects is significant, there are sections for which maintenance does not correspond to the actual need, and also sections for which the maintenance activities carried out are more extensive than the actual need. The random effects could therefore be used to lessen the opportunistic maintenance and make maintenance decisions and the use of limited resources more effective.

REFERENCES

1. Do, M. Comparative Analysis on Mean Life Reliability with Functionally Classified Pavement Sections. *KSCE Journal of Civil Engineering,* Vol. 15, No. 2, 2011, pp. 261–270.
2. Smith, K. L., L. Titus-Glover, S. Rao, H. L. von Quintus, and M. Stanley. *Life-Cycle Cost Analysis of SMA Pavements and SMA Application Guidelines.* Report WHRP 06-11. Wisconsin Highway Research Program, University of Wisconsin, Madison, 2006.
3. Smith, K. L., L. Titus-Glover, M. I. Darter, H. L. von Quintus, R. N. Stubstad, and J. P. Hallin. *Evaluation of the Cost Benefits of Continuous Pavement Preservation Design Strategies Versus Reconstruction.* Report FHWA-AZ-05-491. Arizona Department of Transportation, Phoenix, 2005.
4. Gharaibeh, N. G., and M. I. Darter. Probabilistic Analysis of Highway Pavement Life for Illinois. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1853,* Transportation Research Board of the National Academies, Washington, D.C., 2003, pp. 111–120.
5. Eltahan, A. A., J. F. Daleiden, and A. L. Simpson. Effectiveness of Maintenance Treatments of Flexible Pavements. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1680,* TRB, National Research Council, Washington, D.C., 1999, pp. 18–25.
6. Hall, K. T., Y. H. Lee, M. I. Darter, and D. L. Lippert. Forecasting Pavement Rehabilitation Needs for Illinois Interstate Highway System. In *Transportation Research Record 1455,* TRB, National Research Council, Washington, D.C., 1994, pp. 116–122.
7. Svenson, K. Estimated Lifetimes of Road Pavements in Sweden Using Time-to-Event Analysis. *Journal of Transportation Engineering,* Vol. 140, No. 2, 2014, pp. 1–8.
8. Dong, Q. *Enhancement of Pavement Maintenance Decision Making by Evaluating the Effectiveness of Pavement Maintenance Treatments.* PhD thesis. University of Tennessee, Knoxville, 2011.
9. Wang, Y., K. Mahboub, and D. Hancher. Survival Analysis of Fatigue Cracking for Flexible Pavements Based on Long-Term Pavement Performance Data. *Journal of Transportation Engineering,* Vol. 131, No. 8, 2005, pp. 608–616.
10. Therneau, T. M., P. M. Gambsch, and V. S. Pankratz. Penalized Survival Models and Frailty. *Journal of Computational and Graphical Statistics,* Vol. 12, 2003, pp. 156–175.
11. Ripatti, S., and J. Palmgren. Estimation of Multivariate Frailty Models Using Penalized Partial Likelihood. *Biometrics,* Vol. 56, 2000, pp. 1016–1022.
12. Banerjee, S., M. M. Wall, and B. P. Carlin. Frailty Modeling for Spatially Correlated Survival Data, with Application to Infant Mortality in Minnesota. *Biostatistics,* Vol. 4, 2003, pp. 123–142.
13. Pruitt, S. L., J. A. Tiro, J. M. Ruiz, and S. Inrig. Residential Racial Segregation and Mortality Among Black, White, Hispanic Urban Breast Cancer Patients in Texas, 1995 to 2009. *Cancer,* Vol. 121, No. 11, 2015, pp. 1845–1855.

14. Faradmal, J., A. Talebi, A. Rezaianzadeh, and H. Mahjub. Survival Analysis of Breast Cancer Patients Using Cox and Frailty Models. *Journal of Research in Health Sciences,* Vol. 12, No. 2, 2012, pp. 127–130.
15. Chung, Y., and W. W. Recker. Congested Impact Information on Freeway Accidents. *Intelligent Transportation Systems,* Vol. 16, No. 4, 2015, pp. 2104–2211.
16. Gao, L., J. Aguiar-Moya, and Z. Zhang. Bayesian Analysis of Heterogeneity in Modeling of Pavement Fatigue Cracking. *Journal of Computing in Civil Engineering,* Vol. 26, No. 1, 2012, pp. 37–43.
17. Gagnon, D., G. Dore, and D. Perraton. A Study of Pavement Parameters Variability. Presented at 5th Symposium on Pavement Surface Characteristics of Roads and Airport, Toronto, Ontario, Canada, 2004.
18. Lea, J.D., and J.T. Harvey. Using Spatial Statistics to Characterize Pavement Properties. *International Journal of Pavement Engineering,* Vol. 16, 2015, pp. 239–255.
19. Lea, J.D., and J.T. Harvey. A Spatial Analysis of Pavement Variability. *International Journal of Pavement Engineering,* Vol. 16, 2015, pp. 256–267.
20. Johnson, D.S., P.B. Conn, M.B. Hooten, J.C. Ray, and B.A. Pond. Spatial Occupancy Models for Large Data Sets. *Ecology,* Vol. 94, 2013, pp. 801–808.
21. Neelon, B., P. Ghosh, and F.P. Loebs. A Spatial Poisson Hurdle Model for Exploring Geographic Variation in Emergency Department Visits. *Journal of the Royal Statistical Society: Series A,* Vol. 175, 2012, pp. 1–25.
22. Cox, D.R. Regression Models and Life-Tables. *Journal of the Royal Statistical Society,* Vol. 34, No. 2, 1972, pp. 187–220.
23. Besag, J., and C. Coperberg. On Conditional and Intrinsic Autoregressions. *Biometrika,* Vol. 82, 1995, pp. 733–746.
24. Rue, H., S. Martino, and N. Chopin. Approximate Bayesian Inference for Latent Gaussian Models Using Integrated Nested Laplace Approximations (with Discussion). *Journal of the Royal Statistical Society,* Vol. 71, 2009, pp. 319–392.
25. Cressie, N.A.C. *Statistics for Spatial Data.* Wiley & Sons, Inc, New York, 1993.
26. Pebesma, E.J. Multivariable Geostatistics in S: The gstat Package. *Computers and Geosciences,* Vol. 30, 2004, pp. 683–691.
27. Neuhaus, J.M., and C.E. McCulloch. Estimation of Covariate Effects in Generalized Linear Mixed Models with Informative Cluster Sizes. *Biometrika,* Vol. 98, 2011, pp. 147–162.

Paper III

1

Estimating the marginal costs of road wear[1]

Jan-Eric Nilsson,
Swedish National Road and Transport Research Institute
jan-eric.nilsson@vti.se
Kristin Svenson,
Dalarna University
kss@du.se
Mattias Haraldsson,
Swedish National Road and Transport Research Institute
mattias.haraldsson@vti.se

Abstract: This paper sets out to assess the marginal cost for using road infrastructure using a large set of data with information about sections of the road network, including age, pavement type, traffic, etc. The paper suggests a strategy for identifying major differences in marginal costs across the road network. The analysis provides evidence for that not only heavy vehicles but also cars contribute to road quality deterioration. The hypothesis is that this is due to the widespread use of studded tires in countries with regular freeze-thaw cycles. No indication of deterioration due to time per se is however found.

Keywords: Marginal costs, wear and tear, road reinvestment, Weibull model.

2

1 Introduction

In order to maximize the social welfare from resources expended on infrastructure, two aspects are in focus. The first is to build new roads, bridges etc. once the aggregate benefits of an investment exceed the resources allocated to construction; the second is to charge for the use of existing assets according to the social marginal costs emanating from their use. The focus of the present paper is on the second issue. Perry and Small (2005) demonstrate the overall scope and significance of the task.

To implement efficient charges, appropriate estimates of marginal costs is the first requirement. The present paper contributes to this task by estimating one component of marginal costs. More specifically, the purpose of the paper is to provide a model for, and a measure of the (short-run) marginal infrastructure cost, i.e. costs related to the impact on resurfacing periods of additional vehicles using the infrastructure. Three hypotheses are tested against data; quality deteriorates due to the number of heavy vehicles (hypothesis 1) and the number of light vehicles (hypothesis 2) using a road. In addition, quality deteriorates over time, independently from the extent of traffic (hypothesis 3).

Figure 1 illustrates the complex engineering interactions between road design, usage, and deterioration that have to be disentangled in order to address the purpose of the paper in an intelligible way. The specification of road quality or standard is at the core of this discussion. In Sweden, this analysis of quality is facilitated by annual measurement in terms of rut depth and horizontal evenness using the International Roughness Index (IRI). While the quality values per se are of secondary relevance for the present analysis, a critical assumption is that engineers have established some threshold level of rutting and/or IRI statistics that triggers resurfacing. This eliminates situations with resurfacing taking place at pre-set time intervals or being random.

Except for surface smoothness, road quality differs with respect to asphalt thickness, stone quality etc. Since we have excellent data also in this dimension of quality, it is made part of our analysis. Sweden does not use concrete, only flexible pavements.

Expected traffic is a crucial determinant of quality. The center box in Figure 1 indicates that traffic affects both the design quality of a new road (arrow to the left) and the deterioration of roads, once they start to be used (rightward arrow). The more traffic that is expected on a road, the more robust is the design standard (thickness of sub- and superstructure, etc.). At the same time, the more traffic using an existing road, the faster is the deterioration. However, if the road was initially built to tolerate much traffic, deterioration may still be slow.

3

Figure 1: Interrelationship between design, traffic, and deterioration of roads.
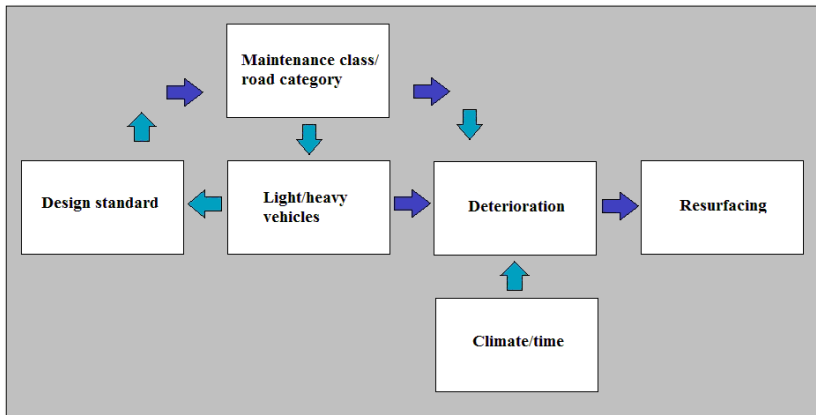
Except for the relevance of traffic for both design standard and deterioration, the combination of design standard and traffic also provides the basis for the strategy for on-going maintenance of a road (the top-most rectangle). Most countries classify roads into being built for use of long distance international or national traffic or for regional or local purposes. There may, however, be sections of an international road that have little traffic while parts of local roads may be heavily used. And irrespective of the details of road categorisation, day-to-day maintenance may be implemented in ways that have consequences for the degree of deterioration and for the subsequent date of pavement renewal.

A further feature of the interrelationship is climate or time per se (the bottom-most rectangle), as road surface standard may deteriorate independent from use. This is a subject where we have failed to find a common view amongst engineers but which will be further analyzed in the paper.

With these interrelationships as a background, the idea is that the larger the number of vehicles using a road, the more damage is inflicted on the infrastructure. Furthermore, if traffic increases relative to ex-ante projections, future periodic maintenance must be advanced in time in order to retain quality. This imposes an additional cost to society, which is the marginal cost related to traffic in focus in the present paper.

An important aspect that cannot be captured by the figure concerns the relationship between vehicle weight and road wear; this association is commonly believed to be non-linear. For this reason, it is necessary to make a distinction between the impact of heavy and light vehicles on road quality. The standard rule-of-thumb is that wear increases according to the fourth power of weight per vehicle axle. This rule, and the relative significance of heavy and light vehicles for quality and the timing of reinvestment, and therefore for estimating the marginal costs, is therefore at the core of the design of a welfare enhancing charging system.

Previous literature: The estimation of marginal infrastructure cost, i.e. costs related to additional vehicles using the infrastructure at large, has a long history. Newbery (1988)

4

focuses on a particular aspect of these estimates, namely the fact that the wear of one vehicle using a road imposes on subsequent vehicles' costs for using a road of slightly lower quality. His fundamental theorem demonstrates that this extra cost under some assumptions is balanced by the date of the future reinvestment being slightly advanced in time. Under these, while not under other assumptions, this between-vehicle externality cancels out.

The leading model formulation for empirical analysis is the book by Small et al (1987), which is used for constructing the base model in our paper. A number of subsequent empirical studies, not least in the engineering science, are based on this method. Lindberg (2006) added to this model by using the concept of deterioration elasticity in order to facilitate empirical estimations; this is defined in more detail in section 2. In his dissertation, Haraldsson (2007) makes use of this analytical trick and also introduces the possibility of using a Weibull distribution for modeling the life length of road network sections. The dissertation also includes a more comprehensive review of issues related to the assessment of marginal costs. Our paper provides an extension and specification of his dissertation. In addition, by making use of a unique database, we are able to estimate not only the marginal costs for road wear but also to address some of the underlying complexities captured by Figure 1.

Using the basic components of the Small et al (1987) model, section 2 specifies the analytical approach and also identifies the type of data necessary for the empirical assessment. We also describe where the previous analysis is extended. Section 3 presents information on costs for re-surfacing used for providing a measure of the average resurfacing cost per square meter of the road. Section 4 applies a time-to-event model for estimating the surface life length and the traffic using roads between "birth" and "death" of a pavement. Section 5 summarizes results.

2 The modeling framework

This section starts by developing the economic model for calculating marginal reinvestment costs (section 2.1). An essential input of this model concerns the life-length of roads and section 2.2 describes the engineering aspects of these calculations. While the first two sections treat all vehicles as being identical, section 2.3 elaborates on the implications for cost estimation of vehicle/axle weight. Section 2.4 summarizes the framework section by formulating the testable hypotheses.

2.1   The engineering model

Figure 2 captures the economic framework for the reinvestment problem. The solid line characterizes the deterioration and worsening quality of a piece of infrastructure as time goes and as more and more vehicles have used it. At some point in time, $t = T$, quality reaches a critical standard (π_f), and as a result the road standard has to be restored, ideally to the original level (π_0). After that, the degradation starts once again.

The pattern of deterioration-rehabilitation cycles is based on expectations regarding future traffic when the road is originally built. The analytical trick of the model is to assume that at some point in time, τ, traffic increases relative to the ex ante belief. The consequence of the unexpected (one-time) addition of traffic, and therefore also wear, is that the critical quality level will be reached slightly earlier than predicted, making it necessary to frontload

5

the rehabilitation activity.[2] Spending on rehabilitation earlier than planned represents a cost to society. Since the frontloading effect continues for the foreseeable future, the rather small cost increase the first period may boost the present value of resurfacing substantially. The extent of the cost increase is related to the frequency of resurfacing activities and the level of the discount rate.

Figure 2. Renewal intervals without and with a marginal increase in traffic at time τ.
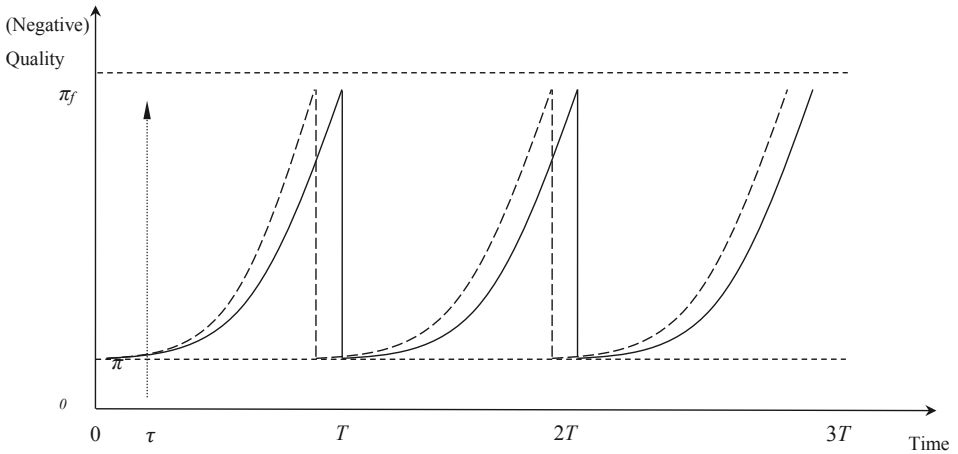
In order to model these effects, let C represent the cost per square metre for a resurfacing activity. A new road surface is laid every T years, and equation (1), where r denotes the discount rate, defines the present value of all future overlay costs (PVC) at time T.

$$PVC_T = \frac{C}{(1-e^{-rT})} \qquad (1)$$

It is, however, necessary to account for that the external shock may take place at any time between the most recent rehabilitation (at $t \approx 0$) and just before the next ($t \approx T$). In eq. (2), the present value of all future pavement renewal costs after $t = \tau$ is therefore discounted by e^{-rv}, $v = T - \tau$ being the remaining life of the pavement.

$$PVC_\tau = \frac{Ce^{-rv}}{(1-e^{-rT})} \qquad (2)$$

---

[2] In Figure 2, traffic growth would mean that the time between resurfacing intervals would gradually become shorter. Newbery (1989) handles this by assuming a constant time interval while rehabilitation cost increases since more must be spent on durability (e.g. thickness of the pavement) in order to keep T constant.

6

Differentiating (2) with respect to traffic (Q_τ)[3] provides the marginal costs (MC).;

$$MC_\tau = \frac{\partial PVC}{\partial Q_\tau} = \frac{\partial PVC}{\partial T}\frac{dT}{dQ_\tau} = -Cr\frac{e^{-rv}}{(1-e^{r\overline{T}})}\frac{\partial v}{\partial T}\frac{dT}{dQ_\tau} \tag{3}$$

Following Lindberg (2004), it is instructive to rewrite this expression in terms of changes in annual average traffic between $t = 0$ and $t = T$, say $\overline{Q}$; in the absence of traffic growth, $Q_\tau = \overline{Q}$. He also introduces the concept of deterioration elasticity (ε), $\varepsilon = \frac{\delta T}{\delta \overline{Q}} \times \frac{\overline{Q}}{T}$. This is a measure of the responsiveness in pavement life to a change in traffic intensity. Since $\frac{\delta T}{\delta \overline{Q}} = \varepsilon \times \frac{T}{\overline{Q}}$ and $\frac{\partial v}{\partial T} = 1$, the relation between a momentary traffic change and deterioration elasticity is given by eq. (3');

$$MC_\tau = \frac{\partial PVC}{\partial \overline{Q}} = -Cr\frac{e^{-rv}}{(1-e^{r\overline{T}})}\frac{\varepsilon}{\overline{Q}} \tag{3'}$$

The average MC over all possible remaining lifetimes from the date of the traffic increase, or analogously over a large number of road sections of different remaining lifetime, is the expected marginal cost taken over a probability density function of v, $g(v)$:

$$E\left[\frac{\partial PVC}{\partial Q_\tau}\right] = -\frac{Cr\varepsilon}{\overline{Q}}\int_0^\infty \frac{e^{-rv}}{(1-e^{rT})}g(v)dv \tag{4}$$

If the pavement deteriorates deterministically with traffic and if the lifetime of a pavement comes to its end exactly when its quality falls to a predetermined level, $g(v)$: is uniform, i.e. $g(v) = \frac{1}{T}$. Under this assumption, eq. (3) collapses to the below expression where the expected marginal cost is equal to the deterioration elasticity times the average reinvestment cost (the quotient in the below equation)

$$E\left[\frac{\partial PVC}{\partial Q_\tau}\right] = -\frac{Cr\varepsilon}{\overline{Q}}\frac{1}{(1-e^{rT})}\left[-\frac{1}{r}e^{-rv}\right]_0^T = -\varepsilon\frac{C}{\overline{Q}T}$$

We assume, however, that pavement lifetime T is not deterministic. Pavement durability is modelled using a Weibull function; the motive and variable definitions is given in the next section. The survival function of a Weibull function implies the following pdf for remaining lifetimes:

$$g(v) = \frac{e^{-\gamma v^\alpha}}{E[T]}, 0 < v < \infty$$

---

[3] The precise meaning of 'traffic' is not important here. In the subsequent empirical applications heavy traffic is, however measured using the number of standard axels (ESAL) and passenger traffic the number of vehicles.

7

Substituting this into eq. (3') gives eq. (4), which is the analytical version of the model we seek to estimate. The first component of this expression is elasticity and the second – cost, life expectancy, and traffic – are the same as in the deterministic version. The third component is related to the discounting and uncertainty with respect to when in the period between an almost new pavement ($t = 0$) and a pavement that is about to be replaced ($t = T$) that the external shock takes place. The fourth component allows for uncertainty with respect to when the pavement's life ends.

$$E\left[\frac{\partial PVC}{\partial Q_\tau}\right] = -\varepsilon \frac{C}{E[T]\bar{Q}} \frac{r}{\left(1-e^{-r\overline{T}}\right)} \int_0^\infty e^{-rv-\gamma v^\alpha} dv \qquad (5)$$

Note that with traffic growth, i.e. if $Q_\tau \neq Q_0$, the numerical result of (4) will differ according to the date of the external shock. While the growth factor does not affect the expression per se, it is addressed in the empirical part of the paper. With x representing traffic growth $Q_\tau = (1 + x) \times Q_{\tau-1}$. Information about Q_o and x makes it straightforward to calculate $\bar{Q}$.

The input for calculating expected marginal costs as depicted by eq. (4) requires information about contract costs (C) or rather average costs ($C/\bar{Q}$) as well as features related to the way in which pavements decay over time. The estimation of average cost is presented in section 3, where we will also be able to define average cost for three different pavement qualities. Before that, the next two sections elaborate on engineering aspects of the equation.

2.2   The engineering model

Svenson (2014) uses the semi-parametric Cox proportional hazards model for analyzing pavement quality deterioration and life-length. One benefit of that approach is that no a priori assumption of a specific distribution of the pavement lifetime is required. The term 'proportional hazard' refers to the fact that the ratio between the hazards – the probability of a new asphalt being spread, given that it has lasted for a certain period of time – of two road sections with different values of one variable (e.g. different pavement types) is constant over time. This is convenient for the estimations since the expected life of a certain pavement or road section is not affected by when it was originally built, i.e. whether it was spread in 1990 or 2005 – its expected life is the same.

The Weibull distribution is the only parametric distribution that has this proportional hazards feature. An advantage of parametric models compared to the semi-parametric Cox model is the ability to directly estimate the effect of explanatory variables on the lifetime. The response variable in the Cox model is the hazard and the lifetime can only be estimated indirectly through the relationship between the hazard and survival functions. The Weibull distribution is flexible regarding the characteristics of the hazard function, and thus it is a common choice when modeling pavement lifetimes (see e.g. Tsai 2003, Haraldsson 2007, Hong 2008, Wang 2008, Dong 2011). In the appendix, we show that the Weibull model estimates of life length are 8 to 16 percent longer than the Cox model's for our dataset.

The Weibull distribution function ($F(t)$), survival function ($S(t)$) and hazard function ($\lambda(t)$) are represented below; the probability density function of remaining lifetimes v was defined in section 2.1. The distribution has two parameters. $\gamma > 0$ is the scale parameter, defining the spread of the distribution, and α is the shape parameter. If $\alpha = 1$ the failure rate

8

is constant over time, while a value of $\alpha > 1$ $(\alpha < 1)$ implies a failure rate that increases (decreases) with time.

$$F(t) = 1 - e^{-\gamma t^{\alpha}}$$
$$S(t) = e^{-\gamma t^{\alpha}}$$
$$\lambda(t) = \frac{-S'(t)}{S(t)}$$
$$= \gamma \alpha t^{\alpha - 1}$$

Explanatory variables can be introduced by adding an exponential factor to the hazard.

$$\lambda(t|z) = \gamma \alpha t^{\alpha - 1} \exp(\boldsymbol{\beta}'\boldsymbol{Z}) \tag{6}$$

It can be shown that, by taking the log of time lnT and substituting $\gamma = \exp(-\mu)$ and $\theta = -\alpha\beta$, (5) corresponds to the log-linear accelerated failure time model (see for instance Kiefer, 1988 or Klein and Moeschberger, 2003):

$$lnT = \mu + \boldsymbol{\beta}'\boldsymbol{Z} + \epsilon/\alpha \tag{7}$$

Where ϵ is an error term following an extreme value distribution.

Small et al (1987) assume that the quality of a (section of) road at time t deteriorates deterministically in the way depicted by eq. (5) where $N = \sum_{t=0}^{T} Q_t$ is the traffic, i.e. the number of (light and heavy) vehicles the road is designed to bear before resurfacing; cf. also Figure 2. Initial road quality has been normalized to zero, $\pi_0 = 0$.

$$\pi_t = \pi_f(\frac{\bar{Q}*t}{N})e^{mt} \tag{8}$$

where π_t is the road quality at time t. Equation (8) shows that the higher the traffic load the faster road quality will reach the critical standard π_f. Further, the exponential part indicates that pavement roughness may increase at a rate $0 \leq m \leq 1$; if $m = 0$, road quality is proportional to cumulative traffic ($Q \times t$). The presence of the m-variable may represent several features. Figure 1 points to that ageing per se may affect the standard, meaning that even a road that is not used would decay. If this is driven by weather or climate, different countries may see their roads deteriorate in different ways. It is reasonable to give (8) a stochastic interpretation. i.e. assuming that decreasing road quality increases the probability of ending the lifetime. That specific formulation however does not correspond to any familiar stochastic model. We therefore use another formulation with similar properties.

9

Replacing the exponential variable e^{mt} in eq. (8) by the power function $t^{\alpha-2}$[4], and initial road quality π_f by parameters $\gamma\alpha$, we get the proportional Weibull hazard eq. (9) which corresponds to the general form presented above. In order to test hypotheses about the contribution to road deterioration from vehicles we have introduced the parameter θ for $\bar{Q}$. In the present application, the hazard rate indicates the chance that a pavement will be replaced at time t given that it has lasted so long.

$$\lambda(t) = \gamma\alpha \frac{\bar{Q}^\theta}{N} t^{\alpha-1} = \gamma\alpha \exp(\theta lnQ - lnN)t^{\alpha-1} \qquad (9)$$

With $\alpha = 2$ and $\theta = 1$ deterioration is proportional to cumulative traffic. If $\alpha > 2$ ($\alpha < 2$) the deterioration per vehicle increases (decreases) over time. This could be a consequence of any non-stationary variable except for traffic; Small et al (1987) refers to it as ageing. $\theta \neq 1$ indicates that vehicles on roads with different traffic intensity do not contribute equally to deterioration. With $\theta < 1$, vehicles on high-traffic roads do not wear down the road as much as vehicles on low-traffic roads. In terms of the hazard in (9) an extra vehicle on roads with high traffic do not increase the hazard as much as an additional vehicle on a low traffic road (ceteris paribus). The opposite is true for $\theta > 1$. The log-linear model for pavement lifetime corresponding to (9) is:

$$lnT = \mu + \beta_Q lnQ + lnN + \epsilon/\alpha \qquad (10)$$

In terms of the economic model, $\beta_Q = -\theta/\alpha$ is the deterioration elasticity. Obviously, if $\alpha = 2$ and $\theta = 1$ so that the hazard is proportional to cumulative traffic, the deterioration elasticity is -0.5. This corresponds to the intuition that doubling traffic should half the lifetime.

2.3   Light and heavy vehicles

Before estimating eq. (4) or (6) empirically, it is necessary to elaborate on the treatment of traffic. The combination of eq. (7) and (8) facilitates an analytical separation of heavy and light vehicles in the deterioration process. Equation (7) is used to handle the fact that vehicle weight affects road deterioration and consequently that light and heavy vehicles may not be handled in the same way in the empirical estimations of road deterioration. y is the number of days per year; using a traffic-per-day statistic is the standard way in the industry to represent traffic information. W_{ia} is the weight (tons) on axle $a = 1, ..., A$ which is divided by 10 for normalisation to Newton. Weight per axle is raised to the power σ to represent the fact that road wear of vehicle of class $i = 1, ..., I$ increases exponentially.

---

[4] Both functions are positive and increasing if t>0 and m>0. For the exponential function, the second derivative is always positive if m>0, while the second derivative varies with the value of m (α). As long as there is no particular motive for assuming an exponential relationship, the two are reasonably similar.

10

$$\mu_i = \sum_{a=1}^{A} \left(\frac{W_{ia}}{10}\right)^{\sigma} \tag{7}$$

$$\pi(t) = \pi_f \frac{\sum_{i=1}^{I} \mu_i \bar{q}_i}{N^{\tilde{\beta}}} t^{\alpha-1} \tag{8}$$

μ_i therefore translates the weight of each vehicle class to a number representing the impact on road standard of that class. The concept is known as "standard axles", and defined using the universally agreed Equivalent Standard Axle Load (ESAL), one ESAL being a single axle of 18 000 pounds (8 164 kg). Equation (8), where $\bar{q}_i$ is the annual number of vehicles of class i, is then simply eq. (5'') incorporating the fact that vehicles of similar weight cannot be handled in the same way.

The value of σ in eq. (7) is of immense importance for the translation from weight per axle to road wear. The conventional wisdom is that $\sigma = 4$, commonly referred to as the fourth power rule, meaning that an increase from 8 to 10 tonnes per vehicle axle does not increase wear of vehicle type μ_i by (10/8=) 25 percent, but by ((10/8)^4=) 144 percent.

In Table 1 the number of standard axles has been computed for two types of trucks in Sweden using the fourth power rule. The first row indicates that there are 4 269 and 176 trucks in Sweden weighing up to 5 tonne and that have two and three axles, respectively. The wear factor – the ESAL number – of both is very low. Therefore, it is obvious that vehicles below 5 tonne have no effect on road wear. Since passenger cars typically weigh less than two tonnes, they are irrelevant from a quality deterioration perspective.

The table, secondly, indicates that adding an axle for a given weight will reduce the road wear. This also carries over to trucks with higher total weight and more axles. Third, the table also illustrates the successive increase of road wear (ESAL) as weight increases. Fourth, numbers in italics refer to vehicles built to carry heavy loads but that exceed maximum load according to regulations. While the regulated weight maximum of a two-axle truck is 18 tonnes, well over 8 000 trucks are constructed for carrying more than that.

The fourth power rule emanates from empirical tests is the US mid-west made in the late 1950's. Re-estimating the original data using more up-to-date econometrics, Small et al (1987) confirms the result, landing a coefficient of 3.7. Since a number of aspects other than vehicle weight may affect road wear, the rule of thumb has been widely challenged. Variations in axle configuration (boggy or single axles), type and configuration of tires as well as road construction etc. may affect the extent and speed of deterioration. No alternative to the power four has, however, won common acceptance.

11

Table 1: Number, weight and ESAL of Swedish two- and three-axle trucks. Source: Sweden's vehicle registry and own calculations.

Total weight, tonne	2 axles		3 axles	
	ESAL	No. of vehicles	ESAL	No. of vehicles
5	0,01	4269	0	176
7,5	0,06	3163	0,01	114
10	0,18	1833	0,04	43
12,5	0,45	5513	0,1	36
15	0,93	3002	0,21	13
17,5	1,73	3548	0,4	95
20	2,02	7907	0,67	216
22,5	2,18	1363	1,08	339
25	-	-	1,64	795
27,5	-	-	2,41	22932
30	-	-	3,41	9268
32,5	-	-	4,7	123
35	-	-	7,25	56

As part of a government assignment trigging the research in our paper, resources have also been made available for using our institute's Heavy Vehicle Simulator (HVS) to test the fourth power rule. Three types of roads with different strength have been built and then worn down in a "lab" by axles of three different weights in order to establish when the surface reaches the critical quality level. The equipment makes about 22 000 passages per full day corresponding to 150 000 passages per week.[5] The trials only generated (3 x 3=) 9 observations which mean that results are not statistically valid. However, the approach point to a way for taking this type of analysis one step further. One striking observation is, moreover, that the measurement results are very close to the fourth power rule-of-thumb; cf. further Erlingsson (2014). This provides further support for not deviating from using the standard fourth power rule of thumb in our empirical estimations.

   Except for axle weight, there is a separate discussion about that passenger vehicles may damage pavements due to their use of studded tires in countries with repeated freeze-thaw cycles. Since we have access to detailed information, it will be feasible to address the possibility that not only heavy but also passenger vehicles are of relevance for when a road is given a new surface layer. This is done by way of representing usage in terms of both heavy vehicles (ESAL) and the number of passenger vehicles in the estimation of pavement life. In equation (6), Q is then separated into Q_{-j} and Q_j, the first accounting for average ESAL of all heavy vehicles and the latter the number of passenger vehicles using each road section.

---

[5] Further information is available at http://www.vti.se/en/vti-offers/accelerated-testing-road-construction/

2.4   Summary

To summarize, eq. (4) establishes the way in which expected marginal costs relating to the impact of traffic on the need for reinvestment is to be calculated. Section 2.2 has elaborated on the way in which road quality deteriorates over time while section 2.3 has emphasized the need to distinguish between heavy and light vehicles. While the difference between vehicle types has a huge impact on the numerical outcome of the estimations, it does not affect eq. (4). The only consequence is that Q (and N) is not conceived of as one number representing vehicles at large but ESAL's of heavy vehicles and the number of passenger vehicles.

The generation of information for estimating marginal costs as defined by equation (4) therefore means that three hypotheses are tested against available data: quality deteriorates due to the extent of heavy vehicles, measured as ESALs (hypothesis 1), the number of light vehicles (hypothesis 2) and time, independently from the extent of traffic (hypothesis 3).

3   Calculating reinvestment costs

The Swedish Transport Administration (*Trafikverket*) tenders both maintenance and reinvestment activities. 285 resurfacing contracts, tendered during 2012 and 2013[6], have been made available for deriving a value of average cost for resurfacing, i.e. a value of $(C/\overline{Q})$ in eq. (4). Contracts range in size from just over SEK 1 million to SEK 65 million; the smallest contracts are below 1000 m^2, the largest being 2.2 million m^2. This information is used to calculate the average cost for the country as a whole as well as for each region; cf. Table 2.

The table separates three pavement techniques or qualities, warm and half-warm pavements and tank lining. These represent different pavement thickness, stone quality etc. with warm pavements being the most expensive and used for roads with much traffic. The much cheaper tank lining is spread on roads with little traffic. The estimate of national average cost is calculated using the relative size $(\sum_i m_i^2)$ as weight for region and type of pavement. In two regions, no tank-lining and half-warm contracts have been tendered during these two years. Since there indeed are roads with these pavement types also in these regions, the national average for the respective category has been imputed in these cells.

The average cost for a contract is SEK 87 per m^2. Contrary to ex-ante expectations, resurfacing using materials defined to be warm are less expensive per m^2 that half-warm pavements. Officials at *Trafikverket* have suggested that the half-warm pavement is used on roads with intermediate traffic levels (between 1000 and 7000 vehicles per average day) that may not have been built to standard from the beginning. If this is correct, the resurfacing activity may in reality represent a rehabilitation project. The resurfacing activity and the ensuing cost is triggered by that the surface standard of the road is inappropriate, and the choice has been made to use this figure in the estimations of marginal costs.

---

[6] Because of low or non-existent inflation during these and the subsequent two years, no price-level adjustment have been made. The approximate exchange rate is €1=SEK9,20

13

Table 2: Average cost per contract in six regions and for three types of pavement. 2012 and 2013, SEK/m². [7] None – no contract using this method has been tendered in this region. Number within brackets has been imputed using the average cost in these cases.

Region		Method			
		Tank-lining	Half-warm	Warm	Total
Middle	Average cost, SEK	26	127	110	98
	No. of contracts	8	9	19	37
	$\sum_i m_i^2$ million	5.1	1.9	4.2	11.3
North	Average cost, SEK	21	148	108	124
	No. of contracts	4	25	19	50
	$\sum_i m_i^2$ million	4.1	2.8	2.9	9.9
Sthlm	Average cost, SEK	31	(124)	100	97
	No. of contracts	1	No tender	29	30
	$\sum_i m_i^2$ million	0.5	-	2.6	3.1
South	Average cost, SEK	38	(124)	78	68
	No. of contracts	13	No tender	42	55
	$\sum_i m_i^2$ million	2.6	-	4.8	7.4
West	Average cost, SEK	21	79	81	71
	No. of contracts	10	9	45	64
	$\sum_i m_i^2$ million	3.4	3	5.7	12.1
East	Average cost, SEK	20	101	86	78
	No. of contracts	11	13	25	49
	$\sum_i m_i^2$ million	5.1	2.1	6	13.2
Total	Average cost, SEK	27	124	90	87
	No. of contracts	47	56	179	285
	$\sum_i m_i^2$ million	20.7	9.7	26.1	56.8

4   Estimating pavement life

Trafikverket's Pavement Management System (PMS) is used for storing data collected during annual road quality measurement activities. The system also registers when a road is "treated" in different ways, including the timing of major pavement renewals. In addition, it includes

---

[7] Costs are in the price level of the respective years, but inflation was around zero at this time. The exchange rate is approximately SEK9=€1 and SEK7.50=$1

14

information about traffic using each road segment. Section 4.1 provides further information about this dataset, section 4.2 details the Weibull model used for estimating pavement life while section 4.3 presents the results.

4.1   Data

The 2012 version of the PMS comprises 390 966 observations of homogeneous road sections that vary in length from over one kilometer to only a few meter. Only sections that are 50 meters or longer have been retained. Sections shorter than 50 meters are often crossings or junctions where the pavement life usually differs from the road network in general. The elimination of short sections in combination with removal due to data inconsistencies has left 266 614 road sections to use in the analysis.

A lot of information is available about each section.[8] This includes which out of five different width classes that the section belongs to as well as the precise type of pavement laid, facilitating the estimation of life expectancy on a very disaggregate level. The previous section, however, demonstrated that information is less opulent about resurfacing costs. Since the only type of information that can be used for estimating marginal costs at a disaggregate level is three main categories of pavement defined in section 3 and since each of the six regions tender these contracts, 18 disaggregate observations of age and costs are made. Except for possible managerial differences, the regions differ with respect to climate, the situation in the north of the country being different from those in the south.

Table 3 provides some descriptive information about the Swedish road network in the year 2012. This information is acquired from the Swedish National Road Database, which unlike the PMS database comprises up-do-date information about all national roads. The fifth column illustrates that the most expensive and resilient hot pavement type indeed is used for resurfacing roads with much traffic.

Section 2.3 emphasized the necessity to separate light (Q_{-j}) from heavy vehicles (Q_j) in eq. (4) and in particular the huge significance of heavy vehicles' weight per axle. The final step in the compilation of data for the estimations is therefore to convert the number of heavy vehicles to ESALs. An extensive system is in place for measuring traffic on network links, including the share of heavy vehicles. While the number of heavy vehicles is known, the information about vehicle weight is poor. However, since 2004, *Trafikverket* collects this type of information at 12 places across the country using a Bridge Weigh-in-Motion system. One week per year, the length of each heavy vehicle is registered as well as the weight of each axle. While there are several shortcomings with the system per se, this provides the type of information appropriate for the present purpose; see further Erlingsson (2010).

Based on suggestions from *Trafikverket* staff, we have converted the observed number of heavy vehicles (j) on road section i (Q_j^i) to ESAL for four different types (k) of roads, using coefficient φ^k. $\varphi^1 = 1.1$ for European highways; $\varphi^2 = 1.0$ for national roads; $\varphi^3 = 0.8$ for county roads if the share of heavy traffic is below 13 percent; $\varphi^4 = 1.5$ for county

---

[8] Svenson (2014) describes this information in more detail.

roads if the share of heavy traffic is 13 percent or higher. Our concluding discussion includes a test of the sensitivity of our results for these assumptions.

Table 3: Descriptive statistics of traffic and road length within each region and surface category registered the year 2012. Source: Swedish National Road Database, NVDB.

Region	Surface category	Traffic (million vehicles/year)	Road length (km)	Average no. of vehicles/road km, million	Thereof heavy traffic (%)
Middle	Cold	6442	8319	30,3	11
Middle	S.D.	745	2504	7,3	8
Middle	Hot	43627	10211	345,9	11
North	Cold	1595	2464	19,8	14
North	S.D.	273	777	19,1	10
North	Hot	43679	4387	1011,6	8
Sthlm	Cold	112	166	6,3	6
Sthlm	S.D.	620	1713	7	7
Sthlm	Hot	62709	10135	745,6	9
South	Cold	1296	1349	22,8	7
South	S.D.	714	2152	17,3	8
South	Hot	113224	14062	664,3	10
West	Cold	1309	1855	12,6	8
West	S.D.	762	1825	8,2	7
West	Hot	48997	9656	407	13
East	Cold	548	682	22,5	9
East	S.D.	817	3102	6,3	6
East	Hot	49356	10009	462,4	13

Notes: S.D.= surface dressing

4.2 Modeling life length

In order to compute the expected marginal present value cost in eq. (4) it is necessary to estimate the deterioration elasticity, ε, and the Weibull parameters α and γ. For each renewal activity, we have information about traffic (Q) as well as the number of years the road link has been in use. Heavy traffic is represented as Q_{ESAL} and the number of passenger vehicles as Q_{car}. Since neither N (the amount of traffic that has used a road until renewal) nor π_f (the actual, critical level of road quality), can be directly observed, a vector of covariates $\boldsymbol{M}$ that may have an impact on the pavement lifetime, including a constant, is used in order to provide a consistent estimate of the traffic coefficient. A linear model is then given by eq. (8) representing the empirical equivalent of eq. (6):

16

$$lnT = lnQ_{car}\beta_{Q_{car}} + lnQ_{ESAL}\beta_{Q_{ESAL}} + \boldsymbol{\beta_M M} \qquad (8)$$

Estimated coefficients $\beta_{Q_{car}}$ and $\beta_{Q_{ESAL}}$ are used for testing hypothesis 1 and 2 respectively. If these coefficients are significantly different from zero they will signal the impact of heavy traffic and passenger vehicles on reinvestment costs. If so, the coefficient values will also be used for calculating the respective marginal costs. In addition, the value of $\hat{\alpha}$ is used for testing the hypothesis that there is an independent effect on the hazard of a road being "treated". Specifically, if $\hat{\alpha} > 2$ the last component of eq. (5') makes the hazard increase at an increasing rate with time. With estimates of $\beta_{Q_{car}}$ and $\beta_{Q_{ESAL}}$ it is also feasible to compute the deterioration elasticities:

$$\hat{\varepsilon}_{ESAL} = \frac{\delta lnT}{\delta lnQ_{ESAL}} = -\hat{\beta}_{Q_{ESAL}} \text{ and } \hat{\varepsilon}_{car} = \frac{\delta lnT}{\delta lnQ_{car}} = -\hat{\beta}_{Q_{car}}$$

4.3   Results

Table 4 provides the results of the estimates of eq. (8). As expected, the large number of observations result in very precise estimates, for instance establishing that cold surfaces as well as surface dressing have statistically significantly shorter life than warm pavements. Moreover, Stockholm's roads last shorter time than roads in the other regions; for instance, roads in the Middle region "live" about 14 percent longer than roads in Stockholm. A possible reason is that even though roads in Stockholm are robustly built, the road construction may not be robust enough for the much higher traffic load in this region compared to other regions.

Based on the results summarised in table 4, hypothesis 3 can be rejected while hypotheses 1 and 2 are not: Both light and heavy vehicles affect the timing of resurfacing activities and consequently the life length of pavements, while there are no aspects related to time or weather that has this impact ($\hat{\alpha} = 1,61$ which is lower than the critical value $\hat{\alpha} < 2$). Bearing in mind the crude way to transform information about heavy vehicles into ESAL, it is noteworthy that the impact of passenger vehicles on surface life is at least as strong as the consequences of variations in heavy vehicles.

Our hypothesis is that the significant value of the coefficient for cars can be rationalized by their use of studded tires. If this hypothesis is correct, it is reasonable that cars' road wear in the north of the country is at a lower level than in the reference region, which is Stockholm in the south-east part of the country. This is so since the road surface furthest north is covered by snow and ice for longer periods than in the southern parts of the country, meaning that the studs do not wear down the pavement for long periods. In addition, roads in the southern parts of Sweden have less harsh winters than up north, meaning that fewer cars use studded tires.[9] Interacting cars and regions provide an indication of that these conjectures may be correct. Compared to Stockholm, the car-region coefficient is some 20 percent higher for regions North, South and West while roads in regions East and Middle last about 10 percent longer.

---

[9] SMHI (2008) indicates that 44 percent of cars in region South had studded tires in 2008 while the average for the rest of the country is close to 80 percent.

17

Table 4: Estimates of surface life length using a Weibull model. 252 309 observations of homogeneous road sections from the Swedish Pavement Management Systems.

	Coefficient	Std. Error	Z	p-value
Intercept	4,03	0,014	282,1	0,0000
$\hat{\varepsilon}_{ESAL}$	-0,09	0,002	-45,4	0,0000
$\hat{\varepsilon}_{car}$	-0,10	0,003	-40,4	0,0000
Hot*	0,00			
Cold	-0,24	0,005	-48,0	0,0000
Surface dressing	-0,14	0,004	-36,3	0,0000
Sthlm*	0,00			
Middle	0,14	0,007	19,8	0,0000
North	0,24	0,008	31,0	0,0000
South	0,19	0,007	29,0	0,0000
West	0,20	0,007	30,0	0,0000
East	0,03	0,007	3,71	0,0002
Log(1/ α)	-0,48	0,002	-249,8	0,0000
α	1,61			

* - reference category.

Figure 3 shows the empirical (in black) and predicted cumulative hazard functions. The cumulative hazard, which is equivalent to the integrated hazard function, is often used as a graphical assessment of the goodness of fit of the parametric model. The empirical cumulative hazard represents the entire data set by calculating the Nelson-Aalen estimator; d_i is the number of events (i.e. resurfacing actions) at time t_i, and n_i is the number of objects (road sections) at risk of the event at t_i:

$$\tilde{\Lambda}(t) = \sum_{t_i < t} \frac{d_i}{n_i}$$

The predicted curves represent mean values for traffic and ESAL and the Stockholm region. Stockholm has among the shortest lifetimes, which is why all predicted curves lie above the empirical. A fairly straight empirical cumulative hazard implies that α (the shape parameter of the Weibull distribution) is close to 1. The estimated value of α is 1,61, which makes the predicted curves bent slightly inwards and the hazard (i.e. failure rate) is increasing with time.

Based on the results in table 4, estimated pavement life length based on car traffic and ESAL is summarized in table 5. Median lifetime is strikingly similar across surface types and regions.[10] Hot pavements live slightly shorter in spite of being more robust. Most probably, the reason is that they are used by much more traffic than roads with other types of surface treatment.

---

[10] Median life is the spot where the survival function $S(t) = 0.5$. The hazard function $h(t)$, which is estimated, is directly related to the survival function since $h(t) = -dlnS(t)/dt$.

18

Figure 3: Empirical and predicted cumulated hazards for surface types.
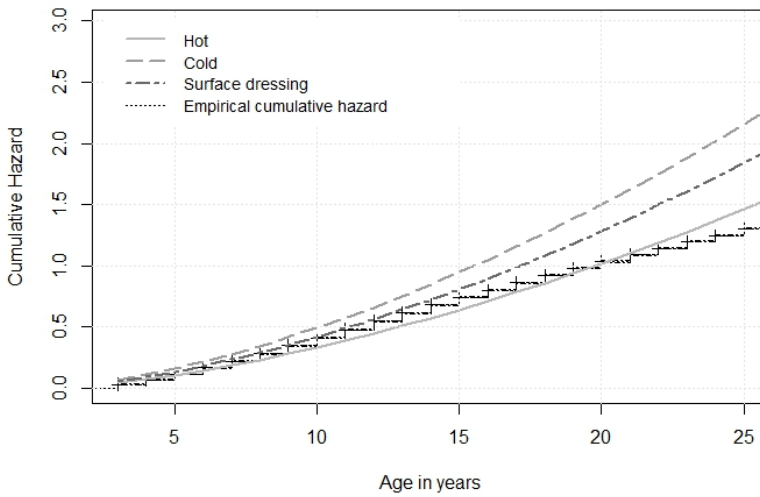
Table 5: Lifetimes estimated from the Weibull model.

Region	Surface	ADT cars	ADT ESAL	Median life (years)
Middle	Cold	344	22	16,7
Middle	S.D.	455	36	17,2
Middle	Hot	1339	105	16,2
North	Cold	326	36	17,8
North	S.D.	242	26	20,8
North	Hot	1244	110	17,9
Sthlm	Cold	283	12	15,7
Sthlm	S.D.	498	27	15,2
Sthlm	Hot	4999	219	11,5
South	Cold	271	11	19,2
South	S.D.	286	13	20,8
South	Hot	1480	98	16,9
West	Cold	339	15	18,5
West	S.D.	486	25	18,8
West	Hot	2597	206	15,1
East	Cold	278	12	16,2
East	S.D.	372	17	16,7
East	Hot	2084	136	13,5
Average		2280	143	17

Notes: ADT = average daily traffic, S.D.= surface dressing

19

5 Calculating marginal costs

Equation (4) is used for calculating marginal costs. In order to elaborate on the logic of the estimations, a particular example is given.

Example: Calculation of the national average marginal cost. *C* is the construction cost, which by Table 2 is SEK 87 per square meter for an average road. *Q* is the traffic using the average road; in 2012 this was 2280 cars and 143 ESAL (trucks) per average day. The average pavement lasts for an average 17 years (*T*). In order to derive total traffic over the life of the average road (*N*), and since car traffic has increased by 1 percent p.a., it is straightforward to establish that over the period there are 2112 cars per average day.[11] With traffic growth at 1.8 percent p.a. for heavy vehicles, the corresponding number of ESALs is 125 vehicles per average day over the life cycle. The average number of heavy vehicles and cars using the average road between its birth and death is therefore (17 years * 365 days * 125=) 776 000 ESALs and (17 years * 365 days * 2112=) 13.1 million cars. The average cost is SEK 87 divided by these numbers, i.e. SEK $1{,}12 * 10^{-4}$ per ESAL and $6.64*10^{-6}$ per car.

The first component of eq. (4), ε, represents the deterioration elasticity, now split in two, i.e. $\hat{\varepsilon}_{cars}$ and $\hat{\varepsilon}_{ESAL}$. The numbers in table 3 indicate that increasing ESAL or number of cars by 10 percent will reduce the service life of pavements by about one percent for both categories. Multiplying the average cost by the respective elasticities, the result is $(0.0888*1{,}12 * 10^{-4} =)$ $9.95*10^{-6}$ for ESAL and $(0{,}1036*6.64*10^{-6} =)$ $0{,}69*10^{-6}$ for cars.

The official discount rate for the transport sector, *r*, is 3.5 percent. With median life being 17 years, the value of $\frac{r}{\left(1-e^{-r\overline{T}}\right)}$ is 0.078. The final component of eq. (4) is the integral $\int_{0}^{\infty} e^{-rv-\gamma v^{\alpha}}\,dv$. This is a means for handling the fact that the external shock, i.e. the non-expected increase in traffic, could materialise at any point of time between the previous and the next date for renewal. The value of the integral is 12.5, and the combination of the last two terms 0.976. Given eq. (4), the marginal cost estimate is $(9.95*10^{-6} * 0.976 =)$ SEK $9.71*10^{-6}$ for each ESAL and $(0.69*10^{-6} * 0.976 =)$ SEK $0.673*10^{-6}$ for each car.

This benchmark estimate of marginal costs is calculated per square meter at the same time as the standard way to represent traffic is by vehicle km. The average Swedish road being 6.75 m wide, and multiplying by 1 000 for transfer from meter to kilometre, the marginal cost for a heavy vehicle using an average road is $(6.75 * 1\,000 * 9.71*10^{-6} =)$ SEK 0.066 per ESAL km and $(6.75 * 1\,000 * 0.673*10^{-6}=)$ SEK 0.0052 per km for cars. **End of example**.

The numerical example is based on an average vehicle using an average road. A detailed calculation of marginal costs is, however, based on 477 262 observations from the Swedish National Road Database, one for each road section. In this, all road sections are given a

---

[11] y*1,01^16=2280 => y=1944, => (2280+1944)/2)

20

weight based on length relative to total road length in order to create the total average. This is the approach used to derive all values in table 6. Comparing the last row in table 6 with the manual average calculated in the above example, it is obvious that the values in the table are much higher. The reason is that the table provides a better accuracy, accounting for actual road length rather than (implicitly) assuming all links to be equally long.

Table 6: Marginal and average cost, SEK per ESAL kilometer and SEK per car kilometer. Price level 2013.

Region, pavement type	ESAL		Car	
	Marginal	Average	Marginal	Average
	Cost		Cost	
Middle, Cold	0,60	7,09	0,067	0,63
Middle, S.D	0,25	3,02	0,022	0,21
Middle, Hot	0,25	2,86	0,043	0,43
North, Cold	0,54	7,38	0,068	0,75
North, S.D	0,14	1,94	0,013	0,14
North, Hot	0,36	4,78	0,014	0,16
Sthlm, Cold	1,14	13,46	0,076	0,72
Sthlm; S.D.	0,25	2,97	0,016	0,15
Sthlm, Hot	0,38	4,39	0,021	0,20
South, Cold	0,67	7,98	0,045	0,44
South, S.D	0,33	3,99	0,020	0,20
South, Hot	0,19	2,27	0,013	0,12
West, Cold	0,42	5,05	0,033	0,32
West, S.D	0,14	1,69	0,009	0,09
West, Hot	0,19	2,33	0,014	0,14
East, Cold	0,68	7,92	0,046	0,44
East, S.D	0,20	2,37	0,013	0,13
East, Hot	0,30	3,50	0,020	0,19
All	**0,32**	**3,78**	**0,027**	**0,27**

Note: S.D. – surface dressing

21

The only available point of reference for benchmarking is Haraldsson (2007); his aggregate estimates are SEK 0.01 for heavy vehicles and 0,001 for cars.[12] This means that our cost estimates are higher than before. One reason may be that heavy traffic is here transformed from the number of vehicles to ESAL, which was not the case in the previous study. Another difference is that elasticities are now -0,09 and -0,10 while they were -0,04 and -0,052 in Haraldsson (2007) for heavy and light vehicles, respectively, i.e. they are now twice as large.[13] Moreover, we use more than twice the number of observations. Finally, although both studies are based on information from the same source, seven more years of observations are now available. We have seen in other, similar studies that there may be a change in maintenance methods during these years that may have consequences for elasticity estimates. It would require further analyses in order to sort out these differences.

6 Discussion

Table 7 provides the policy context for the present paper, summarizing all marginal cost components from our recent report to the government; cf. VTI (2014). Since the present paper has been updated after the submission of the report, the wear & tear costs in table 7 differ slightly from the results in table 6. This also illustrates the need to be magnanimous toward the precise cost levels since the devil certainly is in the detail. With this in mind, the indication is that taxation of petrol used by cars is substantially above marginal costs while the opposite situation is true for heavy vehicles.

Table 7: Average marginal costs for cars and trucks in Sweden, SEK 2013

	Cars	Trucks
Wear & Tear	0,06	0,63
Accident risk	0,01	0,005
CO2	0,12	0,70
Other emissions	0,02	0,20
Noise	0,02	0,09
Congestion	-	-
Aggregate marginal cost	0,22	1,64
Fuel tax	0,45	1,02

The availability of disaggregated data makes it possible also to consider the implications of our results from a road type perspective. Table 8 establishes the costs for Europe, national and county type roads according to the actual mix (and length) of roads with Cold, Surface

---

[12] The costs used in Haraldsson (2007) is based on a reference from 2004. Assuming that the cost per m^2, which is SEK 65, refers to year 2000, and using CPI, the corresponding number for 2012, which is the year used here, is SEK 78. Since the cost for asphalt has increased much faster than consumer prices this compares well with the value used here, i.e. SEK 85 per m^2.

[13] The covariates in the respective equations are, however, not the same.

22

Dressing and Hot pavements. County roads are lesser-used roads while the bulk of long-distance transport and travel in the first place use Europe but also national roads. A substantial proportion of Europe roads is motorways.

Using a weighted average which accounts for the length and width of sections with the respective type of pavement, Table 8 indicates that using heavy vehicles on county roads costs about 150 to 300 percent more than using them on national and European roads built for a higher share of heavy vehicles. The difference is smaller for cars with a marginal cost that is 130 to 190 percent higher for county roads vs. European and national roads, respectively. The origin of this difference is the fact that county roads are not built for a substantial share of heavy traffic. This indicates a case for differentiated charging, which would incentivize heavy vehicles to use the major roads even more than today.

Table 8: Marginal costs for heavy and light vehicles using different road categories

	ESAL	Car
Europe	0,24	0,023
National	0,12	0,016
County	0,36	0,030
National average	0,32	0,027

The results so far refer to an average heavy vehicle. Using information of the type provided by Table 2 would, however, be straightforward to estimate costs for each combination of vehicle weight and number of axles and to calculate a marginal cost for each heavy vehicle class. The Swedish government has, indeed, established a committee for designing a weight-distance tax with these qualities and for considering the costs of implementing this type of tax. The committee is also supposed to consider the appropriateness and costs for differentiating charges across different types of roads in the way indicated in Table 8.

With these results in mind, it is not satisfactory that the information about actual vehicle weight is of meager quality. We have therefore made an assessment of the implications of this uncertainty around the φ^k coefficients applied in section 4.1. Table 6 establishes a national average marginal costs at SEK 0,32. Eq. (4) demonstrates why the deviation is non linear: through Q and through γ in the integral.

23

Figure 4: Variation in marginal cost per ESAL allowing for 50 percent higher and lower average values of φ^k around the benchmark average SEK 0,32.
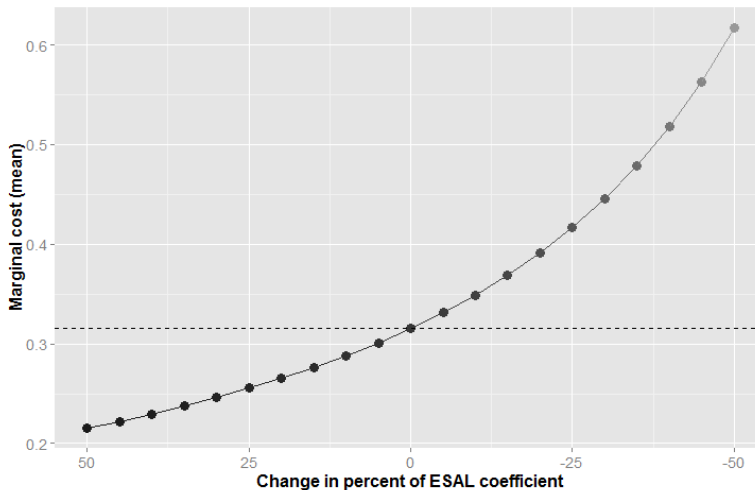
7 Summary

The present paper has estimated the marginal costs for road reinvestment using information at a very disaggregated level. One robust result of the analysis is that not only heavy but also light vehicles affect the periodicity of resurfacing activities. Most probably, this is the consequence of the use of studded tires in a country with freeze-thaw cycles.

While both trucks and cars have an impact on the timing of resurfacing, the analysis has not provided any indications of that time per se is of separate relevance for life length. One possible reason is that vehicle wear triggers the resurfacing activity long before the passage of time has any effects on the aging of the road surface.

However, it is then reasonable to ask which aspects other than vehicle wear that are of relevance for understanding the complex relationships illustrated in Figure 1. Turning the question upside down, if a doubling of traffic does not reduce the lifetime of the pavement by 50 percent, what is then decisive factors for road deterioration?

Based on World Bank research in developing countries during the 1960's, Newbery (1988) discusses different reasons for road quality deterioration. Contrasting his observations with the outcome of the present analysis may imply that roads deteriorate over time and due to weather and climate in different ways in cold and warm climates. It is also important to acknowledge that the heterogeneity across sections of roads is only partly characterized by available data. In particular, very little is known about the quality of the road beneath the top layer. The "life" of roads within a certain region that are classified in the same way and that are used by a similar number of vehicles may still differ substantially due to that they are built on solid rock, on sand or on materials that are not stable during the spring thaw. Constructions

24

may also differ simply because of different skills of the construction teams or since sub-structures built long ago had a lower standard than today. While information is abundant about surface structure, the variation in substructure still leaves the analyst with only partial information of what drives quality deterioration. Svenson et al. (2016) have addressed the stochastic nature of the data at hand. These results show that road sections' lifetimes varies significantly after controlling for several explanatory variables, indicating that latent variables such as substructures have a substantial impact on road lifetimes.

One implicit observation of this underlying heterogeneity has been made with the observation that half-warm pavements actually are more expensive per square meter than warm mixes. This is contrary to the fact that the half-warm mix in isolation is cheaper than the warm mix. The blip in cost data is most probably an indication of that the half-warm type of surface is spread on roads that were not built for the traffic it actually is carrying. As a result, resurfacing of these roads require more extensive rehabilitation activities, increasing the average (and marginal) costs.

Except for the impact on rehabilitation, variations in traffic may also affect the spending on day-to-day maintenance. This possibility is addressed by Haraldsson (2007) and is updated by Swärdh & Jonsson (2014) in a study that is parallel to the present. In addition, Anani and Madanat (2010) consider the possibility of activities which are more intense than maintenance but less so than complete resurfacing. This could, for instance, be substantial pothole mending or other activities with a possible impact on the timing of future resurfacing. Except for this type of activity possibly representing a component of marginal costs, it may affect the analysis of observed life cycles and in that respect represent an additional source of (unobservable) heterogeneity. In the wake of data about these types of activities, it is impossible to extend the analysis in this direction.

8 References

Anani, S. B., S. M. Madanat (2010). Estimation of Highway Maintenance Marginal Cost under Multiple Maintenance Activities. *Journal of Transportation Engineering,* Vol. 136, Issue 10, pp. 863–870.

Dong, Q (2011), *Enhancement of pavement maintenance decision making by evaluating the effectiveness of pavement maintenance treatments*. Ph.D Dissertation, University of Tennessee.

Erlingsson, S. (2010) "Characterization of heavy traffic on the Swedish road network," *Proceedings of the 11th International Conference on Asphalt Pavements*, Nagoya, Japan, 01 – 06 August, CD-ROM.

Erlingsson, S. (2014) "Tunga trafikens samhällsekonomiska kostnader – Accelerade tester av tre vägkonstruktioner," VTI-notat, 22 s. Utkast.

25

Haraldsson, M. (2007). The Marginal Cost for Pavement Renewal – a Duration Analysis Approach. In Essays on Transport Economics. Dissertation in Economics, Uppsala University.

Hong, F., Chen, D., and Mikhail, M. M. (2010). Long-Term Performance Evaluation of Recycled Asphalt Pavement Results from Texas. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2180, pp. 58-66.

Hosmer, D., Lemeshow, S., and May, S. (2008). Applied survival analysis: Regression modelling of time-to-event data. 2nd Edition, Wiley, Hoboken, NJ.

Kiefer, N (1988). Economic Duration Data and Hazard Functions. *Journal of Economic Literature*, Vol. 26, pp. 649-679.

Klein, John P., and Moeschberger, Melvin L. (2003) Survival Analysis – Techniques for censored and truncated data. Springer

Lindberg, G. (2006). Development path of HeavyRoute systems – impact and socio-economic consequences. Deliverable 4.2 for EU Specific Targeted Research Project "HeavyRoute".

Newbery (1988). Road User Charges in Britain. *Economic Journal*, Vol. 98, pp 161-176.

Small, K.A., Winston. C., and Evans, C.A. (1987). Road Work. Washington. The Brooking Institute.

Parry, I.W. H., and Small, K.A. (2005). Does Britain or the United States Have the Right Gasoline Tax? *American Economic Review*, Vol. 95, Issue 4, pp.1276-1289.

SMHI (2008). *Vintervägar med eller utan dubbdäck.* (Winter roads with and without studded tires; in Swedish). SMHI note 134.

Swärdh, J-E., L. Jonsson (2014). Marginal Cost of road operation and maintenance – Swedish estimates based on data of 2004-2012. Working Paper

Svenson, K. (2014). Estimated Lifetimes of Road Pavements in Sweden Using Time-to-Event Analysis. *Journal of Transportation Engineering*, Vol 140, Issue 11, 0414056.

Svenson, K., Machucova, Z., Li, Y., and Rönnegård, L. (2016) Evaluating Needs of Road Maintenance in Sweden with the Mixed Proportional Hazards Model. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2589, pp. 51–58.

26

Tsai, B-W., Harvey, J.T., and Monismith, C.L. (2003), Application of Weibull theory in prediction of asphalt concrete fatigue performance. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 1832, pp. 121-130.

Wang, Y., Allen, D. (2008), Staged survival models for overlay performance prediction. *International Journal of Pavement Engineering,* Vol. 9, Issue 1, pp. 33-44.

27

Appendix

Table A1 shows the difference in expected median lifetime between the parametric Weibull model and the semi-parametric Cox Proportional Hazards model. While the Weibull model can be specified to model the (logarithm of the) lifetime, the Cox model is semi-parametric in the sense that the hazard $\lambda(t)$ is not specified to any particular parametric shape. Median lifetimes from the Weibull model can be obtain through prediction because the hazard has a particular shape and log(lifetime) can be modelled directly.

The Cox model rather models the hazard in the following way:

$$\lambda(t) = \lambda_0(t)\exp(X\beta)$$

To get estimates of the median lifetimes from a Cox model, the relationship between the survival function and the hazard function can be used. A non-parametric estimate of the baseline hazard $h_0(t)$ must be used, for example through:

$$\hat{\lambda}_0(t) = \frac{d_i}{\sum_{k \in R(t_i)} \exp(X_k\hat{\beta})}$$

where d_i is the number of events at time t_i and $R(t_i)$ is the set of individuals that could experience the event of interest at time t_i. The $\hat{\beta}$-estimates are obtained through maximization of a partial likelihood (Hosmer et al. 2008). The median lifetimes are found where the survival function $S(t) = 0.5$, where the survival function is found by its relationship with the hazard function:

$$S(t) = \exp[-\Lambda(t)]$$

where $\Lambda(t)$ is the cumulative hazard function, estimated as;

$$\widehat{\Lambda}_0(t) = \sum_{i:t_i<t} \hat{\lambda}_i$$

$\hat{\lambda}_i = \hat{\lambda}_0(t)$, and the estimated survival function is:

$$\hat{S}_0(t) = \exp[-\widehat{\Lambda}_0(t)]$$

The table demonstrates that lifetimes estimated from the Weibull model are 0.5-2.2 years longer than those from the Cox model. This is due to the parametric form of the hazard specified in the Weibull model, which makes a stronger assumption about the remaining lifetime of the censored observations. With a large share of censored observations, the Cox model is more likely to underestimate the lifetime, while the Weibull model might over- or underestimate it depending on how close the real hazard is to the assumed Weibull hazard.

28

Table A1: Estimated lifetimes in years, comparison between a Weibull model and Cox proportional hazards model.

Region	Surface Type	Median Age Weibull	Median Age Cox	Difference (years)
Middle	Cold	16,7	15	1,7
Middle	S.D.	17,2	16	1,2
Middle	Hot	16,2	15	1,1
North	Cold	17,8	16	1,8
North	S.D.	20,8	20	0,8
North	Hot	17,9	17	0,9
Sthlm	Cold	15,7	14	1,7
Sthlm	S.D.	15,2	14	1,2
Sthlm	Hot	11,5	11	0,5
South	Cold	19,2	17	2,2
South	S.D.	20,8	19	1,8
South	Hot	16,9	16	0,9
West	Cold	18,5	17	1,5
West	S.D.	18,8	17	1,8
West	Hot	15,1	14	1,1
East	Cold	16,2	14	2,2
East	S.D.	16,7	15	1,7
East	Hot	13,5	12	1,4

Paper IV

Detecting Road Pavement Deterioration with Finite Mixture Models

Kristin Svenson[*], Stuart McRobbie[†], Moudud Alam[‡]

Abstract

Budget restrictions often limit the number of possible maintenance activities in a road network each year. To effectively allocate resources, the rate of road pavement deterioration is of great importance. If two maintenance candidates have an equivalent condition, it is reasonable to maintain the segment with the highest deterioration rate first. To identify such segments, finite mixture models were applied to road condition data from a part of the M4 highway in England. Assuming that data originates from two different normal distributions – defined as a "change" distribution and an "unchanged" distribution – all road segments were classified into one of the groups. Comparisons with known measurement errors and maintenance records showed that segments in the unchanged group had a stationary road condition. Segments classified into the change group showed either a rapid deterioration, improvement in condition because of previous maintenance, or unusual measurement errors. Together with additional information from maintenance records, finite mixture models can identify segments with the most rapid deterioration rate, and contribute to more efficient maintenance decisions.

Keywords: Finite Mixture Models, Pavement Deterioration, Road Maintenance

1 Introduction

There is a constant challenge in optimizing the efficiency of maintenance planning. Modern database systems can store road condition data for short segments of an entire road network that makes it possible to analyze microdata for macroscale applications. However, it is not realistic to manually examine and process datasets of such magnitudes, and the use of analytical models that can display relevant information become increasingly important.

Probabilistic modeling of pavement deterioration is a well-explored research topic. For instance, Lethanh and Adey (2012) use exponential hidden Markov models to model deterioration when road condition data is sparse. Hong and Prozzi (2006) applied a Bayesian approach and found it successful in finding parameter distributions for their pavement performance model. However, when making practical decisions on actual maintenance locations, the maintenance planner has to look at the condition data and determine which sections should be maintained as a priority, and which can be left until later. When making this decision the maintenance engineer must consider that sometimes the rate of deterioration is more important than the absolute condition: a section of road could be deteriorating fast, but still be in an acceptable overall condition, while another section could be stable, but in a less acceptable state. Maintaining the first section before it reaches an unacceptable state may be a more efficient use of

---

[*]School of Technology and Business Studies, Dalarna University. E-mail: kss@du.se
[†]Transportation Research Laboratory. E-mail: smcrobbie@trl.co.uk
[‡]School of Technology and Business Studies, Dalarna University. E-mail: maa@du.se

1

resources than simply maintaining the site in worse, but stable condition and allowing the rapid deterioration to continue.

By clustering parts of the network based on the measured change in the condition it is possible to objectively identify road segments that show the highest rates of deterioration. Discussion of clustering methods is a wide topic in both statistics and machine learning. Algorithms for clustering include connectivity based clustering which uses distance as a mean of clustering (see e.g. Ward Jr, 1963; Sibson, 1973; Defays, 1977), k-means clustering which assigns an object to the nearest defined cluster center (see e.g. Steinhaus, 1956; Hartigan and Wong, 1979), and distribution-based clustering which distinguishes clusters in data based on probabilistic models. Distribution-based models can capture correlations and dependencies between variables in a way that machine learning models often cannot. They are therefore a good choice when the theoretical foundation of the model is of importance, as well as the actual clustering.

In the class of distribution-based clustering methods, finite mixture models are the most commonly used (McLachlan and Basford, 1988). Within transportation research, they are not widespread but have been applied by e.g. Park and Lord (2009) who combined mixtures of two different distributions to find sub-populations in motor vehicle crash data from Toronto, Canada. Finite mixture models are more prevalent in other fields such as biochemistry, where they have been used to model protein evolution (Lee et al., 2008); in genetics, where they are used to detect heterogeneity in gene sequences (Pagel and Meade, 2004); and in medicine, where heart rate variability has been modeled with gaussian mixture models (Daeyoung and Lindsay, 2015). Finite mixture models are also valuable for their flexibility when modeling unknown distributions, as proposed by Laitila and Karlsson (2014).

To be able to detect road pavement deterioration with finite mixture models, the road network is considered as being composed of a number of segments (short lengths of contiguous road which can be characterized by a set of condition variables). When considering a time series of condition measurements we can identify two groups within the data: a "change" group, where some or all of the variables show significant change between measurement points and an "unchanged" group where the variables show little or no change. Finite mixture models are used to classify all segment into either of these groups. In literature, this type of mixture is described as an outlier density: two normal distributions with a similar mean but different variances (McLachlan and Peel, 2000, p. 14). By applying a multivariate distribution of several road condition variables, as much information as possible is used to distinguish between real changes and random errors. This could provide useful information about the rate of progression of the deterioration. Knowing which segments have a more severe deterioration, the engineer responsible for the maintenance of the road network can make a more efficient use of resources in determining which sites to maintain.

The aim of this paper is to evaluate if finite mixture models can help identify road segments that have a higher deterioration rate by clustering segments into a change group and an unchanged group.

2 Methods and Materials

2.1 Data material

The data used in this study comes from the Traffic-Speed Condition Surveys (TRACS), which is performed annually on the road network maintained by Highways England. Finite mixture models have been applied to road condition measurement data from January 2013 to January 2015, measured on parts of the M4 motorway and its junction with the A34 in west Berkshire (Junction 14). Only data from lane one, the leftmost lane in each direction, has been used. The

total length of this dataset is 88.24 kilometers divided into 10 meter segments.

TRACS surveys are carried out at traffic speed by a vehicle fitted with a number of advanced measurement systems and sensors. The survey makes measurements which are then processed and turned into a number of condition parameters, including GPS, road geometry, pavement longitudinal and transverse profile, surface texture depth, and cracking. The parameters are reported for every 10 m subsection of the Highways England road network and are stored in a condition database where they can be accessed by maintenance engineers. In spite of the best efforts of the survey contractors, auditors and researchers involved in developing the parameters and systems, it is important to understand that there are a number of sources of error in the measurements. These include systematic errors caused by variations in driving line and environmental conditions, and measurement errors because of random system variability. To minimize possible sources of systematic errors, only surveys from one single contract period is used. Within a contract period, one company is contracted to do the TRACS measurements with equipment that is calibrated and accredited according to defined criteria each year. Between contract periods, different companies using vehicles with different calibration methods can cause systematic measurement errors. For TRACS surveys, a contract period is typically four years.

Another source of systematic error is the alignment of data from successive surveys. If the data is not correctly aligned between each measurement occasion, the potential change that is found could simply occur because different sites are compared. Before the finite mixture models are fitted to the data, it has been through a process of alignment using GPS data in the first step, and profile-based alignment in the second step. The GPS alignment is a relatively coarse method, providing alignment within a couple of meters, while the profile-based alignment uses the correlation between successive years longitudinal and transverse profiles to shift the data until it is optimally aligned. This aligns the data from successive surveys to within a few centimeters, and helps ensure that any apparent change is due to changes in condition parameters, and not simply because of changes in the location of the raw data used to produce the parameters

As a reference for the change group segments, maintenance data from the Highways Agency Pavement Management System (HAPMS) is used. This data consists of sections where some part(s) of the section have been maintained. Exact GPS locations of the maintained sites were not available within the timescale of this research. The maintenance data is used to identify sections that have been maintained between measurement occasions. Segments in the change group located within a section that has been maintained have most likely changed because of maintenance, and not deterioration.

2.1.1 Road condition variables

Eleven variables, measuring different aspects of road condition, have been chosen as the basis for the clustering. The transverse profile is represented by rutting measured in the left and right wheel path respectively. The longitudinal profile is measured using enhanced longitudinal profile variance (ELPV) measured on the near side (NS) and off side (OS) of the road on maximum 3, 10 and 30 meter wavelengths. The texture is represented by the mean values of Root Mean Square Texture (RMST) measured on the near side, middle, and off side of the road. Mean and standard deviations of all variables are presented in Table 1.

3

Table 1: Road condition variables

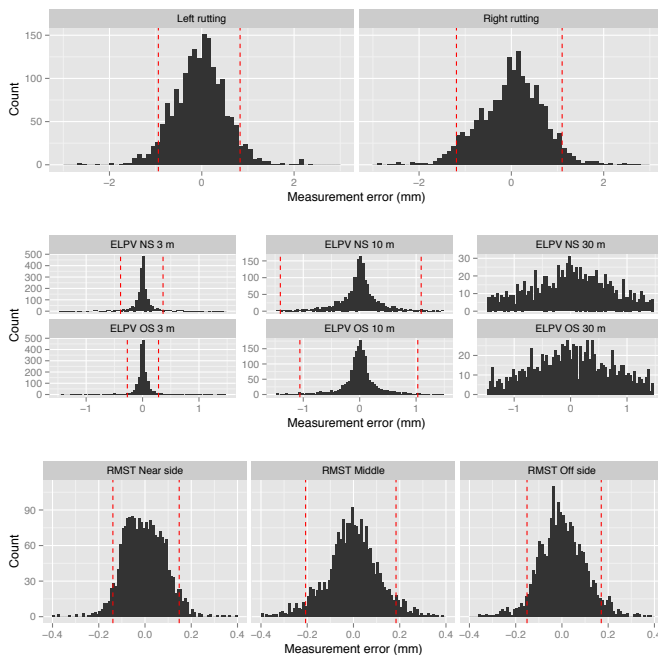| Variable	2013		2014		2015	
	Mean (mm)	Std (mm)	Mean (mm)	Std (mm)	Mean (mm)	Std (mm)
Left rutting	3.3	2.3	3.3	2.1	3.3	1.9
Right rutting	3.5	2.6	3.5	2.4	2.9	2.3
ELPV 3 m NS	0.2	0.3	0.2	0.5	0.2	0.3
ELPV 3 m OS	0.2	0.5	0.2	0.5	0.1	0.2
ELPV 10 m NS	1.0	1.5	0.8	1.4	0.8	1.3
ELPV 10 m OS	1.0	1.8	0.8	1.7	0.8	1.3
ELPV 30 m NS	8.7	10.6	5.7	6.9	5.7	6.5
ELPV 30 m OS	8.3	10.3	5.3	6.9	5.1	6.4
RMST Mean Middle	1.1	0.2	1.0	0.3	1.0	0.2
RMST Mean NS	0.8	0.2	0.8	0.2	0.8	0.2
RMST Mean OS	0.8	0.2	0.8	0.2	0.8	0.2

Figure 1: Histograms of measurement errors for all road condition variables. Red dotted lines show the 90th percentile.

4

To be able to find real change in the road condition variables, the extent of random measurement errors are of interest. These are evaluated by using measurements from a test site, where the accuracy of the survey vehicle is tested. The vehicle has surveyed all variables of a 2.2 kilometer long road section (divided into 10 meter segments) on five occasions in January 2013: the 1st, 12th, 17th, 18th and 25th. No variable is expected to show any measurable change in that time. The difference in measurements between each occasion is calculated and plotted in histograms (Figure 1). For rutting, almost all measurement errors lie within ± 2 mm and 90 percent of the error lies within approximately ± 0.96 mm for left rutting and ± 1.1 mm for right rutting. 3 m ELPV has 90 percent of the measurement errors within approximately ± 0.37 mm for near side measurements and ± 0.27 mm for off side measurements, 10 m ELPV within ± 1 mm for both near side and off side and 30 m ELPV within ± 4.5 mm for both sides. For RMST, 90 percent of the measurement errors are within approximately ± 0.14 mm for near side, ± 0.2 mm for middle and ± 0.16 mm for off side. This gives an indication of how much variability we can expect in a variable in the absence of any genuine change.

2.2   Multivariate Finite Mixture Models

Consider the road condition variables as a multivariate random variable $\mathbf{Y}_i$, $i = 1, \ldots, N$, where N is the number of 10 meter segments. In order to model the change in surface condition, the variables need to be transformed so that the initial condition does not affect the analysis This is done by letting $\mathbf{Y}_i^* = \mathbf{Y}_i - \mathbf{Y}_{\max(1, i-1)}$. Hence the first year of data, 2013, will only be used to create the new variable, which leaves two years worth of actual changes in road condition. If a longer time series were present, a longitudinal model would have been suitable. However, because of the short time series ($t = 2$), the estimation of any within-subject variation will be unstable. Even though the original road condition variables might be autocorrelated, the first difference of the measurements can be assumed to be independent (see eg. Hamilton 1994, p. 360-361 for some examples from economic variables). This assumption cannot be verified with such a short time series as only two realizations, but with the data at hand, it is a reasonable approach. Therefore, we choose to fit multivariate gaussian mixture models, assuming that the change in road condition between the two measurement occasions is a stationary i.i.d. process.

Following the structure used by Fraley and Raftery (2002), the likelihood for a multivariate mixture given data $\mathbf{y} = \mathbf{y}_i, \ldots, \mathbf{y}_n$ with K components is:

$$L_{mix}(\theta_1, \ldots, \theta_K; w_1, \ldots, w_K | \mathbf{y}) = \prod_{i=1}^{n} \sum_{j=1}^{K} w_j f_j(\mathbf{y}_i | \theta_j) \tag{1}$$

where f_j and θ_j are the density and parameters of the jth component in the mixture, and w_j is the probability that an observation belongs to the jth component ($w_j > 0$ and $\sum_{j=1}^{K} w_j = 1$). For a gaussian mixture, f_j is the multivariate normal density ϕ_j parameterized by its mean μ_j and and covariance matrix Σ_j:

$$\phi_j(\mathbf{y_i} | \mu_j, \Sigma_j) \equiv \frac{\exp\left\{-\frac{1}{2}(\mathbf{y_i} - \mu_j)^T \Sigma_j^{-1}(\mathbf{y_i} - \mu_j)\right\}}{\sqrt{\det(2\pi\Sigma_j)}} \tag{2}$$

Fraley and Raftery (2002) describe the clusters as being centered around the mean μ_j with the covariance matrix Σ_j determining the shape, volume and orientation of the clusters. If $\Sigma_j = \lambda I$,

5

giving spherical clusters of the same size, only one parameter is needed to characterize the covariance structure. For the most relaxed type of covariance structure, implying an unrestricted Σ_j, $K(d(d+1)/2)$ parameters are required for d-dimensional data. A general framework for the geometric properties of the covariance matrix was proposed by Banfield and Raftery (1993), which is a decomposition on the form:

$$\Sigma_j = \lambda_j D_j A_j D_j^T \tag{3}$$

where D_j is the orthogonal matrix of eigenvectors, A_j is a diagonal matrix with elements proportional to the eigenvalues, and λ_j is a constant of proportionality. This composition is used by Fraley and Raftery (2002) to put constraints on the cluster covariance matrix and evaluate which structure is most appropriate for the data. D_j governs the orientation of the jth component of the mixture; A_j governs the shape, and λ_j the volume.

The actual clustering is performed with model-based agglomerative hierarchical clustering as described by Fraley and Raftery (2002), in which the approximate classification likelihood is computed as:

$$L_{CL}(\theta_1, \ldots, \theta_K; \ell_1, \ldots, \ell_n | \mathbf{y}) = \prod_{i=1}^{n} f_{\ell_i}(\mathbf{y}_i | \theta_{\ell_i}) \tag{4}$$

where ℓ_i are the labels indicating the classification of each observation, i.e. $\ell_i = j$ if $\mathbf{y}_i$ belongs to the jth component. The fitting algorithm is available in the R-package `mclust` written by Fraley and Raftery (2002). Optimization is based on the EM-algorithm (Dempster et al., 1977; McLachlan and Krishnan, 2007).

2.3 Model Selection

A two-component mixture model ($K = 2$) is fitted to the $d = 11$ dimensional road condition measurement data set. The deterministic choice of the number of components is motivated by the nature of the data. The segments belonging to the unchanged group is likely to have a very narrow distribution with small variance, while the segments belonging to the change group will have a wide distribution with larger variance. The mean is not necessarily different between the two groups because not all variables measuring different aspects of road condition are likely to change at once.

There is always a possibility of identifying more clusters than what is theoretically sound. The Bayesian Information Criterion (BIC) (Schwartz, 1978) is used to evaluate the goodness of fit of several models with up to six clusters and different covariance structures. The overall highest BIC (-83125.37) was found for a model with three clusters and a covariance matrix with varying orientation, shape, and volume. A three-cluster model would have been useful if the road condition variables were all changing in the same fashion, with deterioration showing as a positive change and maintenance showing as a negative change. However, this is not the case for all of the variables. RMST can have a negative change because of both maintenance and deterioration. In a three cluster mixture model, the segments where deterioration is showing in the RMST measurements might be classified into a cluster consisting of segments showing negative change caused by maintenance in other variables.

To avoid losing any deterioration information, we will choose a model consisting of a mixture of two distributions where the change group consists of segments showing both positive and

6

negative change. The mixture models with the highest BIC among the two cluster models is the one with the most flexible covariance structure (BIC = −111840.38). Because of the nature of the two distributions described in the previous paragraph, a flexible structure between clusters in both shape, volume and orientation is theoretically valid.

3 Results

3.1 Clustering results

Of the 8824 road segments, 6136 are classified into the unchanged group and 2688 into the change group. Within the change group, 1663 of the segments had measurements that indicated change both 2013–2014 and 2014–2015. 1025 segments had measurements which indicated a change in road condition only between 2013 and 2014. Not all segments had measurements from all years because some measurement data is missing (assumed to be missing at random).

Table 2 shows the parameter estimation results for both groups. Road condition is expected to be stationary in the unchanged group, and mean values are around zero for all variables except ELPV 30 meter. The change group generally has mean values towards the negative side, indicating that the effect of maintenance is higher in most variables than the effect of deterioration. Variances for the variables in the change group are 5.7 up to 330 times larger than variances in the unchanged group, which is in line with the assumption of an underlying outlier distribution.

Table 2: Parameter estimation results

Parameter	Unchanged group $w_1 = 0.76$		Change group $w_2 = 0.24$	
	μ_1	σ_1^2	μ_2	σ_2^2
Left rutting	0.35	0.38	-0.66	8.34
Right rutting	0.16	0.23	-1.32	10.51
ELPV 3 m NS	-0.007	0.002	-0.009	0.66
ELPV 3 m OS	-0.006	0.002	-0.05	0.53
ELPV 10 m NS	-0.08	0.034	-0.22	4.39
ELPV 10 m OS	-0.08	0.034	-0.36	5.05
ELPV 30 m NS	-1.98	8.37	-4.39	97.3
ELPV 30 m OS	-1.97	8.76	-4.42	50.0
RMST Mean Middle	0.005	0.006	-0.04	0.054
RMST Mean NS	0.02	0.008	-0.15	0.072
RMST Mean OS	0.001	0.008	-0.07	0.036

Figure 2 shows the density plots of the segments clustered into either the change or the unchanged groups. Left and right rutting have long negative tails for the change group, indicating that rutting had improved for a lot of segments between 2013 and 2015 – i.e. these segments have been maintained since deterioration in rutting will always result in positive change. The distribution of the change group matches the findings from the investigation of the measurement errors (Figure 1). If we assume that an actual measurement $x_i^\star$ equals:

$$x_i^\star = x_i + \epsilon_i$$

7

where x_i is the true measurement and ϵ_i is a random error, the difference between two measurements can be written:

$$x_2^\star - x_1^\star = x_2 - x_1 + (\epsilon_2 - \epsilon_1)$$

Under the assumption that $\epsilon_i \sim N(0, \sigma_\epsilon^2)$, the distribution of $\epsilon_2 - \epsilon_1 \sim N(0, 2\sigma_\epsilon^2)$, i.e. the variance of this distribution is twice that of the original measurement error.

The 90 percentiles of the change distribution are $[-6.07 : 2.48]$ mm for left rutting and $[-7.32 : 1.60]$ mm for right rutting. This is slightly wider than the expected range (which is twice as wide as the 90 percentiles of the measurement error distributions around ± 1 mm) and slightly skewed to the negative side. However, the unchanged distribution is much narrower than what could be expected if change was occurring, with 90 percentiles ranging from $[-0.64 : 1.38]$ mm for left rutting and $[-0.68 : 0.87]$ mm for right rutting. This is an indication that the segments in the unchanged group are very unlikely to show either deterioration, maintenance or large measurement errors.

ELPV 3 m and 10 m for off side measurements are shown in the middle panel of Figure 2. The unchanged group has a much narrower interval than the measurement errors in Figure 1, with 90 percent of the observations in an interval of $[-0.08 : 0.06]$ mm for 3 m ELPV and $[-0.44 : 0.15]$ mm for 10 m ELPV (to be compared with the expected ± 0.54 mm for 3 m ELPV and ± 2 mm for 10 m ELPV). The change group interval for off side ELPV 3 m and 10 m are in line with expectations, with 90 percentile intervals of $[-0.44 : 0.27]$ mm and $[-2.45 : 1.26]$ mm respectively. The very flat density plots for the ELPV variables indicate that segments which show change in other variables are likely to show change in ELPV too.

The RMST variable for the middle part of the road has a thicker tail around -0.6 mm for the change group. The tail is slightly shifted to the right for the near side measurements, which has a small peak around -0.3 mm. This result is coherent with how the texture is expected to deteriorate: at first, the road surface is polished (RMST decreases), but later in the deterioration process stones are beginning to come out (RMST increases). The middle part of the road will always be behind the near and off sides in the deterioration process because vehicles drive there less frequently.

All change distributions have a more or less distinct peak around zero. The probabilities for each segment i of belonging to either group is based on the multivariate response variable $\mathbf{Y}_i$, and since the analysis has only three years worth of data, it is likely that not all road condition variables have changed during this time. Some segments will be clustered into the change group although they do not show any change in some variables, but a large change in other variables.
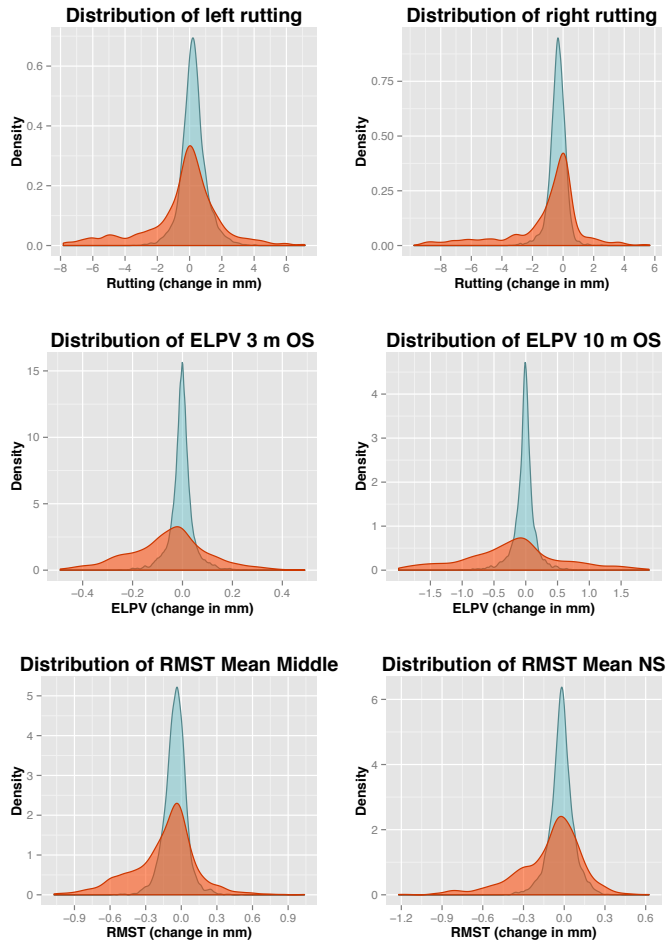
8

Figure 2: Density plots of the segments divided into the change group (red) and unchanged group (blue).

3.2 Mapping clusters

Figure 3 shows the segments that belong to the change group as red or yellow dots. Red dots represent the segments with measurements indicating change both 2013–2014 and 2014–2015, while yellow dots are segments that only showed change in 2013–2014. No segments were classified into the change group based on change in road condition measurements only

9

between 2014–2015. The change group may include segments that have changed either because of deterioration or maintenance, where maintenance often results in a greater change in road condition. In order to determine if the classification is of any value in practice, data regarding maintenance activities was plotted in the road network as shown in Figure 4.
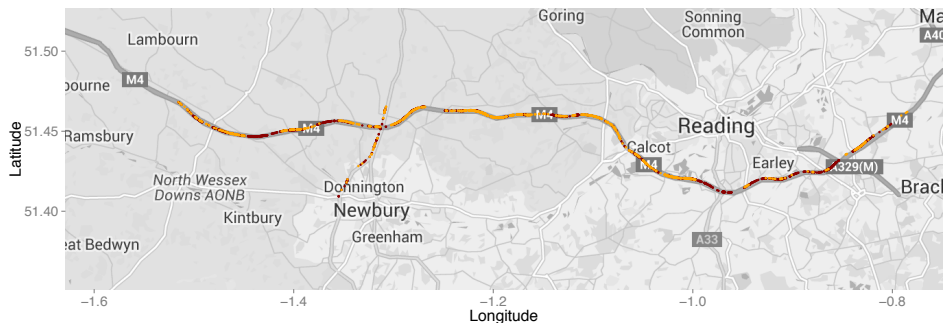

Figure 3: Map of the M4/A34 intersection with yellow and red dots representing segments in the change distribution.

Figure 4 shows a close-up on a junction on the M4. Maintenance that occurred 2013 and 2014 are drawn as blue lines and maintenance that occurred after January 2015 is drawn as green lines. The Highways England network is split into sections, each of which is typically 1–2 km long. Because the available information about maintenance only provided the section where the maintenance work took place, and not the GPS coordinates of the maintenance location, the blue and green lines cover longer sections than where the maintenance was carried out. However, if the maintenance data and the change group segments correlate, this is an indication that the classification can capture real changes.

The red segments beneath the blue line in Figure 4 are likely to have changed because of maintenance. A look at the raw data for these segments shows large negative changes in almost all variables for these segments between 2013 and 2014; e.g. 6-11 mm for rutting, 0.5-1.5 mm for ELPV 3 m and 0.6-1.5 mm for RMST, which are way outside the measurement error interval. The data from 2015 for these segments is missing.

The green lines represent maintenance activities that took place after the last road condition survey in January 2015. That is, the red segments beneath the green lines have not changed because of previous maintenance; they were maintained shortly after the road was surveyed in January 2015. The fact that these segments were maintained soon after the survey implies that they were deteriorating and that the change detected reflects genuine deterioration. Previous research described in e.g. the Highway Development and Management Model (HDM-4) and its precursor the Highway Design and Maintenance Standards Model (HDM-III) provided by the World Bank (Watanatada, 1987; Odoki and Kerali, 2000) shows that initial deterioration is often very low, and after a certain amount of time it increases rapidly in an exponential rather than linear fashion. If a segment is identified as a maintenance candidate object because of bad condition, it has often shown an increasing deterioration pattern.

Looking at the raw data from the segments under the green line, there are positive changes in the range of 2-5 mm for rutting, a strong indication of genuine deterioration. (Figure 1 shows

10

that the amount of expected variability in rut measurements without deterioration was of the order of 1-2 mm). The segments also show negative changes for RMST of 0.3-0.8 mm, which is slightly outside the expected measurement error interval. Negative change can imply both maintenance and deterioration for RMST. The ELPV variables for these segments have only small changes within the expected measurement error interval.

The example in Figure 4 indicates that both changes because of maintenance and changes because of deterioration can be captured by the finite mixture model.

Figure 4: Map of a part of the M4 highway at Junction 14. Lane 1 in each direction is used in the analysis.

4   Discussion

With only three years' worth of data, the dependency structure between road condition measurements cannot be evaluated. It is assumed in this paper that the first difference between measurements is a stationary i.i.d. process. With measurements from more years available, it will be possible to check this assumption, and if it does not hold, a longitudinal model which takes within-subject correlation into account may be used.

The mean vector μ_2 is negative for all road condition variables in the change group. The result is expected because change in road condition caused by maintenance activities (which can

11

account for most, but not all, negative change) is likely to be greater than change caused by deterioration. However, negative change in road condition might also be a sign of deterioration for some variables; in this data that is the case for RMST. If it was certain that all variables showed positive change because of deterioration and negative change because of maintenance, a mixture of three distributions would be worth considering: change caused by deterioration and/or large positive measurement errors, unchanged, change caused by maintenance and/or large negative measurement errors. However, because negative change in RMST can occur because of both maintenance and deterioration this variable could not be included in such a model. RMST is one of the most sensitive variables (in the sense that it is possible to measure deterioration over a period of 1-2 years) and therefore it would be unwise to disregard it. When the mixture models consist of two distributions, defined as a change and an unchanged group, complementary data from maintenance records are necessary to distinguish between change caused by maintenance or deterioration, and to validate the model performance.

The findings of this study also reveal another possible use of the finite mixture models: identifying maintenance activities. A way of doing this is to look at e.g. the development of rutting for segments in the change group, where negative change (i.e. less rutting) in road condition on several adjacent segments is very likely to occur because of maintenance. Reporting errors and missing data are not uncommon in maintenance databases, and this measure could be useful when looking at historical road condition data to identify or verify when maintenance has taken place.

If maintenance data with coordinates is available, it is straightforward to identify segments that have changed because of previous maintenance activities. For the data used this paper, only start and end coordinates of a section (consisting of several 10 m segments) is provided which limits the precision. Apart from maintenance and deterioration, it is also possible that large measurement errors are identified as change. The finite mixture model is not an automatic deterioration detection, but rather a way to identify segments that deviate from a stationary deterioration pattern. As a future research topic, combining predictive pavement deterioration models with finite mixture models could be a way of further explore the possibilities of finding early deterioration.

5   Conclusion

The road segments from the M4 were classified into either a change group or an unchanged group using finite mixture models. The distribution of the segments in the unchanged group showed no sign of deterioration. The segments in the change group were compared with data from maintenance records. It could be concluded that the change group most likely consists of segments that have changed because of either maintenance activities or deterioration since they matched locations of both previous and (at the time of the measurement occasion) future maintenance.

The conclusion of the study is that finite mixture models are successful in identifying segments with a road condition that is deviating from the stationary state of a road surface that is not maintained or deteriorating. Identifying such segments can be helpful in order to prioritize possible maintenance candidate objects when resources are limited, and thus make maintenance decisions more efficient. However, it is not possible to fully distinguish between change because of maintenance, deterioration, or severe measurement errors using the finite mixture models approach on the current variables without complementary data sources.

References

Banfield, J. D. and Raftery, A. E. Model-based gaussian and non-gaussian clustering. *Biometrics*, 49:803821, 1993.

Daeyoung, K. and Lindsay, B. G. Empirical identifiability in finite mixture models. *Annals of the Institute of Statistical Mathematics*, 67(4):745–772, 2015.

Defays, D. An efficient algorithm for a complete link method. *The Computer Journal*, 20(4): 364–366, 1977.

Dempster, A. P., Laird, N. M., and Rubin, D. B. Maximum likelihood from incomplete data via the em algorithm. *Journal of the royal statistical society. Series B (methodological)*, pages 1–38, 1977.

Fraley, C. and Raftery, A. E. Model-based clustering, discriminant analysis and density estimation. *Journal of the American Statistical Association*, 97:611–631, 2002.

Hamilton, J. D. *Time series analysis*, volume 2. Princeton university press Princeton, 1994.

Hartigan, J. A. and Wong, M. A. Algorithm as 136: A k-means clustering algorithm. *Journal of the Royal Statistical Society. Series C*, 28(1):100–108, 1979.

Hong, F. and Prozzi, J. A. Estimation of pavement performance deterioration using bayesian approach. *Journal of Infrastructure Systems*, 12(2):77–86, 2006. doi: 10.1061/1076-0342.

Laitila, T. and Karlsson, M. Finite mixture modeling of censored regression models. *Statistical Papers*, 55(3):627–642, 2014.

Lee, S. Q., Lartillot, N., and Gascuel, O. Phylogenetic mixture models for proteins. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1512):3965–3976, 2008.

Lethanh, N. and Adey, B. T. Use of exponential hidden markov models for modelling pavement deterioration. *International Journal of Pavement Engineering*, 14(7):645–654, 2012. doi: 10.1080/10298436.2012.715647.

McLachlan, G. and Krishnan, T. *The EM algorithm and extensions*, volume 382. John Wiley & Sons, 2007.

McLachlan, G. and Peel, D. *Finite Mixture Models*. John Wiley & Sons, New York, 2000.

McLachlan, G. J. and Basford, K. E. *Mixture Models. Inference and applications to clustering*. Dekker, New York, 1988.

Odoki, J. and Kerali, H. G. *Highway Development and Management Series: Analytical Framework and Model Descriptions*. World Road Association (PIARC), 2000.

Pagel, M. and Meade, A. A phylogenetic mixture model for detecting pattern-heterogeneity in gene sequence or character-state data. *Systematic Biology*, 53(4):571–581, 2004.

Park, B.-J. and Lord, D. Application of finite mixture models for vehicle crash data. *Accident Analysis & Prevention*, 14(4):683–691, 2009. doi: 10.1016/2009.03.007.

Schwartz, G. Estimating the dimenson of a model. *The Annals of Statistics*, 6:461–464, 1978.

13

Sibson, R. Slink: an optimally efficient algorithm for the single-link cluster method. *The Computer Journal*, 16(1):30–34, 1973.

Steinhaus, H. Sur la division des corp materiels en parties. *Bull. Acad. Polon. Sci*, 1:801–804, 1956.

Ward Jr, J. H. Hierarchical grouping to optimize an objective function. *Journal of the American statistical association*, 58(301):236–244, 1963.

Watanatada, T. *The Highway Design and Maintenance Standards Model: User's manual for the HDM-III model*, volume 2. Johns Hopkins University Press, 1987.

14

Paper V

The effect on running time of temporary speed restrictions on a Norwegian railway line

Kristin Svenson*, Andreas Amdahl Seim†, Andreas Dypvik Landmark‡

Abstract

The effect on train running time of temporary speed restrictions varies substantially due to empirical factors such as driver behavior and location of the speed restricted segment. Lacking information about this variation has a negative impact on planners' ability to adjust timetables when a speed restriction is enforced. To estimate empirical effects of temporary speed restrictions, a generalized linear mixed gamma model was applied to data from a Norwegian railway. Results showed that the average effect of a temporary speed restriction is nine percent longer running time between two stations. On station blocks where the normal speed is relatively high, temporary speed restrictions had limited impact on running time. On station blocks where the normal speed is low, speed restrictions resulted in up to three times longer running time. When feasible, drivers tend to increase the speed on the remaining part of the block to reduce speed restrictions' impact on running time. Length of the speed restriction had a very small effect on running time for freight trains, implying that acceleration is more critical. The importance of acceleration was emphasized by the result that a speed restriction before an uphill increased running time for freight trains with twelve percent.

Keywords: Railway Maintenance; Temporary Speed Restriction; Generalized Linear Mixed Model

1 Introduction

Punctuality in railway networks is considered a key performance indicator of the railway industry. There are several possible definitions of punctuality. Rudnicki (1997) defines it as "a feature consisting in that a predefined vehicle arrives, departs or passes at a predefined point at a predefined time". To achieve punctuality, the running time of a vehicle must be predictable, and the timetable must be scheduled accordingly. Scheduling problems with the objective to optimize network capacity and punctuality is a well-studied topic (see e.g. Mees 1991; Higgins et al. 1996; Cordeau et al. 1998; Dariano et al. 2007; Krasemann 2012). The theoretical running time is a deterministic function of distance and speed. However, in real applications, the running time of a train is always affected by more or less predictable external factors. A large field of study is the identification of factors that do have an influence on actual running time, and ultimately punctuality.

In a review of influencing factors on punctuality, Olsson and Haugland (2004) identify the following: number of passengers, occupancy ratio (passengers/seats), infrastructure capacity utilization, cancellations, temporary speed restrictions, railway construction work, departure

---

*School of Technology and Business Studies, Dalarna University, Borlänge, Sweden. E-mail: kss@du.se
†The Foundation for Scientific and Industrial Research (SINTEF), Trondheim, Norway
‡SINTEF, Trondheim, Norway

1

and arrival punctuality, and operational priority rules. Out of these factors, a temporary speed restriction (TSR) – many of which occur due to maintenance work on the railway – is almost certain to have an effect on running time. To reduce the effect of a TSR on punctuality, it is of interest to be able to plan these temporary speed restrictions so that the impact on network level is minimized. Maintenance planners may want to know whether one longer or several shorter segments with reduced speed have a greater impact on the expected running time. Theoretical deterministic models of TSR typically include running time as a function of the speed restriction, normal speed, length of TSR segment, an acceleration coefficient and a deceleration coefficient, which may vary depending on train type (Dingler et al. 2009; Shafiullah et al. 2010). However, few studies have confirmed these approximations using empirical data. In reality, the assumed normal speed, speed restriction and acceleration coefficients may have a great variability between station blocks, weeks, days, and individual trains. The nature of the TSR segment (i.e. curvature or height profile), or potential seasonal effects is not considered in the theoretical model. Also, a deterministic model cannot capture any potential effect of the train drivers' behavior. As stated by e.g. Ochiai and Tomii (2015), drivers may try to reduce the effect of a temporary speed restriction by increasing the speed on the remaining, non-speed restricted part of the station block.

This paper will study the empirical effect of TSRs within the Norwegian railway network by fitting a generalized linear mixed model (GLMM) to train travel and speed restriction data. The benefit of generalized linear mixed models compared to standard linear regression models is the ability to quantify variation among units (McCullagh and Nelder 1989; McCulloch and Neuhaus 2001; Pawitan 2001; Lee et al. 2006; Bolker et al. 2009). In the case of a temporary speed restriction, the effect on running time may vary greatly between different station blocks. GLMMs can also account for data that is non-normal (i.e. skewed), which usually applies to running times.

1.1 Research objective

The aim of this study is to model the effect of temporary speed restrictions on train running time. The data originates from the Norwegian railway network and variables included in the model are season, weekday, train type, speed restriction, the length and height profile of the speed restricted segment, and it's distance to the next station. In order to validate and quantify the effect of the different variables, a generalized linear mixed model approach is used.

1.2 Previous research

Few studies have examined the empirical effects related to maintenance and temporary speed restrictions within a railway network. Budai et al. (2006) create an algorithm which optimizes scheduling of preventive maintenance on railways, assuming that a section where maintenance is carried out is blocked for traffic. However, they do not take into account that some maintenance work may imply temporary speed restrictions before, during, and shortly after the maintenance activity.

Olsson et al. (2002) perform a correlation analysis of temporary speed restrictions and punctuality on the Norwegian Nordland line (Nordlandsbanen) during 13 weeks in the spring of 2002. They find a significant negative correlation between TSR and punctuality of -0.11 in the northbound direction, and no significant correlation in the southbound direction. One explanation for this counterintuitive result is the 4 percent increase in running time included in Norwegian timetables. Olsson assumes that there exists a threshold effect of TSR on punctuality, but this threshold is not investigated in the study.

2

Gorman (2009) estimates total train running time with linear regression in order to explain congestion. Slow orders (i.e. speed restrictions) are one of the explanatory variables. Slow orders (the total number of slow order minutes on a particular date) showed a significant, positive effect on train running time; implying that the existence of slow orders increases train running time. The definition and use of slow orders as an input variable to estimate running time derive from Krueger (1999), who developed a Parametric Capacity Model for the Canadian National Railway. Gorman advocates the use of empirical estimation and statistical methods over simulations, as they are less laborious and computationally intensive. He also states that "[...] statistical analysis of the physical system leads to insights to complement the understanding of the system based on other methods."

Kecman and Goverde (2015) use three different data-driven predictive models (least trimmed squares robust linear regression, regression trees, and random forest) to estimate train running and dwell times. They also use local models for particular train lines and station blocks, concluding that local models usually are more accurate than global models. However, the ability to generalize the results from a global model is appealing and useful, especially when data is scarce.

2 Method and materials

2.1 Data material

The data material used in this study consists of 42,382 unique train travels on the Dovre line (Dovrebanan) between stations Dovre and Hove (Figure 1) from December 12, 2009, until December 12, 2015. A train travel is defined as a scheduled trip between a start and an end station, which may include stops at stations in between. A train travel is either northbound or southbound and has a unique ID number. When a train passes a station without stopping, the departure time from the station is recorded. If the train stops, the arrival and departure times from the station are recorded. Running time of the train is defined as the time between either departure from the previous station to departure from next station (for trains that only pass the station), or departure from the previous station to arrival at next station (for trains that stop at the station). In total, the dataset contains 676,495 observations of train running times on 29 station blocks. Each block is either southbound or northbound direction.

The train travel data was combined with data regarding temporary speed restrictions during the same period. There are 72 unique segments with a temporary speed restriction, located within different station blocks, that may last from one day to several weeks. These restrictions affect 106,055 observations (i.e. running times), of which 15,336 are temporary speed restrictions within station areas.
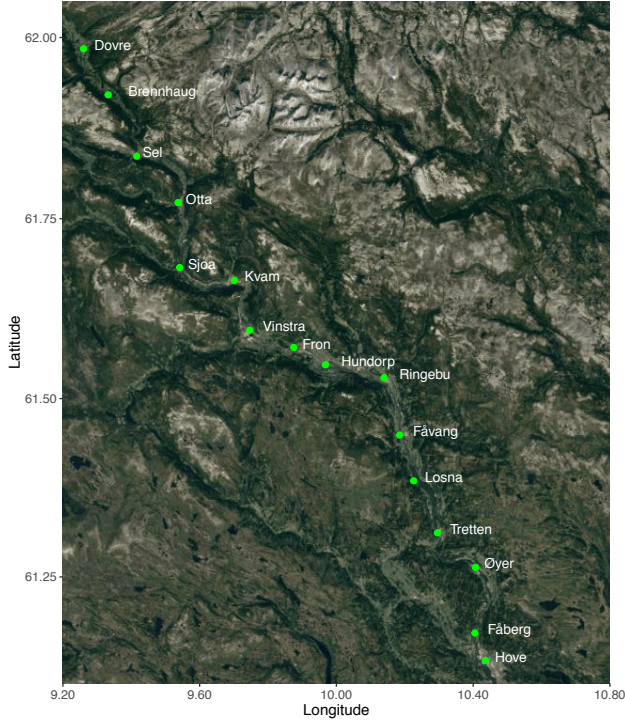
3

Figure 1: Map of Dovre line (Dovrebanan).

The qualitative variables in the model are season, TSR at station, train type, and weekday. Season is defined as a binary: winter (November – March) or summer (April – October). TSR at station implies that there is a temporary speed restriction within the station area. Train type is classified as either freight train or passenger train. Weekend is defined as Saturday and Sunday, while Monday – Friday are weekdays. Because only a small share of all trains had experienced a TSR en route, an indicator variable (No TSR) was introduced to account for potential non-linear effects of the TSR-variables.

Among the quantitative variables, normal speed is calculated as the average speed between two stations on a week with no temporary speed restrictions on that block. Δspeed is defined as the difference between normal speed and the temporary speed restriction limit. The variable TSR height difference is created to capture the possible effect of an uphill just after a TSR segment, where the trains accelerate up to normal speed again. TSR height difference is defined as the sum of positive differences in vertical meters between the following two kilometers after a TSR segment. Up to two kilometers may be necessary for a freight train to gain full speed.

4

Table 1: Summary statistics, qualitative variables.

Variable	True	False
Winter	274,975	403,959
TSR at station	15,336	663,598
Freight train	341,451	337,483
Weekend	96,111	582,823

Table 2: Summary statistics, qualitative variables.

Train type	Running time (min)				Normal speed (km/h)			
	Mean	Std	Min	Max	Mean	Std	Min	Max
Freight train	8.0	3.7	1.2	59.9	56.6	10.1	15.7	87
Passenger train	7.4	3.0	1.2	59.7	70.1	14.1	27.9	98
	ΔSpeed TSR (km/h)				TSR length (km)			
	Mean	Std	Min	Max	Mean	Std	Min	Max
Freight train	13.8	9.91	0	42.6	1.24	1.45	0.20	4.65
Passenger train	13.7	12.24	0	57.1	1.24	145	0.20	4.65
	TSR height difference (m)				TSR distance from station (km)			
	Mean	Std	Min	Max	Mean	Std	Min	Max
Freight train	7.7	16.7	0	86	4.7	2.4	0	8.9
Passenger train	7.8	16.6	0	86	4.7	2.4	0	8.9

2.2   Generalized Linear Mixed Models

A generalized linear mixed model was chosen to evaluate the effect of temporary speed restrictions on the Dovre line. A mixed model adjusts for data that are correlated in clusters by estimating random effects of the cluster units. On the Dovre line, the running time of different trains between two particular stations in a particular direction will be highly correlated. To account for this correlation, a random intercept was fitted to each station block (in total 30 section blocks). Random effects can also be used to fit unit specific slopes for the explanatory variables. In this sense, a mixed model can fit both a global (fixed effects) model and a local (random effects) model.

The generalized linear mixed model (GLMM) is an extension of the linear mixed model and allows for a response variable a from non-gaussian distribution. In the GLMM, the linear predictor is linked to the mean response through a link function, such as the identity, inverse, or logarithm. Running time is often modelled as a log-normal distribution, mainly because it is left truncated (i.e. non-negative) and can have a possibly thick tail (see e.g. Calfee et al. 2001, Rakh et al. 2010, Yuan et al. 2010, Westgate et al. 2013). However, distributions such as gamma (Polus, 1979) – which is also skewed and bounded by zero – and normal are also common (Noland and Polak, 2002).

Three competing models were fitted to the data, assuming normal, gamma, and log-normal distributed data. The models were evaluated based on residual plots and the Akaike Information Criterion (AIC) (McCullagh and Nelder 1989; McCulloch and Neuhaus 2001; Jiang 2007; Verbeke and Molenberghs 2009). Residuals were slightly skewed in all three models with the AIC favoring the gamma distribution. The gamma distribution exhibits a variance that is pro-

5

portional to the square of the mean response (i.e. variation increases as running time increases). This is an assumption that may not be applicable for all running times. Stefánsson (1996) proposes that the slope of a linear regression with gamma distributed data (here: the running time) in a log-log plot of variance vs. mean of a predefined homogeneous unit should be close to two (theoretically, $\log(\text{variance}) = 2 \times \log(\text{mean})$). A fairly homogeneous unit along a railway should be the running time between the station blocks. The log-log plot (Figure 2) shows a slope which is close to two, although the linear fit is not perfect ($r^2 = 0.53$).
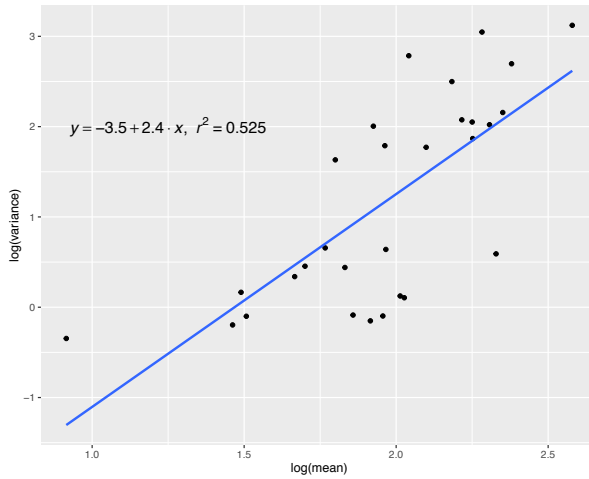

Figure 2: Scatterplot and regression line of log-variance vs. log mean for each station combination.

The canonical link function of the gamma distribution is the inverse. However, as this can be a difficult link for computational and interpretational reasons, the much more common log-link was chosen instead. This link implies a multiplicative effect on running time of the explanatory variables. With N observations, $\mathbf{y}$ (length $N \times 1$) is the response vector of running times between stations. Conditional on the explanatory variables, $\mathbf{y}|\mathbf{b} \sim Ga(\mu, \phi\mu^2)$ where the Gamma distribution is parameterized using the mean and the variance, respectively. ϕ is the dispersion parameter. A gamma GLMM with p fixed effects (one for each explanatory variable and possible interactions) and q random effects (one for each unique combination of stations) is written as:

$$\log(\boldsymbol{\mu}) = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{b} \tag{1}$$

where $\mathbf{X}$ is a $N \times p$ design matrix for the fixed effects, $\mathbf{Z}$ is a $N \times q$ design matrix for the random effects, $\boldsymbol{\beta}$ is a $p \times 1$ vector of fixed effects, $\boldsymbol{b}$ is a $q \times 1$ vector of random effects, $\boldsymbol{b} \sim N(0, \mathbf{G})$. Because not all station blocks have had a speed restriction, potentially random slopes are nested within the binary factor of a TSR occurring or not. The variance-covariance matrix $\mathbf{G}$ will be a diagonal matrix:

6

$$\mathbf{G} = \begin{bmatrix} \sigma_{b,int}^2 \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \sigma_{b,slope}^2 \mathbf{I} \end{bmatrix} \tag{2}$$

where $\sigma_{b,int}^2$ is the variance of the random intercepts and $\sigma_{b,slope}^2$ is the variance of the random slopes. For computational reasons, the random effects are assumed to be uncorrelated. As shown in Figure 3, there is no evident correlation between the intercept and the slope for the thirteen station blocks where a TSR is present.
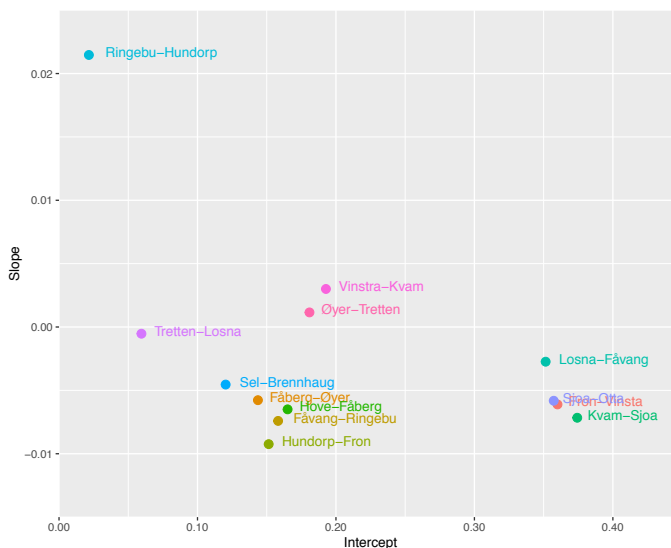

Figure 3: Correlation between random slopes and random intercepts for all station blocks with at least one TSR.

The likelihood function is crucial to obtain the maximum likelihood estimates of the parameters. The likelihood, i.e. the marginal density of $f(\mathbf{y})$, is:

$$f(\mathbf{y}) = \int f(\mathbf{y}, \mathbf{b}) d\mathbf{b} = \int f(\mathbf{y}|\mathbf{b}) f(\mathbf{u}) d\mathbf{b} \tag{3}$$

which can be difficult to evaluate since $f(\mathbf{y}|\mathbf{b})f(\mathbf{u})$ is a very complex function in a GLMM. In a linear mixed model, the marginal distribution of $\mathbf{Y}$ can be computed directly as a multivariate normal. In the case of a gamma GLMM, an approximate optimization algorithm needs to be applied in order to obtain the maximum likelihood estimates of the parameters.

The model in (1) was fitted both in R-package `lme4` which uses Laplace approximation (Bates et al., 2015), and in R-package `MASS` which uses penalized quasi-likelihood (Venables and Ripley, 2002). The Laplace approximation had a slow convergence, but the penalized quasi-likelihood converged, and both methods gave very similar results regarding both the fixed effects and the residual and random effect variances.

7

3 Results

3.1 Fixed effects

Table 3: Maximum likelihood estimates.

Parameter (p)	Parameter Estimate (β_p)	Standard Error	T-value	P-value	
Intercept	1.44	0.05	32.0	< 0.0001	***
Winter	-0.026	0.0006	-45.6	< 0.0001	***
Weekend	-0.017	0.0008	-21.7	< 0.0001	***
Distance (km)	0.07	0.0008	92.2	< 0.0001	***
No TSR	-0.09	0.003	-29.7	< 0.0001	***
Freight train	0.09	0.0006	145.0	< 0.0001	***
TSR station	-0.006	0.006	-1.12	0.26	
ΔSpeed (km/h)	0.003	0.004	0.92	0.36	
TSR length (km)	0.015	0.0009	15.2	< 0.0001	***
TSR height difference (>0 m)	-0.015	0.003	-5.5	< 0.0001	***
TSR height difference (>15 m)	-0.10	0.004	-24.9	< 0.0001	***
TSR distance from station (km)	-0.007	0.0005	-13.7	< 0.0001	***
TSR station×Freight train	-0.06	0.006	-9.9	< 0.0001	***
TSR ΔSpeed×Freight train	0.003	0.0001	23.6	< 0.0001	***
TSR length×Freight train	-0.015	0.001	-15.3	< 0.0001	***
TSR height diff. (>0 m)×Freight train	-0.026	0.003	-7.9	< 0.0001	***
TSR height diff. (>15 m)×Freight train	0.21	0.005	41.4	< 0.0001	***
TSR distance from station× Freight train	-0.016	0.0004	-31.2	< 0.0001	***
Dispersion parameter: ϕ	0.082				
Random intercept variance: $\sigma^2_{b,int}$	0.0053				
Random slope variance: $\sigma^2_{b,slope}$	0.000017				
Significance codes: < 0.0001 ***, 0.001 **, 0.01 **, 0.05 .					

Because the chosen link function of the gamma GLMM is the natural logarithm, regression coefficients have a multiplicative effect on running times, i.e. the effect on running time of a one unit change in any of the explanatory variables is e^{β_p}. Running time increased with $100 \times (e^{0.07} - 1) = 7$ percent per kilometer traveled, and on average, freight trains have $100 \times (e^{0.09} - 1) = 9$ percent longer running times than passenger trains. Running times in winter and at weekends are approximately 2 percent less than in summer and on weekdays. The estimates of the TSR variables are explained in detail in the Discussion section.

3.2 Random effects

In order to capture the between-station block effect on running time of a TSR and create a local model for each station block, a random intercept and a nested random slope for the variable ΔSpeed were included in the model. The nesting implies that a random slope was fitted only to station blocks with at least one TSR. The random effects can be used to make local rather than global predictions – i.e. the running time between two specific stations can be predicted under various Δspeed.

8

A likelihood ratio test between a model with a random intercept only and a model with a random intercept and a nested random slope for the variable Δspeed strongly favored the latter model (χ^2-statistic of 5017 on 2 degrees of freedom using approximate Laplace likelihood estimation).

Figure 4 shows how predicted running times of a freight train depends on the speed reduction for six different station blocks. The explanatory variables for the fixed effects are set according to the specification in Table 4, with distance set to the actual distance between stations.
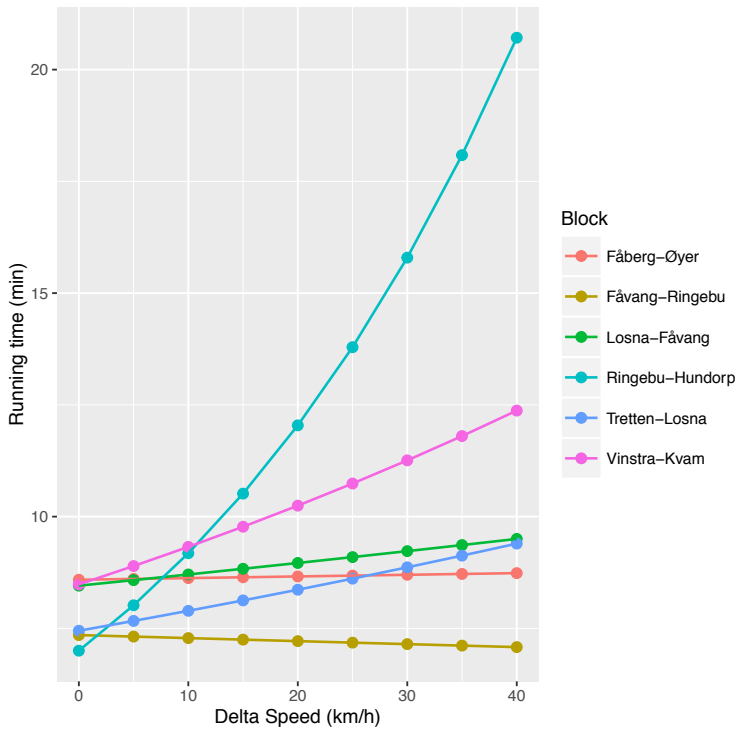

Figure 4: Predicted running times depending on Δspeed for six between-station blocks.

9

Table 4: Values in predictive data.

Variable	Value
Winter	FALSE
Weekend	FALSE
No TSR	FALSE
Distance (km)	5.8-10.2
TSR distance fr. station (km)	4
Freight train	TRUE
TSR station	FALSE
TSR length (km)	0.5
TSR height difference (m)	1-14

4 Discussion

4.1 Fixed effects (global model)

We studied the effect of TSRs on running time for trains on the Dovre line between December 12, 2009, and December 12, 2015, using a generalized linear mixed model approach. The model successfully quantified how a number of different variables modify the effect of TSRs.

For the Dovre line, we found that running times in winter and at weekends were approximately 2 percent less than in summer and on weekdays. Likely explanations are the fewer maintenance actions taking place during the winter months (November – March) and comparatively lower traffic volume during weekends.

The binary variable "No TSR" was found to be significant, and showed that trains which are not subject to any temporary speed restrictions have on average had nine percent shorter running times than trains that are affected by a TSR.

For passenger trains, neither a temporary speed restriction at a station or the effect of Δspeed were significant. However, Δspeed was also introduced as a random slope to account for variation between station blocks. The non-significant global effect but significant between-block variation implies that the effect of Δspeed on running time varies substantially between station blocks.

Each kilometer in length of a TSR segment increased running time with 1.5 percent for passenger trains. This effect is rather small and indicates that the distance the train is required to drive at reduced speed is less important than the impact of acceleration/deceleration. The effect of a positive height difference just after the TSR was evaluated as a categorical variable because the effect is assumed to be non-linear. For passenger trains, both categories showed a slightly negative effect. This implies that the uphills after the TSRs in our data set on the Dovre line did not affect the acceleration capacity of passenger trains.

The variable "TSR distance from station" was significant with a negative sign, implying that the further away from the next station the temporary speed restriction ends, the shorter are the train running times. This result is another indication of the importance of acceleration: if a TSR is very close to a station, the driver will not be able to accelerate up to normal speed before entering the station area.

It can be assumed that freight trains, being heavier and having a slower acceleration, are affected differently by TSR than lighter passenger trains with faster acceleration. Therefore, interaction effects were included in the model between freight trains and all TSR variables. All interaction effects were highly significant which confirms the assumption.

10

The interaction effect of freight train and Δspeed had a small but significant effect on running time. This implies that on average, the magnitude of the speed restriction affects freight trains more than passenger trains. The length of the TSR segment had a small negative effect on running time for freight trains, giving an even stronger indication that the length of the TSR does not have a great impact on running time. Once freight trains have reduced their speed, the length of the speed restriction is of less importance. For freight trains, a positive height difference of more than 15 meters on the two kilometers adjacent to a TSR segment increased running time with $100 \times (e^{0.21-0.1} - 1) = 12$ percent. Being heavier, it is reasonable that the acceleration capacity of freight trains is more affected by uphills than passenger trains. If a TSR is located just before an uphill, the timetable should be adjusted to account for the slower acceleration of freight trains.

Freight trains have an even larger effect of the variable TSR distance from station than passenger trains, with a $100 \times (e^{-0.007-0.016} - 1) = -2.3$, i.e. a 2.3 percent decrease in running time for every kilometer between the TSR segment and the next station. This increased effect is also most likely due to the slower acceleration of freight trains – freight train drivers can catch up time by driving faster, but only if they have enough space to do so.

4.2   Random effects (local model)

The presence of significant variation between station blocks suggests that there are certain features associated with each unique block that is not captured by the fixed effects. One such feature might be the curvature of the railway or other station-specific characteristics.

Between some stations, the effect on running time of the speed restriction captured by Δspeed was found to be very limited. Two examples of such blocks are illustrated in Figure 4 by Fåberg–Øyer and Fåvang–Ringebu, where the latter block even had a slightly negative slope. On other station blocks, the speed restriction had a large impact on running time. Most notably is the block between Ringebu–Hundorp where a speed reduction of 40 km/h on a 500 meter long segment increased running time from 7 to 21 minutes.

It is possible that the train driver can reduce the effect of a TSR by driving faster on the remaining part of the station block. The mean normal speed of freight trains on the two blocks where the effect of Δspeed was limited is among the highest on the Dovre line: 63.9 km/h on Fåberg–Øyer and 65.6 km/h on Fåvang–Ringebu. On blocks where Δspeed had a large impact, mean normal speed is among the lowest – 50.9 km/h on Ringebu–Hundorp and 53.7 km/h on Vinstra–Kvam. This implies that between stations where the mean normal speed is higher, the driver has the possibility to increase speed on the non-speed restricted parts of the station block. On blocks where it is not feasible to drive very fast under normal circumstances, a speed restriction will have a stronger effect on the running time because drivers have less opportunity to increase their speed elsewhere.

4.3   Limitations

The Dovre line was chosen as a case study because 15 percent of all running times 2010–2015 had been affected by a temporary speed restriction. These TSRs were spread over several different station blocks. On other Norwegian railway lines, e.g. Bergenbanen and Sørlandsbanen, only 2-3 percent of the running times were affected by a TSR the same period, partly because TSR data was missing after 2013. Lack of TSR data makes it difficult to find significant effects in a global model. It also has implications for a local model. With few TSRs on a station block, each individual TSR will have a very large impact on the results. This limitation also applies to the Dovre line: TSRs for certain station blocks are found only in a few locations at a specific point in time. From an inference perspective, TSRs would ideally have occurred over the entire

block, varying in time, length, speed restriction, etc. However, in practice, this is not the general case. A TSR at a sensitive location on a station block can have a large impact on running time, whereas a different TSR at a less sensitive location on the very same station block can have almost no impact. With a limited number of TSRs per station block, the parameter estimates become less robust to these variations and generalizations should be made with care.

4.4   Future research

As a future research topic, a hierarchical model that evaluates how temporary speed restrictions affect train running times on an entire railway line – i.e. from start station to end station – is of interest. Another possibility is to add a random effect for each train travel, which would quantify how the variation in running time depends on features connected to individual trains. However, this would require estimation of over 40,000 random effects in a dataset like the one for the Dovre line, and it would take work on the computational capacity to fit such a model.

5   Conclusion

The results of the study can be used to understand how empirical effects – TSR location, the height profile of the railway track, and driver behavior – will affect running time when a temporary speed restriction is imposed on a station block. However, a sufficient amount of TSR data spread over different station blocks is crucial in order to obtain robust results in a global model. Generalizations based on the local model, i.e. random effects, should be done with care since single TSRs can be highly influential.

References

Bates, D., Mächler, M., Bolker, B., and Walker, S. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1):1–48, 2015. doi: 10.18637/jss.v067.i01.

Bolker, B. M., Brooks, M. E., Clark, C. J., Geange, S. W., Poulsen, J. R., Stevens, M. H. H., and White, J.-S. S. Generalized linear mixed models: a practical guide for ecology and evolution. *Trends in ecology & evolution*, 24(3):127–135, 2009.

Budai, G., Huisman, D., and Dekker, R. Scheduling preventive railway maintenance activities. *Journal of the Operational Research Society*, 57(9):1035–1044, 2006.

Calfee, J., Winston, C., and Stempski, R. Econometric issues in estimating consumer preferences from stated preference data: a case study of the value of automobile travel time. *Review of Economics and Statistics*, 83(4):699–707, 2001.

Cordeau, J.-F., Toth, P., and Vigo, D. A survey of optimization models for train routing and scheduling. *Transportation science*, 32(4):380–404, 1998.

Dariano, A., Pacciarelli, D., and Pranzo, M. A branch and bound algorithm for scheduling trains in a railway network. *European Journal of Operational Research*, 183(2):643–657, 2007.

Dingler, M., Lai, Y.-C., and Barkan, C. Impact of train type heterogeneity on single-track railway capacity. *Transportation Research Record: Journal of the Transportation Research Board*, (2117):41–49, 2009.

Gorman, M. F. Statistical estimation of railroad congestion delay. *Transportation Research Part E: Logistics and Transportation Review*, 45(3):446–456, 2009.

Higgins, A., Kozan, E., and Ferreira, L. Optimal scheduling of trains on a single line track. *Transportation research part B: Methodological*, 30(2):147–161, 1996.

Jiang, J. *Linear and generalized linear mixed models and their applications*. Springer Science & Business Media, 2007.

Kecman, P. and Goverde, R. M. Predictive modelling of running and dwell times in railway traffic. *Public Transport*, 7(3):295–319, 2015.

Krasemann, J. T. Design of an effective algorithm for fast response to the re-scheduling of railway traffic during disturbances. *Transportation Research Part C: Emerging Technologies*, 20(1):62–78, 2012.

Krueger, H. Parametric modeling in rail capacity planning. In *Simulation Conference Proceedings, 1999 Winter*, volume 2, pages 1194–1200. IEEE, 1999.

Lee, Y., Nelder, J. A., and Pawitan, Y. *Generalized linear models with random effects: unified analysis via H-likelihood*. CRC Press, 2006.

McCullagh, P. and Nelder, J. A. *Generalized linear models*, volume 37. CRC press, 1989.

McCulloch, C. E. and Neuhaus, J. M. *Generalized linear mixed models*. Wiley Online Library, 2001.

Mees, A. I. Railway scheduling by network optimization. *Mathematical and Computer Modelling*, 15(1):33–42, 1991.

Noland, R. B. and Polak, J. W. Travel time variability: a review of theoretical and empirical issues. *Transport reviews*, 22(1):39–54, 2002.

Ochiai, Y. and Tomii, N. Punctuality analysis using a microscopic simulation in which drivers' behaviour is considered. *Journal of Rail Transport Planning & Management*, 5(3):128–145, 2015.

Olsson, N. O. and Haugland, H. Influencing factors on train punctuality – results from some Norwegian studies. *Transport Policy*, 11:387–397, 2004.

Olsson, N. O., Røstad, C. C., and Veiseth, M. Nordlandsbanen – fortsatte analyser. Technical Report S01002831, SINTEF Industrial Management, Trondheim, Norway, 2002.

Pawitan, Y. *In all likelihood: statistical modelling and inference using likelihood*. Oxford University Press, 2001.

Polus, A. A study of travel time and reliability on arterial routes. *Transportation*, 8:141–151, 1979.

Rakh, H., El-Shawarby, I., and Arafeh, M. Trip travel-time reliability: issues and proposed solutions. *Journal of Intelligent Transportation Systems*, 14(4):232–250, 2010.

Rudnicki, A. Measures of regularity and punctuality in public transport operation. In Papageorgiou, M. and Pouliezos, A., editors, *Transportation Systems 1997*, pages 661–666. Elsevier, 1997. ISBN 0080429319.

13

Shafiullah, G., Ali, S. A., Thompson, A., and Wolfs, P. J. Predicting vertical acceleration of railway wagons using regression algorithms. *IEEE Transactions on Intelligent Transportation Systems*, 11(2):290–299, 2010.

Stefánsson, G. Analysis of groundfish survey abundance data: combining the GLM and Delta approaches. *ICES Journal of Marine Science: Journal du Conseil*, 53(3):577–588, 1996.

Venables, W. N. and Ripley, B. D. *Modern Applied Statistics with S*. Springer, New York, fourth edition, 2002. URL `http://www.stats.ox.ac.uk/pub/MASS4`. ISBN 0-387-95457-0.

Verbeke, G. and Molenberghs, G. *Linear mixed models for longitudinal data*. Springer Science & Business Media, 2009.

Westgate, B. S., Woodardand, D. B., Matteson, D. S., Henderson, S. G., et al. Travel time estimation for ambulances using Bayesian data augmentation. *The Annals of Applied Statistics*, 7(2):1139–1161, 2013.

Yuan, J., Goverde, R., and Hansen, I. Evaluating stochastic train process time distribution models on the basis of empirical detection data. *Timetable Planning and Information Quality*, page 95, 2010.

DALARNA
UNIVERSITY