Estimating the noise reduction effect of low- noise asphalts as a function of their lifetime and environmental factors

GEO 511 Master's Thesis

Author
Jan Guddal
14-701-676

Supervised by
Prof. Dr. Robert Weibel
Bühlmann, Erik (eb@gundp.ch)

Faculty representative
Prof. Dr. Robert Weibel

15.01.2020
Department of Geography, University of Zurich

Department of Geography

Estimating the noise reduction effect of low-noise asphalts as a function of their lifetime and environmental factors

GEO 511 Master's Thesis

Author
Jan Guddal
14-701-676

Supervisor and faculty member
Prof. Dr. Robert Weibel
University of Zurich

External Supervisor
Erik Bühlmann
Grolimund + Partner

Date of Submission 15.01.2020
Department of Geography, University of Zurich

Contact

Prof. Dr. Robert Weibel
University of Zurich
Department of Geography
Winterthurerstrasse 190
CH - 8057 Zürich

Erik Bühlmann
Grolimund + Partner
Road surface acoustics, research & development
Thunstrasse 101a
CH - 3006 Bern

Jan Guddal
guddal.jan@gmail.com
Wartstrasse 15
CH – 8400 Winterthur

Abstract

In Switzerland, approximately one eighth of the population is exposed to harmful road noise every day. Low noise asphalts (LNA) are among the most effective methods of reducing road noise, but their noise reduction effect decreases over time. Little is known about the exact causes of the decrease in the noise-reducing effect of LNAs. Therefore, the objective of this thesis was to find out which factors lead to the loss of the noise reducing effect of the LNAs and how a generalizable model can be generated to predict the performance of aging LNAs. 55 independent variables were constructed. Finally, stepwise forward selection was used to select 10 variables (one time-based, one traffic load, four climatic and four physical LNA parameters) for the final model. Different models were evaluated regarding performance and interpretability. Finally, the Linear Model Tree was perceived and chosen as an ideal compromise between performance and interpretability. The interpretation of the model suggests that age is the most important and mechanical stress from traffic the second most important variable to explain the decrease in noise reduction of an LNA. The results indicate that exposure to heavy vehicles is more important than exposure to passenger cars. Furthermore, results indicate that frost has a detrimental effect on LNAs. Findings also show that the consideration of spatial autocorrelation is important for the adequate interpretation of results.

The thesis demonstrates that it is possible to create a robust, detailed model for predicting the noise reduction effect of an LNA in Switzerland. To achieve this, variables should be chosen that describe age, traffic load, climatic environmental factors and the physical parameters of the LNAs. A non-linear function is recommended to describe the relationship between the variables. Spatial autocorrelation does not need to be considered by the model to be robust and detailed, but is recommended and indispensable for interpretation. During data pre-processing, it seems advisable to clean the data from outliers and ignore tracks younger than approximately 1.5 years.

Acknowledgements

My sincere gratitude goes to the supervisors of this thesis, Prof. Dr. Robert Weibel and Erik Bühlmann. Prof. Dr. Robert Weibel's flexibility, experience, helpful advice and openness towards topics somewhat atypical of geography enabled me to write this thesis.

Thanks to Erik Bühlmann's expertise in the field of road surface noise, he was always able to rapidly explain theoretical ambiguities to me whenever necessary.

In addition, I would like to express my gratitude to Prof. Dr. Reinhard Furrer, who took the time to review my statistical approach and to provide useful advice.

Table of Content

List of Abbreviations

CG: Conventional Group
CPX: Close Proximity Method
DT: Decision Tree
FOEN: Swiss Federal Office for the Environment
GB: Giga Byte
GBT: Gradient Boosted Trees
GHz: Giga Hertz
GIS: Geographic Information System
GLM: Generalized Linear Model
GPS: Global Positioning System
HV: Heavy Vehicle
H1:   Heavy Vehicle Test-Tire
LMT: Linear Model Tree
LNA: Low Noise Asphalt
MAE: Mean Absolute Error
MSE: Mean Squared Error
NAO: Noise Abatement Ordinance
OLS: Ordinary Least Square
PC: Passenger Car
PI: Prediction Interval
P1:   Passenger Car Test-Tire
RF: Random Forest
RMSE   Root Mean Squared Error
SDA: Semi-Dense Asphalt
SEM: Spatial Error Model
SMA: Stone Mastic Asphalt
SPB: Statistical Pass By
TRI: Terrain Ruggedness Index
VIF: Variation Inflation Factor
WHO: World Health Organization
WRAcc: Weighted Relative Accuracy

1 Introduction

1.1 Motivation

In their literature survey on the health effects of noise Kurppa and Ising (2004) concluded that numerous empirical results show long-term health risks associated with noise. If, for example, the daytime immission level exceeds 65 dB(A), there is a trend towards an increased cardiovascular risk. Furthermore, noise from airplanes or trucks can also be classified as a hazard signal during sleep and induce the release of stress hormones. According to the noise stress hypothesis, chronic stress hormone dysregulations as well as the increase of established endogenous risk factors of ischemic heart disease under long-term environmental noise exposure are observable. The World Health Organization (WHO) (2011) estimated that the effects of traffic-related noise accounts for over 1 million healthy years of life lost annually to ill health, disability or early death in the western countries in the WHO European Region. Switzerland was early aware of the noise problem and in 1986 drafted the Noise Abatement Ordinance (NAO). However, the recent report of the Swiss federal Office for the Environment (FOEN) (BAFU, 2018) states that despite substantial investments in noise abatement measures, in the year 2015 more than one Million of Swiss inhabitants – which is approximately an eight of the entire Swiss population – were still affected by harmful road traffic noise during day and night. This means that especially with regard to road traffic noise, the communes and cantons fail to achieve full compliance with the NAOs regulations. This was particularly controversial as Article 17 (4) (b) of the NAO provides that, as from 1 April 2018, the federal government will no longer provide financial assistance for road rehabilitation. This decision was taken on the assumption that the communes and cantons would have sufficient time over a period of just over 30 years (1986-2018) to meet the requirements of the NAO. As a consequence, the NAO was adapted for the second time (first time in 2002) so that the government is obligated to provide financial assistance for road rehabilitation until the 31 December 2022 (Article 21 (1)). The fact that in 2018 the government has extended the deadline for financial support for road rehabilitation, which should lead to compliance with NAO regulations, by only 4 compared to 16 years in 2002, and that the Swiss Noise League (Lärmliga Schweiz, 2018) opened a pool of lawsuits against communes and cantons in 2018 shows that the pressure on communes and cantons is increasing strongly.

The possibilities for communes and cantons to reduce road noise are limited. The Swiss National plan of measures to reduce noise emissions (2015) recommends that communes and cantons reduce speed and install low noise asphalts (LNA) in order to reduce the number of people affected by noise. Since according to the sonBase (2018) report, the potential for reducing the number of people affected by noise by a general speed reduction of 20 km/h on all roads is 30% and the potential for installation of LNAs on critical road sections is estimated

at more than 50%, installation of LNAs seems to be the best method to reduce road traffic noise. The reason for the strong noise reduction potential of LNAs is that LNAs combat noise at its source. the tire-road noise. Tire-road noise is the most important source of road traffic noise in the medium to high speed range (Licitra et al., 2018). However, LNAs have a major disadvantage: the durability of their noise reducing effect durability.

As the next section will show, the current state of knowledge regarding the exact causes and processes linked to the measured loss of effectivity is incomplete.

1.2 Problem Statement

There is already research that had the objective of creating an aging model for LNAs. Bendtsen et al. (2010) compared the aging process of five different LNAs types from a sample of approximately 100 pavements in Denmark and California by putting the pavements' age and the amount of traffic on them in relation to the noise increase. The end result was a universal formula that, however only, takes into account age and the amount of traffic on the pavements in order to predict the acoustic performance after a specific number of years the pavement was built. This universal formula is deficient, as the work by Licitra et al. (2018) shows. Licitra et al. (2018) had access to climate and traffic data and compared linear- with logarithmic regression models. The conclusion was that logarithmic models perform better than linear models, and that treating different pavement types and different climate regions as independent variables leads to a more detailed model. This shows that a universal formula for all pavements, such as the one created by Bendtsen et al. (2010), is not representative. But Licitra et al. (2018) used only seven different tracks, their work cannot be considered as representative as well.

The studies by Bendtsen et al. and Licitra et al. are symptomatic of the problem of research on aging processes in LNAs. Their work has the characteristics of a case study, hence has not enough data, be it the number of measuring sections or variables concerning external (e.g. weather, traffic) and internal (e.g. cavity content, grain size mixture) factors that influence the aging of the pavement. This leads to unrepresentative results.

Research that has come closer to a representative modelling of aging are the studies of Bühlmann et al. (2015) & Hammer et al. (2015). The study by Hammer et al. used a total of 371, tracks with three different pavement types. Additionally a conventional pavement type that functioned as a control group (a pavement which is not of the low-noise category) with 328 samples was added. The total of 699 pavements are spread across the entire country of Switzerland. For each pavement type a fitted logarithmic function, where the acoustic performance and years since installation in relation with pavement specific constants (void and subgroups) were respected. For each logarithmic function a prediction interval (PI) of 83.3% was created. The PIs differ strongly; the three values are -3.5±2.6 dB(A), -4.6±2.4 dB(A)

and -0.2±2.4 dB(A). Since a difference of 3 dB corresponds to a doubling of the sound energy, the PIs are not precise enough to be suitable for an aging model. This shows that an individual logarithmic function for each pavement type is not precise enough as well.

Bühlmann et al. (2015) used in their study the same data as Hammer et al. (2015). In addition, environmental data were used, which enabled the authors to produce 14 different environmental variables which were hypothesized to have an influence on the aging process of LNAs. By using a principal component analysis, the study concludes that frost has the strongest negative impact on acoustic durability of an LNA, followed by traffic load. Nevertheless, according to the authors, the results would probably change if models with more data were created.

In sum the research accomplished so far does not deliver a sufficiently detailed aging model for LNAs. This observation is supported by Sirin's (2016) study, which provides a detailed overview of aging models. Sirin concludes that although researchers have presented mathematical models that fit their respective data, there is a clear need to develop a more general and theoretically sound noise reduction prediction model.

It can be assumed that the lack of significance in previous research is due to insufficient data and too few influencing factors respected per model. This is why a study with a large number of data points, more attributes and more advanced statistical methods is needed.

1.3 Research questions

As the previous section shows, much is not yet understood about the acoustic aging of LNAs over time. Section 2.2 will show that some hypotheses exist, but the literature on the individual hypotheses is rather sparse, as the mathematical models created often only fit the data of the respective study. Since this thesis has a large database available for this field of research, the aim is to close the research gap. This means that this work should provide more generalizable results and use novel methods, i.e., go beyond the traditional linear, monotonic models of previous studies. Consequently, it would be wrong for this thesis to formulate concrete hypotheses, since this would limit the hypothesis space due to the limited available knowledge and would thus make it more difficult to find generalizable results. Hence, the research questions of this thesis have a general approach to the area under study and indirectly ask for specific relationships between the variables. The first research question is as follows:

"Which factors lead to the loss of the noise reducing effect of the LNAs?"

The difficulty in answering this research question lies in creating a model that performs well over the entire data set and should also be simple to interpret.

In order to create a model that performs well, different strategies must be examined. Therefore, the second research research question is as follows:

"How can a generalizable model to predict the performance of LNAs be created?"

This means that a model must be fitted to the data, allowing the model to reliably estimate unprecedented data in the whole area of Switzerland where LNAs are commonly installed.

1.4 Thesis Structure

Following this introduction, the second chapter provides an overview of the work related to this thesis. In addition, Chapter 2 provides a brief theoretical introduction to the research field of low-noise road surfaces. The third chapter describes the data and their pre-processing. The fourth chapter systematically describes the structure of the model design. In the fifth chapter the results and their interpretation of the model created in Chapter 4 follow. Thereafter, a concrete statement is given on the extent to which the research questions of this work have been answered. The sixth chapter concludes the thesis, summarizing the main points and providing an outlook on future work.

2 Background

2.1 Introduction to road surface acoustics

The purpose of this section is to provide readers with limited or no knowledge of road surface acoustics with a concise overview of the topic. In addition, the terms "sound, noise" (Section 2.1.1) and "LNA" (Section 2.1.3) are defined for the purpose of this thesis.

2.1.1 Definition of sound and noise

Sound is a term used in various study fields, which can lead to confusion. The Oxford Dictionary defines sound as "*vibrations that travel through the air or another medium and can be heard when they reach a person's or animal's ear*". This definition perfectly reflects the confusion of the definition, since it is a mix of the following two definitions:

Natural scientific definition: "sound is a vibration that typically propagates as an audible wave of pressure, through a transmission medium such as a gas, liquid or solid." (Jung et al. 2018)

Social scientific definition: "sound is the reception of such waves and their perception by the brain." (Ronan 1967).

The difference is that sound is either defined in such a way that it is audible to humans and receives a subjective evaluation from the individuals who received it, or that sound means the propagating vibration, which can also have frequencies that are not audible nor rated by humans.

The term "noise" is originating from the human scientific definition of sound. Noise is sound that was perceived as irritating by an individual. In this thesis sound is defined according to the natural scientific definition. But since the term noise is commonly used in the tire-noise literature, it is also used in this thesis. However, in this thesis, the term noise is subordinate to the natural scientific definition of sound. Since the classification of sound as noise is subjective and the objective of this thesis is not to evaluate to what extent the vibrations caused by the interaction between tire and road surface are irritating for individuals, it is simply assumed that if sound is produced by the interaction between tire and road surface, the sound is automatically defined as noise, even if it might not be annoying for a particular individual.

2.1.2 How tire-road noise is produced

A broad and detailed collection of knowledge about tire-road noise can be found in the book *Tire/Road Reference Book* by Sandberg and Ejsmont (2001). Bernhard and Wayson (2005) wrote a more concise literature overview on which this section is based.

There are four main mechanisms that produce tire-road noise.

The *tread impact* (Figure 2.1) occurs when the running tread hits the road surface, the vibration caused by this hit is comparable to a small rubber hammer hitting the road surface. The macrotexture of a road surface influences the radial vibration as well. The more uneven the texture is, the stronger radial vibrations are produced.
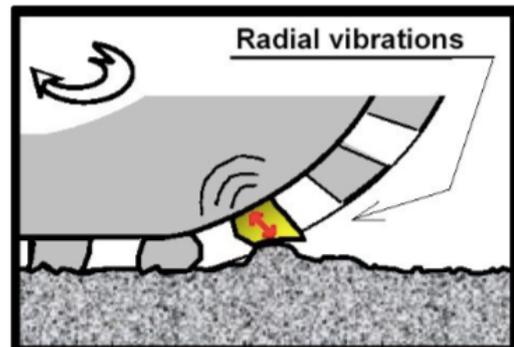
Figure 2.1: Vibration caused by tread/road surface impact

The *air pumping* (Figure 2.2) effect is caused in the contact patch between road surface and tire. In the contact patch, air is sucked in and deformed by the grooves in the tread pattern on the front of the tire. The air carried in these channels is compressed and pumped out at the rear of the tire, creating an aerodynamically generated sound. This phenomenon is similar to the sound produced by clapping one's hands.
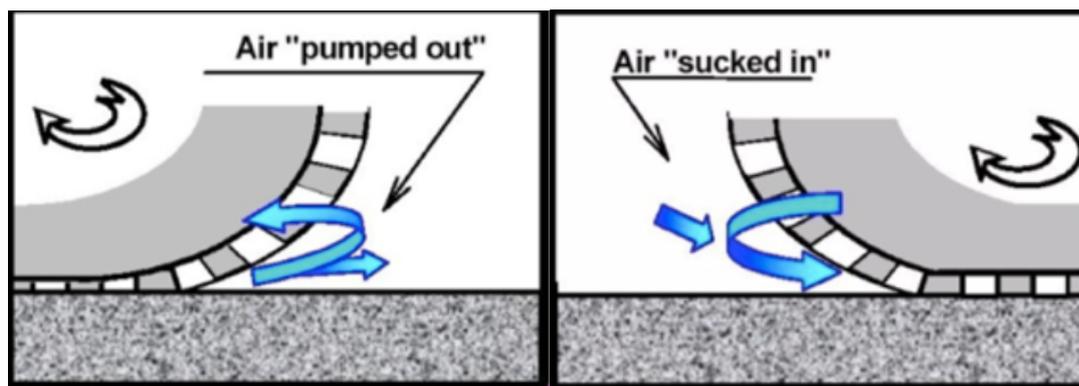
Figure 2.2: Air pumping at the entrance and exit of the contact patch

The *Stick-slip* (Figure 2.3) noise is caused by acceleration or braking. Acceleration and breaking cause deformation of the tire carcass in the contact patch and generate strong horizontal forces at the interface between the tread block and the road surface. If these horizontal forces exceed the friction limits, the tread block slips briefly and then sticks to the road again. This action of slipping and sticking can take place very quickly and generates both noise and vibration. This phenomenon is observed in the gym when sports shoes squeak on a playing field.

Figure 2.3: Stick-slip motion of the tread block on the road surface
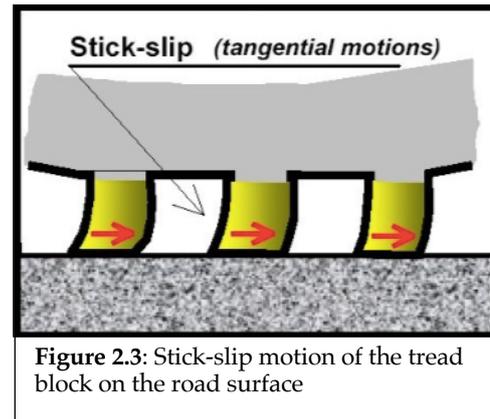
Adhesion (Figure 2.4) noise occurs due to the contact between the tread block and the road surface which results in adhesion between the tread block and the road surface. The phenomenon can be compared with the behavior of the suction cup. When the tread block leaves the contact surface, the holding force holds the tread block back for a short moment. The subsequent loosening of the tread block from the road surface, hence the overcoming of the adhesion, causes both acoustic energy and vibration of the tire carcass.

Figure 2.4: Adhesion between the tread block and road surface at the exit of the contact patch
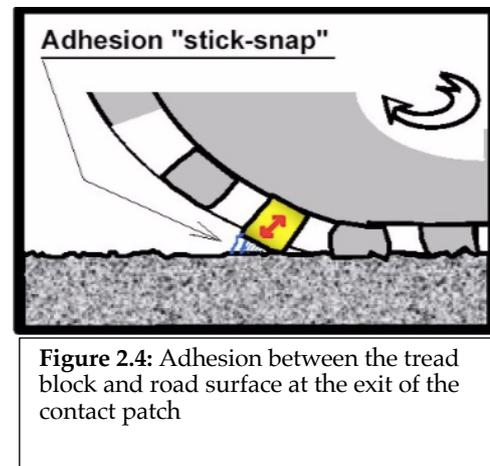
In addition to the mechanisms that generate noise, there are other mechanisms that amplify it, such as the horn effect. The geometry of the tire over the road surface is a natural horn, as shown in Figure 2.5, although the shape is not a classic horn. However, sound produced by a source mechanism near the horn neck is amplified by the horn. Further amplifying effects are described in detail by Sandberg and Ejsmont (2001).

Figure 2.5: The horn effect created by the tire and road surface
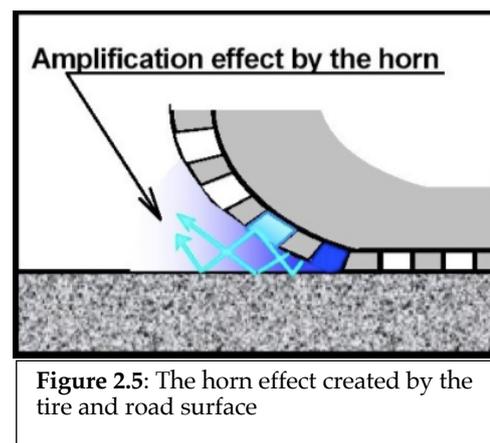
2.1.3 Low-noise asphalts (LNAs)

Sandberg (1999) already addressed the problem of the use of the term LNA in his literature review. Sandberg sees the problem in the fact that in principle a road surface does not cause noise in the resting state; noise develops only with the interaction of a vehicle with rolling tires. Therefore, it is the rolling tire that causes the noise, so theoretically any road surface can be called an LNA.

Sandberg himself suggests the definition: "A 'low noise road surface' is a road surface which, when interacting with a rolling tire, influences vehicle noise in such a way as to cause at least 3 dB(A) (half power) lower vehicle noise than that obtained on conventional and 'most common' road surfaces." This definition is not applicable for this thesis, since this thesis intends to recognize why road surfaces have difficulties achieving the goal to produce a noise reduction of 3 dB(A). This means that a road surface in this thesis is defined as an LNA if the road surface's purpose is to fulfil the definition given by the standard, which reads as follows: "over its entire service life (12-15 years) at least 1 dB less noise is generated than with a conventional road surface according to the road noise model StL86+ (details in Section 2.1.4)". In addition, the noise reduction at the start of use must be at least 3 dB, which acoustically corresponds to halving traffic (Schweizerischer Verband der Strassen- und Verkehrsfachleute VSS Zürich 2013, Figure 2.6).
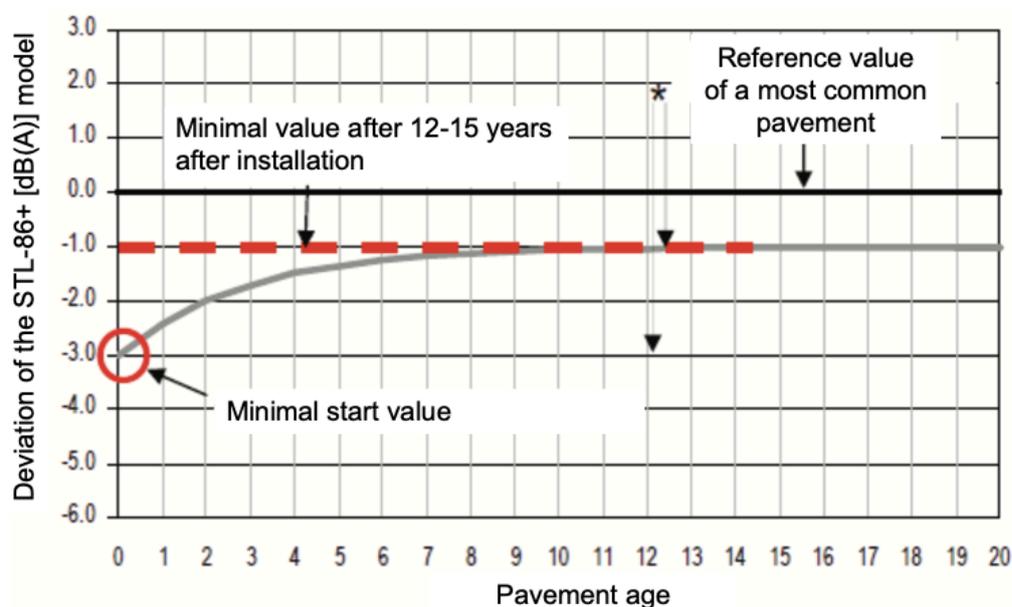
Figure 2.6: Schematic representation of the Swiss LNA definition

In order to meet the requirements of the NAO (or Sandberg's definition), engineers have developed various designs. Again, a detailed description of these types can be found in Sandberg's tire-noise Reference book. Since there is hardly any data (Section 2.2.2) on the

physical parameters of the road surfaces available for this thesis, an exact execution of these types is dispensed with. Another problem with the physical parameters is that the results were derived from case studies; in addition different countries have different standards and definitions for road surface designs, which renders a literature review even more difficult. Beyond that, an LNA must meet further criteria, such as cost-effectiveness, safety or mechanical and structural resistance, which further complicates the design. For example, the correlation between skid resistance (which leads to more grip and therefore more safety) and tire-road noise is negative (Ongel et al. 2007). Therefore, in the following only the rough basic concept of an LNA design is explained and how it contributes to noise reduction.

The most important feature of an LNA are the voids accessible at the surface (Hammer and Bühlmann 2017) and the LNA's texture (Wayson 1998). The voids can reduce mechanisms such as the air-pumping effect, as the air is now sucked not only into the grooves of the tread pattern but also into the voids of the road surface, resulting in less air pumping and consequently less noise. Furthermore, voids have the property of generally absorbing sound, since the sound waves are trapped like in a house and are not directly reflected back (Bernhard and Wayson 2005). Additionally, LNAs are highly effective in reducing tire-road noise, as sound waves are absorbed every time they hit the surface and thereby efficiently reduce the horn-effect (Sandberg 1999).

It is difficult to determine the exact effect of the texture, as the effect of the texture comes to bear mainly in combination with the tread pattern. The facts indicate that the optimization of car tires is different from that of truck tires (Sandberg and Descornet 1980). More recent studies, however, show that fine (Descornet 2000, Bennert et al. 2005) and smooth textures (Hansen et al. 2004, Bennert et al. 2005) generally contribute more to noise reduction than coarse und uneven ones.

2.1.4 StL86+ model and CPX measurements

As described in Section 2.1.3, in Switzerland the acoustical performance of a road surface is typically characterized with reference to the Swiss road traffic noise emission model StL86+. The Stl86+ model, as any acoustic model, can be divided into source and propagation properties.

For source characteristics, the radiated sound level depends on the amount of traffic, the heavy vehicles (HV) share and the speed.

The formula used is:

$$L = A + 10 \times \log\left(\left(1 + \frac{v}{50}^3\right) \times \left(1 + B \times Eta \times \left(1 - \frac{v}{150}\right)\right)\right) + 10 \times \log(M)$$

L: *Energy equivalent continuous sound level in $dB(A)$*

A, B: *empirical constants, $A = 43$, $B = 20$*

v: *velocity*

Eta: *HV share (as a fraction of total traffic)*

M: *Traffic volume (number of vehicles per hour)*

The formulae for calculating the properties of sound propagation, such as geometric damping (by propagation), ground damping, air damping and obstacle damping are described in detail in Environment Series No. 60 (1987).

In summary, the StL86+ model is intended to provide noise predictions by calculating an energy equivalent continuous noise level value in dB(A) for each receiving point.

With regard to close proximity (CPX, more details in Section 3.1) measurements, the acoustic quality of the road surfaces will be indicated as a deviation from the reference road surface of the Swiss StL-86+ reference emission model. The conversion of the tire/road noise levels measured at close proximity to the source (the tire) to the effect that a road surface has on total traffic noise emissions are undertaken by correlating the values obtained with the CPX method with the statistical pass-by (SPB) method. The regression models created for the conversion were collected separately for the vehicle categories passenger cars (PC) (N1) and heavy vehicles (HV) (N2), as well as for the CPX reference speeds 50 km/h and 80 km/h. Each regression model is based on at least 25 SPB measurements, which are also subject to certain standards. These energy equivalent continuous sound level values based on SPB measurements were performed on road sections on road surfaces representative of common road surfaces in Switzerland at normal speed regimes. The CPX index values CPXp and CPXh must be converted into StL-86+ values using the regression equations.

CPX-reference speed 50 km/h:

$$PC(N1): \quad KB_{\frac{50km}{h},PC} = 1.2468 \times CPXp - 112.30$$

$$HV(N2): \quad KB_{\frac{50km}{h},HV} = 1.3617 \times CPXh - 116.16$$

KB: *Correction of pavement influence compared with the reference emission model StL* $-86+$ *as a function of vehicle category and CPX reference speed*

$CPXp$: *CPX index value for passenger cars*

$CPXh$: *CPX index value for heavy vehicles*

Detailed information on all corrections and standards concerning Swiss CPX and SPB can be found under Schguanin and Ziegler (2006), Appendix 1c.

2.1.5 Difference between P1 and H1 tire

Bühlmann's (2019a) study, which tested the acoustic conformity of the test tires, showed that for the heavy vehicle test-tire (H1) the variation of the spectral noise levels in some of the octave bands was considerably greater compared to the passenger car test tire (P1). The standard deviations of the noise levels for the P1 tire are rather small and homogeneous over the noise spectrum and close to the expected measurement uncertainty. The reason why the P1 tire has a smaller standard deviation of the noise levels is that the P1 tire is designed for use in various tests in the automotive industry. In order to ensure the comparability of these tests, the tire will be made available to car manufacturers with virtually unchanged rubber compounds over the long term. This also leads to a satisfactory acoustic agreement of the tires. Unlike the P1 tire, the H1 tire is a market tire. According to tire manufacturers, it is a common practice for tire manufacturers to adjust rubber compounds over the life of a product line, making it difficult to repeat measurements with the H1 tire. Therefore, the acoustic values for measurements with the passenger car tire (variable name = STNL1) are used as target variable in this thesis.

2.2 Related Work

This section provides an overview of previous research on durability of LNAs. Variables investigated so far, and their proven or suspected effects are discussed individually.

2.2.1 Reduction of noise reducing effect at the micro scale

To understand why the factors mentioned in Sections 2.2.2-2.2.5 contribute to the deterioration of LNAs over time, it is necessary to explain what exactly happens with asphalt at the micro level, i.e. the pores and texture. The two main factors that lead to a worsening of the noise reduction effect of an LNA are clogging and rutting.

The blockage of pores can occur due to agricultural pollution (Gradziejczyk 2016), mud, dust, spilled oil or the like (Descornet 2000). Clogging the pores leads to LNAs' voids to lessen, hence the general ability of voids to absorb the sound waves lessens as well.

Rutting is the permanent deformation of a road surface, such as aggregate losses. Rutting is normally caused by heavy traffic (Elvik 2003) or, more generally, by strong mechanical stress. Rutting leads to an uneven and coarse road surface, which is unfavorable for noise reduction.

2.2.2 Physical parameters

The main focus of LNA research is on the observation of physical parameters. Although research has not yet found the recipe for the perfect LNA, some findings are globally accepted. First, the larger the maximum aggregate size, the greater the tire-road noise (EAPA 2007).

Secondly, a smooth surface leads to a reduction of the tire-road noise (Hansen et al. 2004, Bennert et al. 2005). However, completely smooth surfaces have the disadvantage that they have a lower mechanical resistance and rutting occurs earlier.

Third, larger air-void content leads to less tire-road noise (Sandberg and Ejsmont 2002). But also, the air-void content cannot be arbitrarily large. At a certain size (depending on other physical parameters such as aggregate size, bitumen or texture), the road surface becomes fragile to mechanical stress.

Fourth, open-graded asphalt mixtures have a stronger noise reducing effect than dense-graded asphalt mixtures (Bennert et al. 2005).

Fifth, bitumen. When a new LNA is installed, a dark bitumen film is pulled over the aggregates. This bitumen film is a result of the production process of any asphalt mixture and serves as protection as it is worn away by the traffic rolling over it. About the exact noise-reducing effect of bitumen there is almost nothing known, because most measurements take place when the bitumen film has worn off (Bendtsen et al. 2010). The bitumen is rather used to keep the road surface stable for as long as possible (until it has just worn off) before the mechanical stress and other environmental factors can add to the LNA. The aging behavior is therefore delayed.

In Switzerland, the product "semi-dense asphalt" (SDA) has established itself as the best option for LNAs in urban areas. The aim of SDAs is to achieve the highest possible noise reducing effect, with the lowest possible void content and the largest possible aggregate size (large aggregate size is not primarily noise reducing but more resistant to mechanical stress).

Physical parameters in Switzerland. According to Bühlmann (2019b) current experience shows that SDAs with maximum aggregate sizes of 4 mm (SDA4) or 8 mm (SDA8) at lower speed ranges are more effective at reducing noise levels than conventional road surfaces such as SMA11. In the first 12 months after construction, SDAs can reduce noise emissions by about -6 to -10 dB(A). In general, semi-density asphalts with a maximum aggregate size of 4 mm are approximately 2 dB quieter than semi-dense asphalts with a larger aggregate size of 8 mm. The SDA 4 has slightly higher construction costs and a shorter service life. With regard to the long-term effectiveness of SDA road surfaces, long-term CPX measurements show that an SDA4 maintains an efficiency of -3 dB at the end of its service life, while SDA8 has a value of -1 dB(A).

Although Bühlmann (2019b) observed unambiguous differences between SDA4 and SDA8, the Swiss norms used to designate the LNAs could be even more precise. Bühlmann and Hammer (2017) assume that the maximal filler proportion (sieve 0.063 mm) and maximal sand proportion (sieve 2mm) content of SDAs are the decisive factors for long-term acoustic performance. In the Figure 2.7 the black lines serve as thresholds. All gray curves are road surfaces. A road surface is defined as SDA4 as long as its sieve curve is between the black lines. In the figure it can be observed that exactly at the 2 mm sieve, the minimum and maximum allowed proportions are far apart. This is particularly problematic because the 2 mm sieve is a decisive factor in determining the long-term acoustic performance of a road surface. As a result, products with the same norm but from different companies may differ strongly in their final performance.
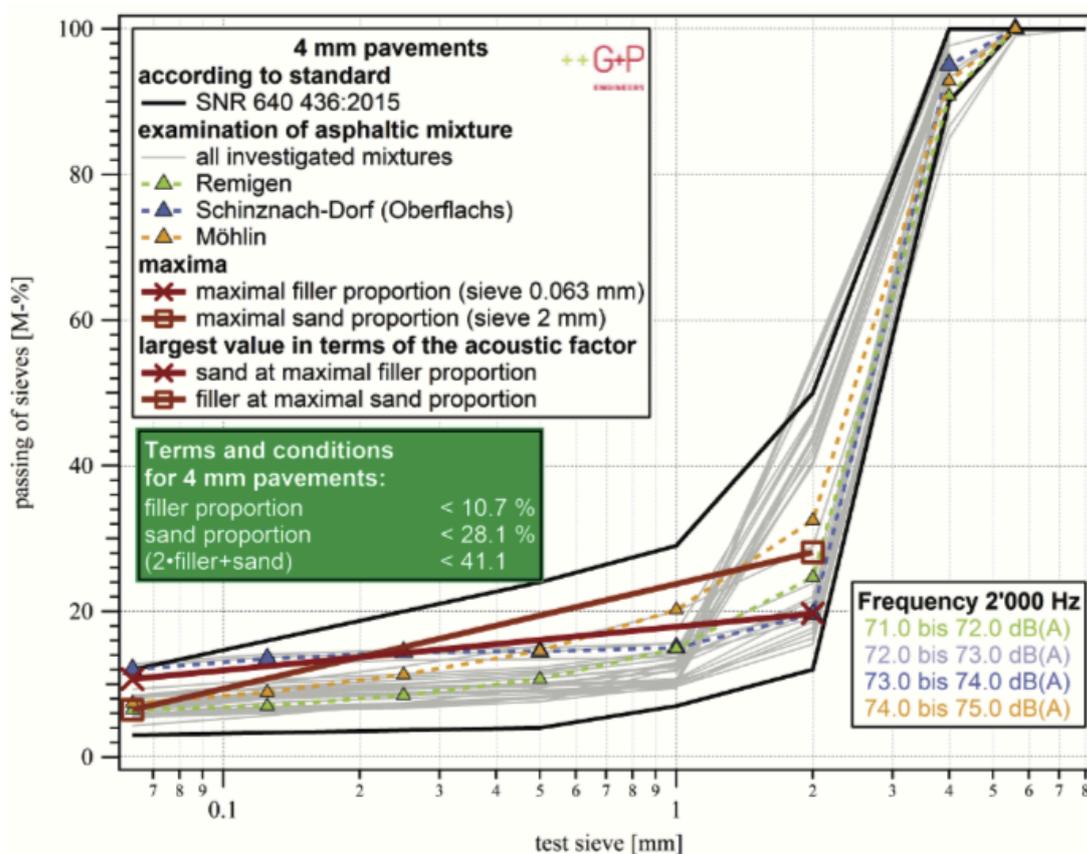
Figure 2.7: Maximum filler and sand content for the design of 4 mm semi-dense asphalts. Grading curves of the underlying sample are grayed out; minimum and maximum.
Thresholds specified by the Swiss standards in black; examples shown in color with corresponding
Surface image of the road surface on the upper side (void content mixture between 12.3 and 14.4 %).

2.2.3 Traffic load

In the first road noise models (in the period of the 1990s) the traffic load was already a strongly focused variable. The Nordic Model (TemaNord 1996), which was developed and used by Scandinavian countries, provided relatively precise results as it included detailed road surface corrections. However, the model only accounted for hot mix asphalts. In addition, the model failed to distinguish normal cars from HVs.

To distinguish normal passenger cars from HVs seems to be indispensable with an aging model for LNAs as the study by van Blokland et al. (2014) shows. HVs have a stronger negative influence on the reduction of the noise reducing effect of an LNA than passenger cars. This could be discovered, because on the motorway the different lanes were considered isolated. The slowest lanes, i.e. where the most trucks are, showed the most wear and tear. Despite this, it can be assumed that it is not solely the exposure to heavy vehicles that is responsible for the decrease in the noise reducing effect of LNAs.

However, the effect of HVs is not apparent in all cases. In Bühlmann's et al. (2015) study, total traffic seems to be more important than the number of HVs. In addition, the climatic factors in this study were more important than the traffic load for the aging of the LNAs.

2.2.4 Climate

The influence of climatic variables on the aging of LNAs is not as undisputed as the work by Bühlmann et al. (2015) suggests. The study by Irali et al. (2015) concludes that temperature has a minor influence on the aging of LNAs. Whereby their study examined only 6 test tracks without specifying their spatial distribution.

Licitra et al. (2018) also examined only 7 test tracks, but these were located in climatically different regions. The study concluded that test tracks with low traffic values, but with many ice days and many freeze-thaw cycles, have high rates of acoustic decay. Bühlmann's et al. (2015) study, which had 371 test tracks available in climatically diverse regions of Switzerland (e.g. Alps vs. lowlands), found that frost and altitude have a significant impact on the noise reduction effect.

The comparison of these 3 studies indicates that the influence of climatic variables becomes apparent when the test tracks are located in climatically different regions. It can therefore be assumed that climatic variables have an influence on the aging of LNAs. However, which climatic variables cause the greatest stress for an LNA has not yet been established. It is also unclear whether traffic load or climatic stress are more harmful for LNAs, since Licitra et al. (2018) and Bühlmann et al. (2015) contradict each other in this respect.

2.2.5 Agricultural land and construction sites

The entry of dirt leads to blockages in the pores of an LNA. Gradziejczyk (2016) finds evidence in his study that the closeness to farmland probably leads to a higher dirt input and that an LNA therefore loses its noise reducing effect faster. Bühlmann et al. (2015) found no negative influence on the noise reducing effect of an LNA with decreasing distance to farmlands. The hypothesis that proximity to farmland has a negative influence on the noise reduction effect of an LNA is therefore unclear and will therefore also be considered in this thesis.

A similar observation could be made on construction sites, as dirt often enters the streets there as well. In addition, heavy vehicles can be found on construction sites, which could generate additional harmful mechanical stress. Therefore, this thesis also examines whether the closeness to construction sites has an influence.

2.3 Background for modelling ecological phenomena

This section is intended to provide an overview of the possible methods that could contribute to achieving the research questions of this thesis.

This thesis is a scientific thesis. Science's four traditional claims are rationality, truth, objectivity and realism (Gauch 2003). The natural sciences can be divided into various categories. One category that occurs in several subdivisions is the empirical category, to which this thesis belongs. An empirical approach uses the methods from its superordinate category, the formale sciences such as mathematics and logic, to transform information about nature into formulas that can be explained as clear statements about the "laws of nature" (Lagemaat 2005). A law of nature is a statement that describes or predicts a range of natural phenomena (Oxford English Dictionary 2005). To describe these "laws of nature" there are a vast number of models.

2.3.1 Ordinary least square regression

One of the simplest methods to describe a natural law is the ordinary least square (OLS) regression. The vast majority of the work cited in Section 2.2 shows that the most popular model in the area of LNA research is ordinary least square regression. This is not surprising, as ordinary least square regression is simple to understand and apply. Another advantage is that even if the relationship between two variables is not linear, the rough direction of the relationship can still be surmised. However, the simplicity of the application also entails risks. With an ordinary least square regression model it should be assumed that the data has no multicollinearity, heteroscedasticity, autocorrelation, or nonlinearity. Especially with regard to autocorrelation, there has been criticism for decades that spatial autocorrelation is often neglected by scientists when investigating spatial data and that ordinary least square regression is therefore applied incorrectly (Poole and O'Farrell 1971). Autocorrelation violates the assumption of random sampling, consequently the standard errors of the estimated coefficient are distorted and the results of the test statistics are no longer reliable.

2.3.2 Spatial regression models

According to Legendre (1993), if spatial autocorrelation is present in data, a researcher has two main options. Either one removes data points from the total dataset until spatial independence is guaranteed (this is not recommended as it often results in the loss of expensive data) or one modifies the statistical method to take spatial autocorrelation into account.

The three main reasons for spatial autocorrelation are:

1. The value of the target variable in a region might impact the value of the target variable in a neighboring or close region.
2. The values of the independent variables in a region might impact the value of the independent variables in a neighboring or close region.
3. The residuals might affect the residuals in a neighboring or close region (spatial heteroscedasticity).

Anselin's (2012) work provides an overview of different models that take spatial autocorrelation into account. Frequently, a spatial-lag model is used to correct spatial autocorrelation that has arisen due to the first two of the above reasons. A spatial error model is usually used to consider spatial autocorrelation based on the third reason.

If models that take spatial autocorrelation into consideration are not available, one may also rely on permutation tests.

Kanevski et al. (2009) consider the research field of geostatistics to be well-established, i.e. the topic of spatial autocorrelation and the development of models that take this into account. Furthermore, the current trend in spatial data, namely that there is an overload of data, means that one has to go beyond geostatistics and use more advanced methods, namely machine learning.

2.3.3 Machine learning

Kanevski et al. (2009) justify the need for machine learning in the study of environmental phenomena, in addition to the overload of spatial data, by three reasons: first, many environmental studies are embedded in a high-dimensional geo-feature space; second, the relationship between variables in most real-case studies are highly nonlinear; and third, the nonstationarity of many real environmental spatial data requires new, efficient and adaptive approaches. Machine learning in the field of environmental studies is not yet as elaborated as geostatistics, as can be seen by browsing the recently published literature.

Mahesh and Surinder (2009) modelled daily evapotranspiration using an M5 model tree; Cracknell and Reading (2014) compared 5 different machine learning algorithms (none is an M5 model tree) to classify lithology of remote sensing data. The algorithms used in these two studies are as different as the preprocessing procedures. There are no standardized procedures, only the evaluation of the models is done by permutation tests. The effect of spatial autocorrelation also seems to have been researched to a certain extent. Gahegan (2000) believes that machine learning algorithms use data very selectively and learn a function that combines the data in the most effective way to identify a particular objective. When autocorrelation affects the predictability of a condition, the machine algorithm learns to give

less weight to the affected attributes. This statement is supported by Brenning (2005), who in his work attempted to identify the future landslides using machine learning algorithms - one of which takes into account spatial autocorrelation. He states that a statistically correct model is not required to achieve good predictive properties. However, Brenning also argues that if statistical inference to model coefficients is a secondary goal for analytical purposes, an adequate representation of the spatial autocorrelation structure is imperative.

2.3.4 Overview

There is no such thing as a cookbook that prescribes which model must be used in an empirical study and how exactly. It is much more the responsibility of the researcher to be aware of the multitude of possible models in order to be able to choose from this multitude a method that helps to achieve the given research objectives. Equally important, however, is that the researcher is aware of the limitations of the chosen model. This is the only way empirical research can deliver meaningful results.

3 Data and preprocessing

This chapter provides an overview of the construction and cleaning of the variables and the general data preparation. A brief description of the meaning and units of measurement of the variables can be found in the Data Description Report in Appendix A1. The data quality report is attached in Appendix A2. The spatial distribution is visualized on maps in Appendix A3.

3.1 The Grolimund + Partner database

The Grolimund + Partner database contains the location as x-, y-coordinates, the measurement date, the road surface type, the road surface ID, the canton in which the road surface is situated and the target variable "STLN1".
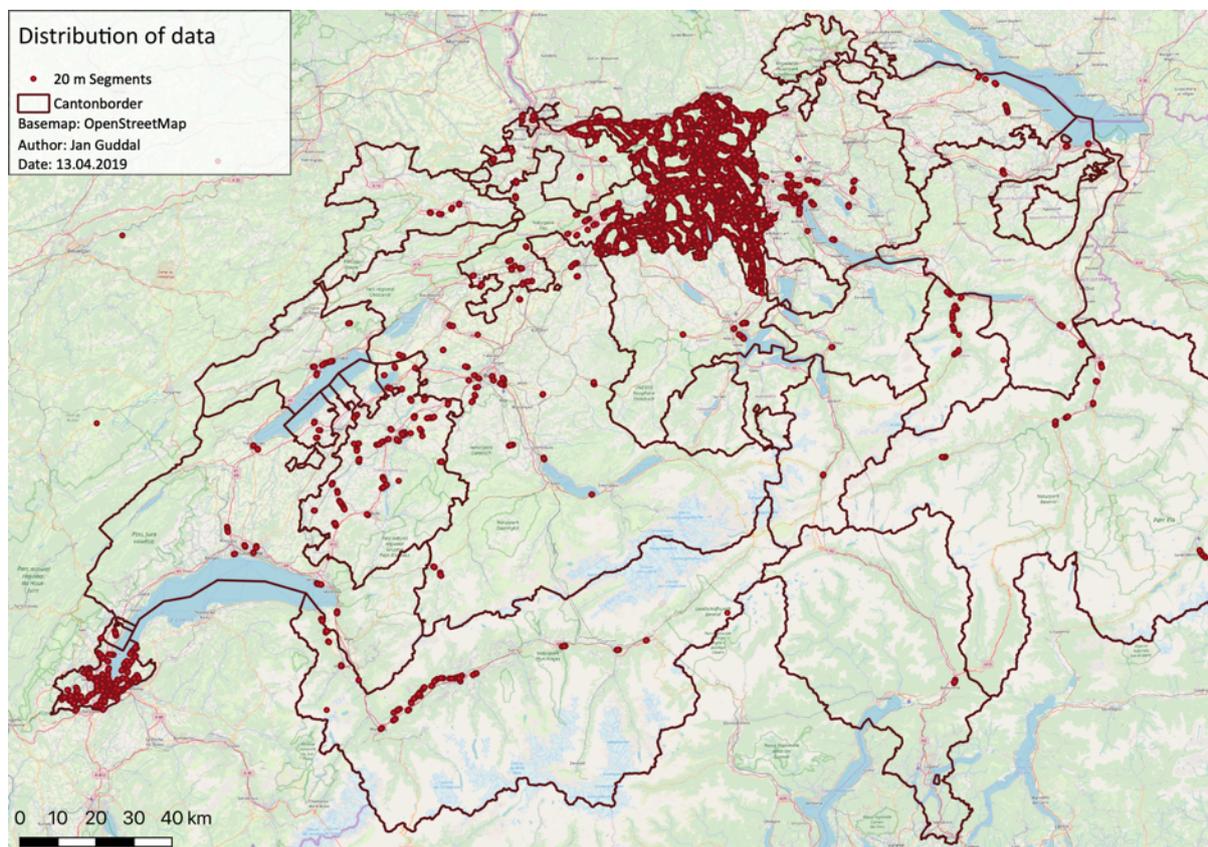
Figure 3.1: Visualization of the spatial distribution of the Grolimund + Partner raw data

3.1.1 Target variable "STNL1"

The "STNL1" variable is measured by the CPX method with a passenger car tire positioned in each wheel track in compliance with the EN ISO 11819-2:2017 standards. The CPX method measures the tire-road noise in two separate sound-insulated chambers inside the measurement trailer in the immediate vicinity of the tires, each with two microphones.

For a normed passenger car tire, sound levels are continuously recorded over the entire measuring section. The microphone signals are recorded at a frequency of 8 Hz and averaged energetically for each measuring segment (which has a normed length of 20 m). The test run is carried out at a constant speed (reference speed of 50 km/h).

The measured CPX values are transformed as described in Section 2.1.4, so that the values finally represent the acoustic deviation from the reference road surface of the StL86+ model (negative value means that the measured road surface segment is quieter than the reference road surface).

From a data analysis point of view, these measured values have a considerable measurement error. Bühlmann (2019a) investigated in his study the repeatability of the CPX measurements of the data used in this thesis. Bühlmann found that using the correction schemes of the current ISO standards, the practical repeatability of the CPX method is 0.45±0.3 dB. The R^2 of the total road traffic noise emissions is 0.87.

The Grolimund + Partner database has the advantage that the data quality is high. Segments that were too slow (at speeds < 45 km/h), measurements taken at too cold temperatures (< 5°C) and loud noise caused by pedestrian crossings, manhole covers, squeaking of tires in tight curves or the like were manually excluded.

But the Grolimund + Partner database has a major disadvantage: The implementation is customer-oriented. Therefore, time attributes function as primary keys, which allows the user to quickly recognize which measurement belongs to which order of the company. A location identifier was only introduced several years after initial deployment and was not consistently reworked. An algorithm had to be created to create a vector ID, namely the Coordinate ID, which contains all measurements ever measured at a specific vector.

3.1.2 Attribute construction

Coordinate ID. The primary key must not be the time but the place. Therefore, an attribute, the Coordinate ID, must be created. The Coordinate ID represents the location of a segment of a road surface. Suppose a segment has the Coordinate ID "1". This means that all measurements that took place exactly at the location "1" also have the value "1" for the attribute Coordinate ID. The first step to create the Coordinate ID is to spatially join each starting coordinate of a vector with each possible starting point of another vector within 15 meters. The search radius is 15 meters, since not every vector is exactly 20 meters long, smaller deviations in the 0.5 meters range occur frequently, in addition the vectors of different measurements are shifted horizontally as well as vertically (thus are never parallel to other vectors of another series of measurements). With the 15 meters it is assured that each starting point of a vector joins all potential addable starting points (Figure 3.2).

The tolerant spatial-join of the first step has the consequence that also starting points of the oncoming lane are joined to the points, which leads to the second step. In order to exclude

points of the opposite lane, the aspect angle in the x-y plane was calculated for each vector. Vectors of the oncoming lane should have a significantly different aspect angle than those on the same lane. However, the inclusion of aspect led to problems. This is because the GPS can show disturbances such as bidirectionality during measurements. This leads to the fact that the aspect angle criterion partly wrongly joins points of the oncoming lane or points actually belonging to the same lane are not joined (This is the case in the example of Figure 3.2 – 3.4, where the vecAngle differences are too small to identify which segment belongs to witch lane). Nevertheless, this results in a gain of information that makes it possible to create an exclusion procedure by linking the various temporal primary keys.

The third step (Figure 3.3 – 3.4), the exclusion procedure takes for each initial coordinate of a measured lane the nearest starting coordinates from all other potential segments (resulting from the first two steps) searched within a search radius of 25 meters. Then the next coordinate of the primary lane is selected, again the start coordinates of the potential lanes are searched, but also the start coordinates of the vector following the start coordinate of the previous joined vector. The algorithm checks whether a segment is correctly joined to a lane by controlling if the value of the SID of the potential point is exactly + 1 of the SID from the point with the 25 meters search radius. This makes it clear which potential lanes lie on the opposite lane, because the algorithm can no longer merge vectors of an opposite lane, since these vectors lie outside the 25-meter distance. The oncoming lanes grow in the opposite direction, so the points of these oncoming lanes cannot be joined at the end.

When the segments are successfully joined to a lane, they are removed from a stack, thus ensuring that no duplicates are created. The end result are the unique coordinates of a vector that contains all the correct start-coordinates of vectors of the same segment measured at a different time, the Coordinate ID.

In another function, the average of all these coordinates is used to create the averaged coordinate and finally calculate the midpoint coordinate of the start and end point of each segment. Those midpoints are used to spatially join all the attribute data.
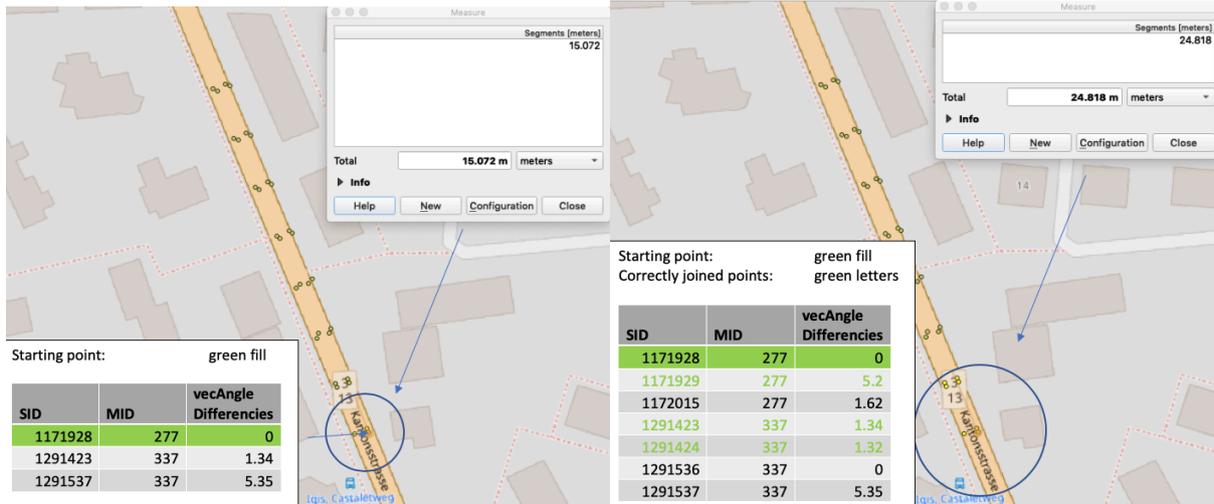
Figure 3.2: Step 1 of Coordinate ID algorithm, joining all potential addable starting points to the main starting point (green fill)
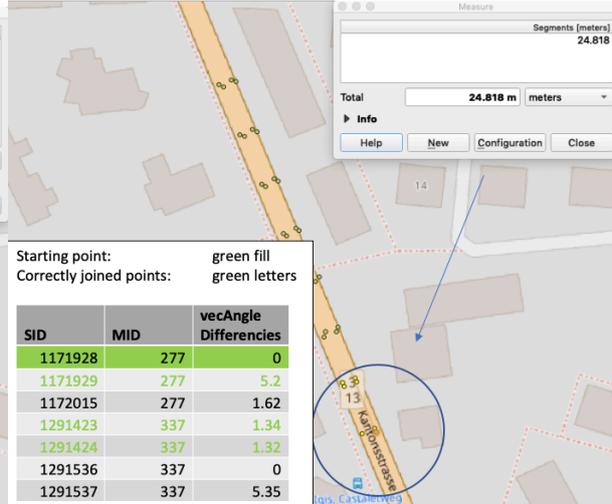
Figure 3.3: Step 3 of Coordinate ID algorithm, joining all potential addable segments inside a 25 meters radius. In this case, the algorithm cannot rely on the vecAngle differences attribute, therefor no change happened in step 2 of the algorithm and all segments inside the search radius are joined for the moment.
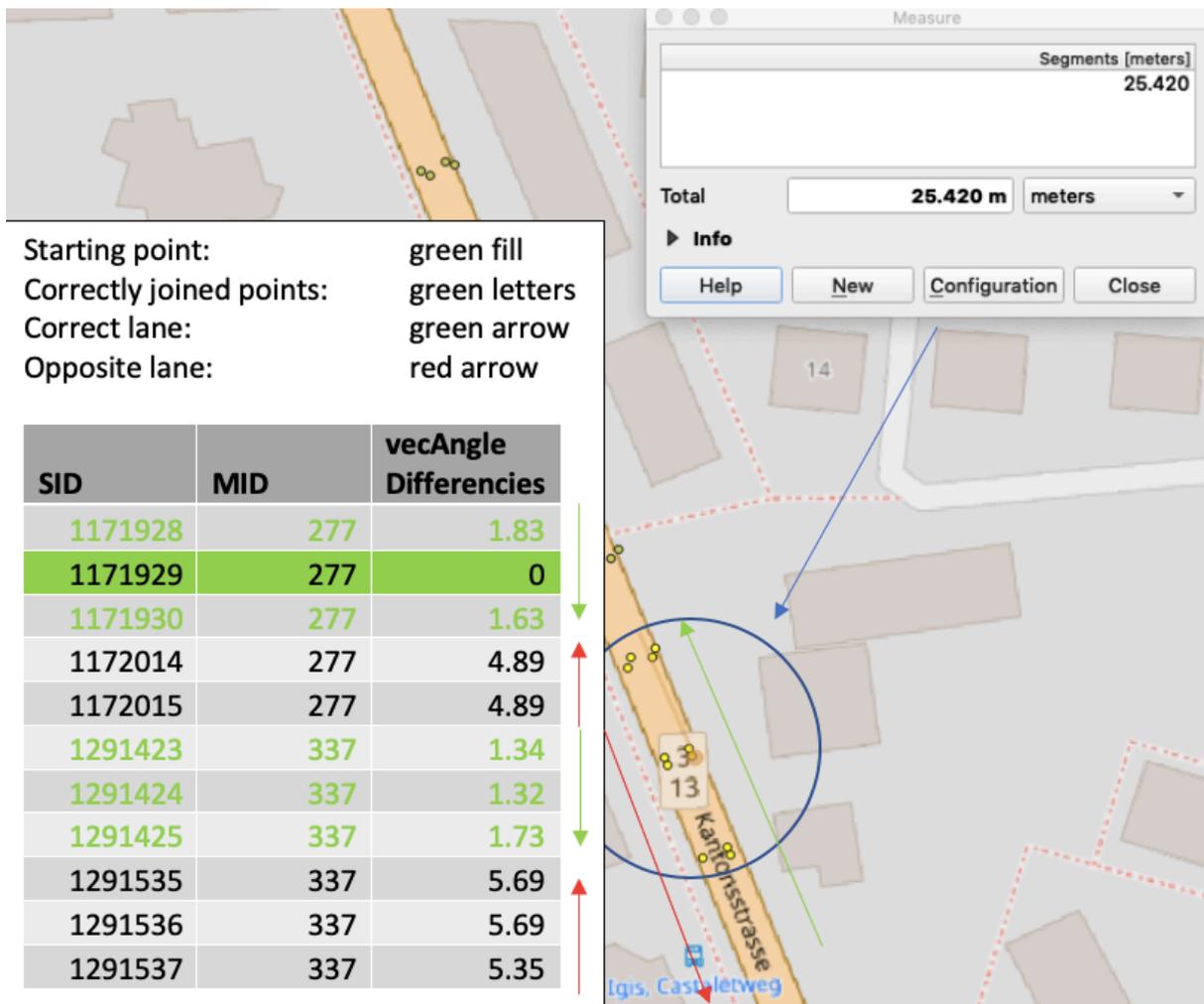
Figure 3.4: Step 3 of Coordinate ID algorithm, the algorithm checks whether a segment is correctly joined to a lane by controlling if the value of the SID of the potential point is exactly + 1 of the SID from the point with the 25 meters search radius.

vecAngle and deltaAngle. In order to investigate whether curves differ from straight sections (e.g. by stronger shear forces), the two attributes vecAngle and deltaAngle were created.

The attribute vecAngle is the 2D azimuth calculated from the start and end coordinates of a segment. 0° is defined as the geographical north. The rotation is clockwise.

The attribute deltaAngle is the difference between the vecAngle value of a segment with the vecAngle value of the immediately preceding segment. The greater the deltaAngle value, the sharper the bend.

3.1.3 Cleaning

Road surface type and road surface ID. The customer-oriented implementation of the database is also noticeable in the road surface ID. All segments belonging to the same road surface have the same road surface ID. However, a road surface was always assigned a new road surface ID during a repeat measurement. By joining the segments of different measurement series to a Coordinate ID, the different road surface IDs can be identified. The lowest value was selected and all other segments with the same Coordinate ID were assigned this value for the road surface ID attribute.

When joining the different measurements, it can also be seen that the attribute road surface type for a location is not consistent. Possible reasons are modified data of the customer which were entered at a different time during the repeat measurement, an assignment to a supercategory (e.g. Famsi defined as 4mmMV), or the installation of a new road surface.

If the assignment to a supercategory took place, then the subcategory is assigned to the Coordinate ID. The reason for this is that, as described in Section 2.2.2.1, the standards for a road surface type are rather too tolerant. If a road surface within one category would differ significantly from the other road surfaces, then this road surface type could be recognized and treated separately and as a separate category.

If a new road surface has been installed (recognizable if later measurements are much quieter than earlier measurements), a new road surface ID with the previous value + 10'000 is created. This ensures that each road surface has a unique road surface ID.

3.2 Traffic data

All attributes related to traffic volume are derived from the sonBase GIS noise database. The SonBase data is segmented into straight, unevenly long lines characterized by traffic volumes. For each Coordinate ID the values of the geographically closest line were joined.

Intersections. Where lines of the sonBase dataset cross each other, nodes were created. These nodes represent intersections. The Intersection attribute contains the shortest distance of a segment to the nearest intersection.

Bus stops. The public transport stop data set of the Swiss Federal Office of Transport includes the complete set of bus stations. For each segment that has a distance <= 20 m to a bus stop, the value is set to 1. Segments with a distance > 20 m receive the value 0.

3.3 Climate data

Temperature and precipitation. Data from MeteoSwiss Spatial Climate Analyses were used to generate temperature and precipitation variables. With a monthly time interval and a spatial resolution of 1 square kilometer, the attributes average temperature, standard deviation of temperature, total amount of precipitation, average precipitation, and standard deviation of precipitation were calculated for each segment. In order to determine the installation month of a road surface, the few data available were used. Of a total of 2442 road surfaces, only 119 were known to have been installed during a specific month. The analysis of these 119 road surfaces showed that the vast majority were installed in May, June, July, August and September. Therefore, the month of July was determined as default value.

Climate indicators. The 5 climate indicators frost days (days on which the temperature falls below 0°C), ice days (days on which the temperature remains below 0°C), summer days (days on which the temperature reaches 25°C or higher), heat days (days on which the temperature reaches 30°C or higher) and tropical nights (days on which the temperature does not fall below 20°C) and their averages and standard deviations were generated.
The data set was obtained from the Swiss Federal Statistical Office, which documents data from 13 measuring stations distributed throughout Switzerland. Each point was assigned a measuring station from which the values of the 5 indicators were derived.
The determination of the most representative measuring station for a point was determined as follows:

1. A data set with high spatial resolution from MeteoSwiss, containing the average number of frost days per year, was obtained.
2. Each segment was assigned the average number of frost days of the MeteoSwiss dataset with a spatial join, using the closest distance criterion.
3. For each segment, the measuring station with the smallest deviation of frost days was indicated. If two or more stations have had exactly the same deviation, the values of the closest station were joined.

Solar radiation. Since temperature, precipitation and the 5 climatic indicators have a much lower spatial resolution than the segments (1 square kilometer vs. 20 m) the solar radiation was calculated to approximate the microclimatic conditions. The solar radiation was calculated using the ArcGIS Desktop (Version 10.5) toolbox "Points Solar Radiation". The SwissToppo's digital surface model with a grid size of 4 square meters served as the input surface grid. The midpoints of the 20 m segments served as input point shapefile. The direct, diffuse and duration of direct solar radiation were calculated for each month of 2018. These months were used as a reference for all months of the other years. The total, average and standard deviation were then calculated for direct, diffuse and duration of direct solar radiation. 32 azimuthal directions were used to calculate the field of view for a point. This number of directions is appropriate for complex topography. The incoming diffuse radiation was for the sake of computational speed assumed to be the same from all directions.

3.3 Distance to farmland and construction sites

The "Areal Statistics" developed by the Swiss Federal Statistical Office was used for this purpose. A shapefile was created which only contains the codes 221 (= arable farming), 146 (= construction sites) and 145 (= dismantling). For the variable distance to farmland (distanceAgrar) the shortest distance to the closest location with the code 221 point was taken. For the variable distance to construction sites (distanceConstruction) the shortest distance to codes 145 or 146 was used.

3.4 Topographic data

All topography attributes were generated using the SwissAlti3D terrain model with a raster resolution of 4 square meters.

3.4.1 Attribute construction

First, the z-coordinate was added by spatial join with the points. With the use of the z-coordinate (elevation) the other attributes could be constructed. The terrain-roughness-index (TRI) was calculated using the ArcGIS Desktop (Version 10.5) toolbox "Calculate Terrain Ruggedness Index (TRI) on DEM" with a 20*20 moving window.
Slope and aspect were calculated within a 3*3 moving window. The difference between aspect to the vecAngle variable is that the aspect represents the exposure of a segment in three-dimensional space, whereas vecAngle represents the exposure in two-dimensional space.

3.4.2 Cleaning

The slope shows outliers. This is because if a road bordered steep terrain and the GPS signal was not exactly positioned on the road, the slope of the steep terrain bordering the road was calculated. All values with a slope > 20° or < -20° were replaced by the average slope of the entire track to which the segment with the falsely calculated slope belongs. The thresholds were chosen as ± 20° because LNAs are normally installed in densely populated urban areas where roads have barely a slope steeper than 20°.

3.5 Data preparation

This section describes how the data was made compatible for the machine learning algorithms.

3.5.1 Standardization

Since the ranges of the variables (e.g. avgDirect vs. avgT) can be very large, the variables were standardized using z-transformation.

3.5.2 Numerical coding

If the machine learning tool or algorithm can only consider numerical attributes, the nominal variables Canton and PavType were dummy coded (= creating binary pseudo-variables for each attribute value).

3.5.3 Binning

The variables Intersection, distanceAgrar and distanceConstruction were not selected for any of the models (Section 4.5) in the variable selection. Since, for example, it could be that the strongest mechanical load is not experienced directly at the intersection itself, but rather near the intersection, since braking and acceleration of the vehicles take place there, all three distance-based variables were nominally coded. Using K-Means, with k = 5, the variables were divided into 5 classes with the names "—","-","=", "+", "++"("++" is the class with the longest distances, "—" the class with the shortest distances).

However, none of this binning has led to a change, which means neither intersection, distanceAgrar nor distanceConstruction were selected by the variable selection algorithm, although they were binned. Hence, the binning was annulled.

In order to maximize the interpretability of the model, the road surface types were binned. The 5 groups SDA4 with low Voids, 8-12% volume content, (4mmLV), SDA4 with medium Voids, 16% volume content, (4mmMV), SDA4 with high Voids, 20% volume content, (4mmHV), SDA8 with low Voids, 8-12% volume content, (8mmLV) and SDA8 with medium

Voids, 16% volume content, (8mmMV) were formed. This happens after sampling the data set for the reasons described in Section 4.2.

3.5.4 Outliers

Although the data were manually cleaned, they contain outliers. The problem is that the reason for the remaining outliers is not known. Only hypotheses can be made. For example, the weather conditions during the installation of a road surface could have been unfavorable. Perhaps a company has experimented with the physical parameters of its products without changing the name or using a bitumen other than usual.

Outliers can provide important information about the process or system (Basu and Meckesheimer 2007). Since it is not possible to know why the remaining outliers are outliers and this thesis is exploratory, they were not removed from the data set.

4. Modelling

4.1 Overview of the modelling process

Since this thesis has a strong focus on quantitative modeling, it does not follow the classical standard approach Introduction, Method, Results and Discussion. In this chapter a systematic explanation is given of why a particular modelling technique, test design, hyperparameters and variables were chosen. Consequently, the results in this chapter are primarily intended to quantitatively justify the decisions made in the modelling process.

This chapter shows the modelling process as follows. Section 4.2 describes the planned plan for sampling, training, testing and evaluation of the models. In Section 4.3 the modelling technique is determined by comparing different modelling techniques in terms of performance and interpretability. Section 4.4 describes the hyperparameter optimization of the selected modelling technique and Section 4.5 describes the variable selection. All these steps (Sections 4.2-4.5) lead to Section 4.6, the creation of the final model, which is used to answer the research questions. The results and their discussion of the final model (Section 4.6) are described in Chapter 5.

4.2 Generate test design

The focus of this thesis is on the LNAs. Since the Grolimund + Partner database has stored a number of data of conventional road surfaces, these conventional road surfaces are investigated separately as a conventional group (CG).
As shown on the maps (Appendix A3) or in the data quality report (Appendix A2, canton variable), the data are spatially unevenly distributed. To quantify the effects of this uneven distribution, the Moran's Index of the target variable was calculated using the R package ape by Paradis and Schlief (2018). As the Moran's Index could not be calculated for all data at once due to the limited computing power available, the Moran's Index was calculated for four subsets of the training data. The average of the Moran's Index values of the four subsets is 0.41 and significant, which means there is spatial autocorrelation present in the data. A more detailed discussion of the effects of autocorrelation on the interpretation of the results is given in Section (5.1). The spatial autocorrelation is caused by two problems: Firstly, the typical environmental influences of data with a high spatial density (e.g. Aargau) are weighted more strongly than data with a low spatial density (e.g. Alpine region). The classical data mining approach would now prefer to weight the variables differently (i.e. those data which are underrepresented would be weighted more heavily) or to regionalize the data according to

environmental factors and create test and training data for each region separately. However, an adequate climatic classification and weighting of the data is difficult. For example, during the summer months parts of the Valais (a canton in the Alps) can experience temperatures above 30 degrees Celsius over a longer period of time as well.

The second problem is that most cantons obtain their road surfaces from different construction companies. As already described in Section 2.2.2.1, the tolerant norms of the types of road surface lead to the fact that the products of different companies behave differently. This is a problem as cantons generally only purchase their LNAs from certain companies, meaning that the types of LNAs often vary from canton to canton. Additionally, the most common products invariably receive more weight than the rare ones. However, since there are 32 different LNA products in total included this data set, the countermeasure for balancing unbalanced data — bagging — would reduce the amount of data per bag to such an extent that machine learning algorithms would hardly provide meaningful results in this area of application.

To quantify the size of the error, based on the problem of spatial heteroscedasticity and regionalization based on climatic conditions, the best performing model of Section 4.3, the Gradient Boosted Tree Model, was used. Three models with the following three observations were tested.

1. Training on Aargau data and testing on Geneva data will perform poorly because road surfaces in Geneva are strongly different from those in Aargau although the environmental influences are similar.

2. Training on Geneva data and testing on Valais data will perform poorly because the environmental influences in the two cantons differ greatly (traffic load larger in Geneva, winters harsher in Valais) although Geneva and Valais use the same road surface products in the majority of cases.

3. Training on Aargau data and testing on Zurich data will perform well, as the two cantons are exposed to similar environmental influences and use similar road surface products.

Training Test	Geneva	Valais	Zurich
Aargau	4.338	X	1.537
Geneva	X	3.863	X

Table 4.1: Overview of the regional test results, values are the RMSE performance metric in dB(A)
a model, X means that no model of this certain case was tested

When considering the results, the three aforementioned observations seem to be correct and training and test data should be as balanced as possible with regard to environmental

influences and road surface types. Hence, sampling considered that the relative distribution of LNAs per canton is equal in the training and test data.

In addition, all data points of a road surface, i.e all segments that have the same PavID, are only found in the test data or only in the training data. This is to counteract the spatial autocorrelation.

4.3 Select modelling technique

The models Generalized Linear Model (GLM), Random Forest (RF), Decision Tree (DT), and Gradient Boosted Trees (GBT) were created using RapidMiner Studio version 9.2 (Mierswa, I., & Klinkenberg, 2019). All RapidMiner Studio models underwent the RapidMiner hyperparameter optimization (details in Section 4.4) and variable selection (more details in Section 4.5).

The linear model tree (LMT) was implemented using the Python source code by Dillard (2018). The linear model tree underwent a separate hyperparameter optimization (Section 4.4) and variable selection (Section 4.5). The mean squared error (MSE), the root mean squared error (RMSE) and the mean absolute error (MAE) were chosen as performance metrics to compare the models.

Performance Metric	GLM	DT	LMT	RF	GBT
MSE	4.088	3.936	3.105	3.014	2.897
RMSE	2.022	1.984	1.762	1.736	1.702
MAE	1.544	1.449	1.304	1.053	1.243

Table 4.2: Overview of the test results (in dB(A)) for model comparison

The results show that RF, GBT and LMT perform best. It is apparent that RF has the lowest MAE, but has an RMSE between the RMSEs of LMT and GBT. This means that RF estimates many values very precisely, but deviates relatively strongly from the actual value in the case of larger errors. This is a sign of overfitting. Additionally, an LMT is easier to interpret than an RF, so RF is not determined as the final model technique.

Interpretability plays an important role in this thesis, the more interpretable model is given precedence. According to Gill and Hall (2018), linear, monotonous models such as GLMs are highly interpretable. Non-linear, non-monotonic models such as GBTs, RFs and LMTs are the more difficult models to interpret. In contrast to the GBT and RF, however, LMT has the advantage that each leaf represents an individual linear model (the LMT model for this thesis has individual ridge regressions (Tikhonov 1943) as leaves), which in the end facilitates high interpretability. Therefore, LMT is chosen as the model for this thesis.

4.4 Hyperparameter optimization

RapidMiner hyperparameter optimization was set on the classical way of hyperparameter optimization, the grid search. Grid search (Hsu et al. 2003), is an exhaustive searching through a manually specified subset of the hyperparameter space of a learning algorithm. Suppose a modeling technique has two continuous hyperparameters x and y. The grid search then trains a model in the cartesian product of these two sets with each pair (x, y) and evaluates their performance on a sustained validation set. The best performing pair then sets the values for the hyperparameter.

The hyperparameter optimization of the LMT model was performed with the variables occurring in the decision tree induction. Adaptable hyperparameters are the minimum number of data points that must reach a leaf during the training phase (Figure 4.1, 4.2, y-axis) and the minimum improvement of the MSE (Figure 4.1, 4.2, x-axis) to create a new leaf. Since the LMT has a relatively long runtime (details in Section 4.5), a mix of manual and grid hyperparameter optimization was chosen. The results are visible for LNAs in Figure 4.3 and for CG in Figure 4.4. The LNAs show two local minima of the MSE. At $(y = 3300 / x = 0.1)$ and $(y = 1000 / x = 0.1)$. If y becomes smaller, more leaves, i.e. more ridge regressions, are formed. The more ridge regressions there are, the more difficult it is to eventually interpret the model, however. Therefore, $y = 3300$ and $x = 0.1$ were determined as the ideal size for the LMT-LNA model hyperparameters.
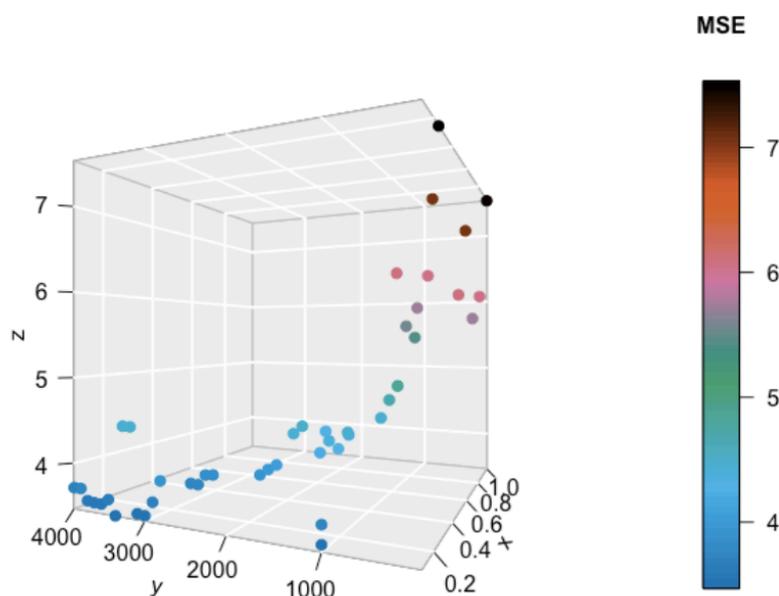
x = minimum MSE split improvement | y = minimum node size | z = performance metric MSE

Figure 4.1: Visualization of the MSE for the optimal hyperparameters of the LMT model for the LNA data

The hyperparameter x plays a minor role for conventional road surfaces (Figure 4.2). The most important factor is y. However, y is much less sensitive for conventional road surfaces than for LNAs. As with LNAs, as few leaves as possible should be created with conventional road surfaces. Therefore, the largest y (26'000) was chosen at the local minimum. As with the LNAs, x is 0.1.
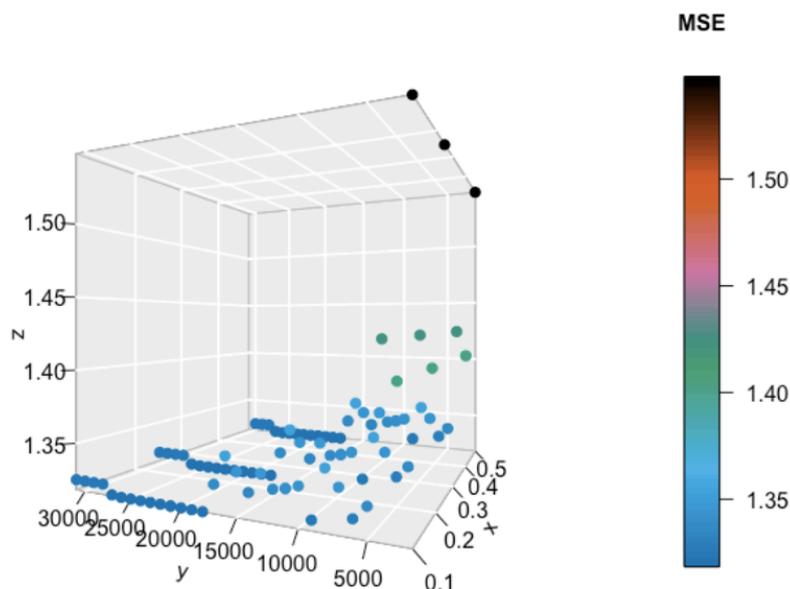
x = minimum MSE split improvement | y = minimum node size | z = performance metric MSE

Figure 4.2: Visualization of the MSE for the optimal hyperparameters of the LMT model for the CG data

After setting the minimum number of data points per leaf (y) and the minimum MSE improvement for a split (x), a grid search was used to search for the optimal alpha of the ridge regression. As shown in Table 4.3, tweaking alpha does not result in significant differences, therefore the default value (alpha = 1) was used.

alpha	0	0.25	0.75	1	2	4	8	16
Training MSE	2.881	2.881	2.881	2.881	2.881	2.880	2.881	2.882
Test MSE	3.106	3.106	3.106	3.106	3.106	3.106	3.105	3.102

Table 4.3: Overview of alpha optimization

4.5 Variable selection

Variable selection is the process of selecting a subset of relevant variables, predictors for use in model construction in order to avoid the curse of dimensionality, overfitting and simplify

the interpretation. Table A.1 in the Appendix shows which variables were taken into account in the variable selection and which ones were not, and for what reason.

The RapidMiner variable selection uses the two deterministic greedy feature selection algorithms forward selection and backward elimination. Forward selection always adds the variable that minimizes the MSE the most. Backward elimination functions similar as forward selection. Backward elimination starts with all variables and always removes the one variable which maximizes the MSE the most. Stepwise forward selection and backward elimination are disputed methods regarding variable selection. Harrell and Frank (2015) show in their work that forward variable selection has several deficiencies. The main problem is that forward variable selection only takes place on the basis of one assumption. In this thesis, the assumption is that if a variable out of the total set of variables is the one that lowers the MSE the most, it is the best of the total set of variables and should be selected as long as the MSE decreases by at least 0.05 for the LNAs. Consequently, the following problem may arise: If $X(1)$ is the best individual variable, it is not guaranteed that either *{X(1), X(2)}* or *{X(1), X(3)}* are better than *{X(2), X(3)}*. Therefore, a forward selection algorithm may select a variable set different from that selected by exhaustive searching. With a poor selection of the input variables, the prediction *Y(q)* of a query *X(q) = {X(1), X(2), … ,X(m)}* may be significantly different from the true *Y(q)* (Deng 1998). However, the advantage of the greedy forward variable selection is that it requires less time than exhaustive variable selections.

The computer used to calculate the models of this thesis has a 1.4 GHz processor and 8 GB of memory. The runtime to create one LMT is approximately 10 minutes. This thesis has a total of 55 different variables to choose from, a complete run through all variables with the forward selection approach takes approximately 10.7 days. Since an exhaustive search would extend the running time even further, the forward variable selection is used to determine the variables in this thesis.

4.5.1 LNA variable selection

From the points plot (Figure 4.3) of the LNA variable selection, the elbow point cannot be clearly identified. However, in the numerical representation in Appendix A.4.1.1, the gap where the MSE hardly improves, namely after the variable sdHD, is well recognizable. Hence, in Figure 4.4.1. this point is highlighted in red.

In order to counteract the disadvantages of the stepwise forward selection, further models were manually tested for their performance, which do contain or, conversely, not contain certain variables close to the elbow point. The resulting Table A.4.1.2 is presented in the Appendix. This table is arranged in such a way that the top model has the smallest difference between training and test error, i.e. the least overfitting. It is noticeable that the two top models perform almost equally well. Whereby the top model is not the one that would suggest the variable selection (sdHD was replaced by avgDiffuse). To compare which model is more stable, a 5-fold cross-validation (Table A.4.1.3) was carried out with both of the two least overfitting models. Cross-validation shows that the model that performs better and has less overfit is the same model as the original forward variable selection suggested. Therefore, the following variable set is chosen to create the final LNA model: nMonths, 4mmMV, 4mmHV, 8mmLV, 8mmMV, sdDirect, DTV_LKW, avgFD, sdHD.
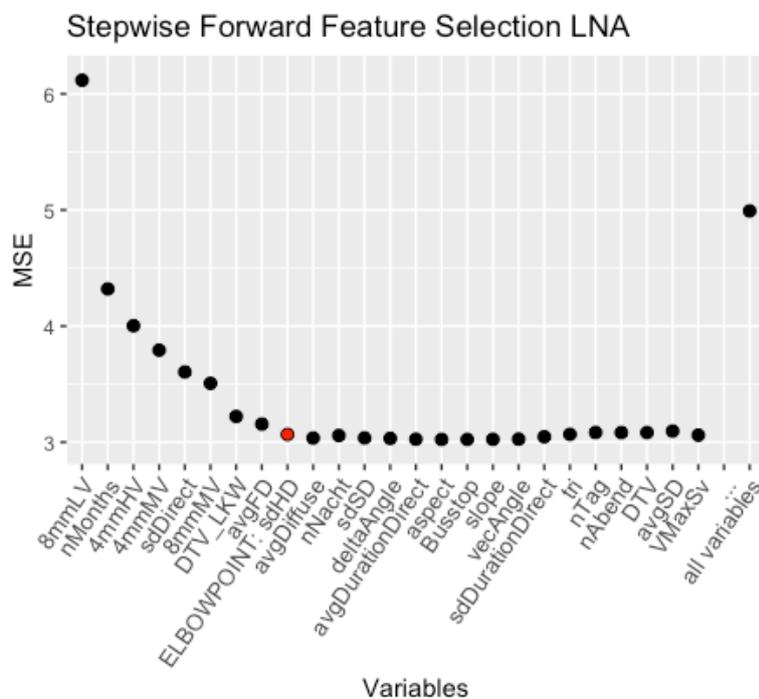
Figure 4.3: Elbow plot of LNA variable selection, MSE in dB(A)

4.5.2 CG variable selection

From the points plot (Figure 4.4) of the CG variable selection, the elbow point can be recognized. In the numerical representation in Table A.4.2.1, the gap where the MSE hardly improves after adding the variable sdT, is also recognizable (hence, this variable has been highlighted in Figure 4.4.2). In the resulting Table A.4.2.2, it can be observed that the underfitting hardly changes after the model with the variables: nTag, VMaxSv, avgDiffuse, DSAK, avgDirect, avgTN. If now the variable sdT is added, the underfitting does not change much, but the MSE decreases relatively strongly (Table A.4.2.1) for the ratios of the CG dataset. Therefore, the following variables are used for the final CG model: nTag, VMaxSv, avgDiffuse, DSAK, avgDirect, avgTN, sdT.
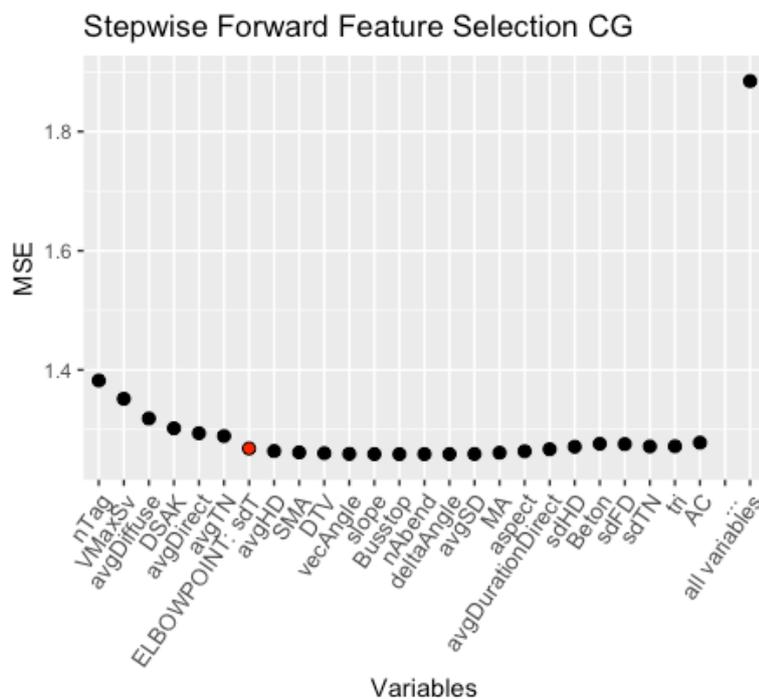
Figure 4.4: Elbow plot of CG variable selection, MSE in dB(A)

4.6 Final model

The final model is a Linear Model Tree whose leaves are ridge regressions. The hyperparameters for each road surface group are summarized below.

4.6.1 LNA

Training data:	57 % of the dataset
Test data:	43 % of the dataset
Minimum number of data points per leaf for the training data:	3300
Minimum mean squared error improvement to create a new split:	0.1 dB(A)
alpha hyperparameter:	1

Variable set = nMonths, 4mmMV, 4mmHV, 8mmLV, 8mmMV, sdDirect, DTV_LKW, avgFD, sdHD

4.6.2 CG

Training data:	52 % of the dataset
Test data:	48 % of the dataset
Minimum number of data points per leaf for the training data:	26'000
Minimum mean squared error improvement to create a new split:	0.1 dB(A)
alpha hyperparameter:	1

Variable set = nTag, VMaxSv, avgDiffuse, DSAK, avgDirect, avgTN, sdT

5 Results and discussion

5.1 LNA

5.1.1 Model evaluation

Performance metrics. The performance metrics for the final LNA model are summarized in Table 5.1.1.1.

Performance metrics	Training Error	Test Error
MAE	1.281	1.304
RMSE	1.697	1.762
MSE	2.880	3.106
R²	0.693	0.643

Table 5.1.1.1: Overview of the performance metrics of the final LNA model in dB(A)

Table 5.1.1.1 suggests that the model has a slight overfit. Furthermore, the table shows that a standardized performance metric is important. An MAE of approximately 1.3 dB(A) seems fine, but when considering R^2 it can be observed that the model can explain just 65-70 % of the variation. An R^2 of 65-70 % might be an acceptable generalizable capability, since the model is capable of delivering such a precision across the entire country of Switzerland. However, as discussed in Section 3.5.4, the data were not cleaned up with respect to outliers. To find out if these outliers are distributed randomly over the tracks, or if usually an entire track is an outlier, the residuals were clustered with k-Means, where k equals 5. K was set to 5 because studies like Dawes' (2008) showed that a five-point scale facilitates interpretability. The interpretability is important in order to be able to interpret what causes how severe outliers. All residuals that were among the two most extreme clusters, i.e. either too loud or too quiet, were counted per track. Figure 5.1 shows that the residuals are usually randomly distributed over the distances (the highest bar by far is the one with only one extreme residual per track). Nevertheless, Figure 5.1 also suggests that there might be entire tracks that are outliers. In order to find out why entire tracks can be outliers, the method of subgroup discovery (Görz et al. 2012) was applied (see below). For the subgroup discovery, each attribute was clustered with k-Means, where k equals 5. Each of this cluster functioned as a new dummy coded variable. Additionally, each canton, road surface type, and leaf (more information in Section 5.1.2) was dummy coded as well. For each prediction error cluster, a separate subgroup discovery model was created, where the prediction error cluster was the target variable and the clusters of all other attributes were selected as independent variables (composition in Appendix 5.1.1). Table A.5.1.1 shows that the extreme outliers account for merely 15% of the

total data set. Table 5.1.1.2 shows the performance of the model if the data had been cleaned from outliers.

Performance metrics	Training Error	Test Error
MAE	1.129	1.198
RMSE	1.544	1.598
MSE	2.302	2.770
R²	0.751	0.680

Table 5.1.1.3: Overview of the performance metrics of the final LNA Model in dB(A) without outliers.

It is interesting to note that R^2 improves in training and test data by 10 %. An R^2 of 75-79 % is a satisfying result, considering that the best possible R^2 value is approximately 87 % because of the target variable's large measurement error (Section 3.1.1). Hence, the influence of the outliers on the performance is strong. Therefore, it is all the more important to understand why certain data points are outliers. Due to the very long calculation time of the variable selection procedure described in section 4.5, there no attempt was made to create a new model in which the outliers are removed from the data during preprocessing, which would have required rerunning the entire procedure of Chapter 4, including the selection of modelling technique, hyperparameter optimization, and variable selection. To test whether the above assumptions that outlier cleaning improves R^2 of the model by about 10%, however, a model was created in which the standard deviation was calculated for each pavement measurement and all segments outside the standard deviation were removed from the data. The hyperparameter optimization for this model suggested a minimum MSE split improvement of 0.1 and a minimum node size of 2000. Table 5.1.1.3 shows the performance of the model with the cleaned data by excluding segments with a value outside the standard deviation of a pavement measurement.

Performance metrics	Training Error	Test Error
MAE	1.060	1.073
RMSE	1.544	1.598
MSE	2.383	2.554
R²	0.795	0.755

Table 5.1.1.4: Overview of the performance metrics of the final LNA Model in dB(A) with the cleaned data by excluding segments with a value outside the standard deviation of a pavement measurement.

Note that by simply excluding segments with a value outside the standard deviaton of a pavement measurement, R^2 improves in training and test data by 5 %. If greater efforts were made, such as cleaning the data with outlier detection during preprocessing as input to

hyperparameter optimization and variable selection, performance could be further improved and possibly actually achieve the assumed 10 % improvement in R^2.
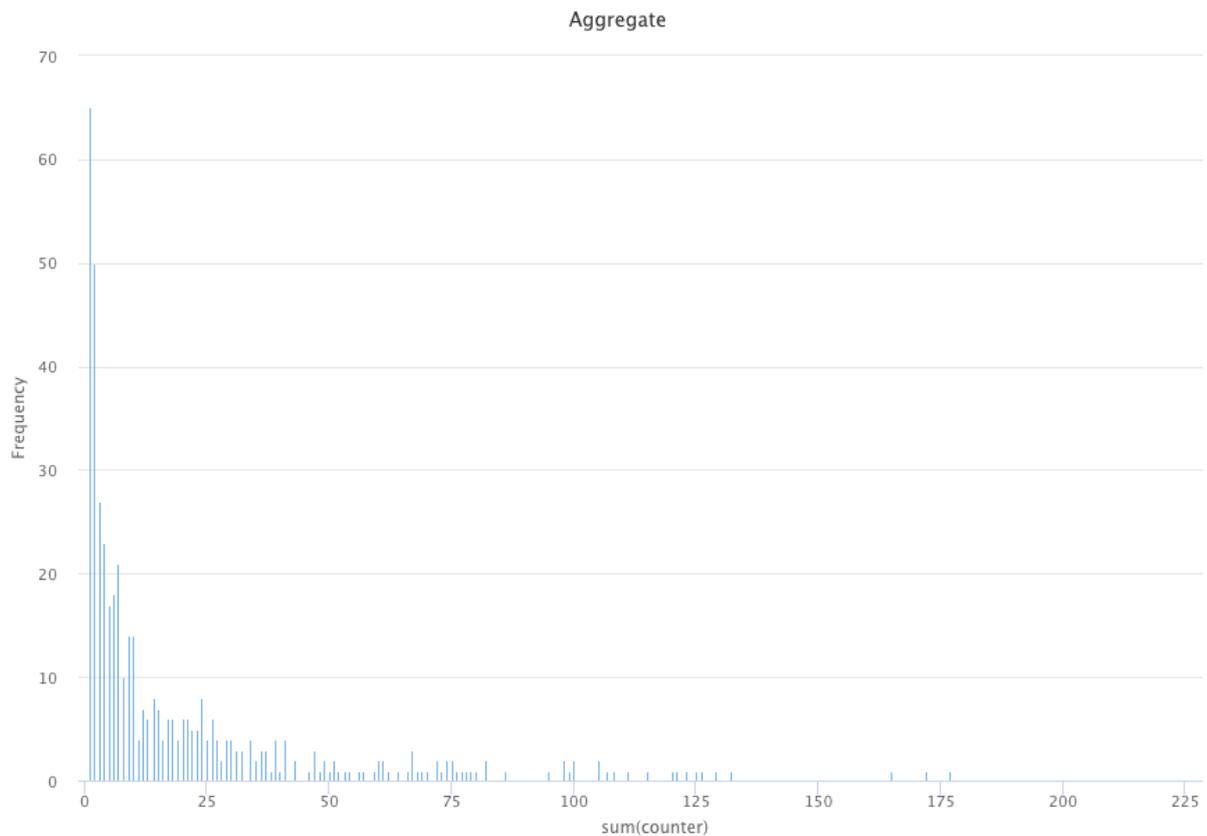
Figure 5.1: x-Axis: number of outliers per track over all years, y-Axis: number of tracks

Subgroup discovery and outlier explanation. According to Görz et al. (2012) subgroup discovery can be defined as follows:

Subgroup discovery, tests the quality of hypotheses $q(h)$, with $q(h) = \sqrt{g} \times |p - p'|$, with

p = the probability with which the target attribute in the subgroup of h has a certain value;

p' = the probability with which the target attribute in the total population has this value;

g = the relative size of the subgroup of h to the total population S and $g = |h|/|S|$.

The advantage of subgroup discovery is that the method can also be used if the target attributes for a classification are distributed highly unequally, which is the case for our data, as the two most extreme outlier clusters account for 15 % of the entire data set. In addition, subgroup discovery describes the groups, which simplifies interpretation. The subgroup discovery models were created using RapidMiner Studio, version 9.2 (Mierswa & Klinkenberg 2019). K best rules, was chosen as mode, while k equals 5. The weighted relative accuracy (WRAcc) of a rule was chosen as the utility function, which determines the five best rules.

Since WRAcc measures generality[1], precision[2] and interest[3] of a subgroup, it is a reasonable compromise of these three factors and therefore a recommendable criterion to create subgroups (Herrera et al. 2011). The results of the five subgroup discovery models are reported in Table A.5.1.2.

The five rules of the most negative residuals ("Cluster Pred --" in Table A.5.1.2) suggest those data points that the model wrongly estimates as far too loud (group mean of -3.084 dB(A)), indicating that most of these data points are either rather young (RegressionLines = TLL, nMonths = --; group mean of 19.8), are located in the canton of Geneva, or of the road surface type Nanosoft4.

These subgroups thus show the effect of spatial autocorrelation. Nanosoft4 road surfaces are a 4mmHV product that is mainly installed in the canton of Geneva. Consequently, the model in the canton of Geneva often wrongly estimates the road surface as too loud. This is crucial, as Table 5.1.1.4 shows.

Canton	Geneva	Valais	Fribourg	Vaud
Relative % of total 4mmHV per canton	59.6	18.9	16.8	5.7
Nanosoft4 share	55.3	32.8	48.3	9.9

Table 5.1.1.4: Overview of the cantonal distribution of the 4mmHV and the Nanosoft4 road surface

Since more than 50 % of the total 4mmHV data is located in Geneva and more than 50 % of this 4mmHV data is of surface type Nanosoft4, approximately one fourth of the entire 4mmHV data is likely to be a strong outlier that was estimated as too loud.

The five rules of the most positive residuals ("Cluster Pred ++" in Table A.5.1.2) represent those data points that the model wrongly estimates to be much too quiet (group mean of 4.008 dB(a)), indicating that most of these data points are rather young (nMonths = -- , avgFD = --, sdHD = --). AvgFD and sdHD are both variables that have to develop first over time (more detail on this in Section 5.1.2) and have a right-skewed distribution. Low values of these variables speak above all for a young age of the road surface.

Bias. Bias can be defined differently. In most research fields, bias refers to a systematic error in quantity, which is caused by a false modelling approach, while trying to estimate the true value of a certain parameter.

Regarding machine learning bias is defined as the difference between the estimated value of a model and the true value of the target variable. However, low bias can cause low

---

[1] Quantifies the quality of individual rules according to the individual patterns of interest covered.
[2] Measures the tradeoff between the true and false positives covered in a lineal function or measures the tradeoff of a subgroup between the number of examples classified perfectly and the unusualness of their distribution.
[3] Measures are intended for selecting and ranking patterns according to their potential interest to the user by calculating novelty, gain or significance.

performance due to high variance. Variance means the difference of fits of a machine learning model between different data sets. In the evaluation of a machine learning model, bias and variance are therefore considered, both of which should be as low as possible. The bottom line is to know how well the model will perform when deployed. There are many possibilities to measure bias and variance, but the most common method is probably cross-validation. Kohavi (1995) recommends a 10-fold cross-validation in his work. Since in this thesis the samplings are double stratified and strongly unbalanced, it is difficult to obtain an acceptable stratification with 10 folds, therefore a 5-fold cross-validation is used. The R^2 95% Confidence Interval of the 5-fold cross-validation equals 62.6 % $\pm$ 2.4 %. This shows that the method to create the model is robust. This does not imply that the results analyzed in Sections 5.1.2 and 5.1.3 are statistically significant. But since R^2 of the test data of the final model is close to the mean of the 95 % confidence interval of the 5-fold cross-validation, the modelling approach seems to be robust and with the critical analysis of the outliers by means of subgroup discovery, cautiously formulated observations can and should be made.

This means that with regard to the analysis of the model in the following sections, the non-machine learning definition of bias, i.e. bias implying a systematic error, must be considered. In concrete terms, it must always be noted that 4mmHV road surfaces in the Geneva region are considered to be too loud and due to the large proportion of 4mmHV in Geneva, 4mmHV are generally considered to be too loud by the model, and young road surfaces (in the order of < 24 months) can hardly be predicted well. In addition, the present spatial autocorrelation within the dataset has the consequence that significance tests regarding coefficients could be erroneously accepted or rejected. In general, the values of the coefficients can vary strongly, therefore only the rough direction of the coefficients should be considered.

5.1.2 Tree structure

When considering the created tree (Figure 5.1.2), three main groups that influence the aging behavior of LNAs can be identified: Age, climate and traffic. The letters in the leaves stand for the allocation path that a data point has taken. T stands for top, L for left and R for right. Consequently, a data point that ends up in the TLL leaf has a value less than 46 for the nMonths variable and less than 67317.85 for the sdDirect variable.
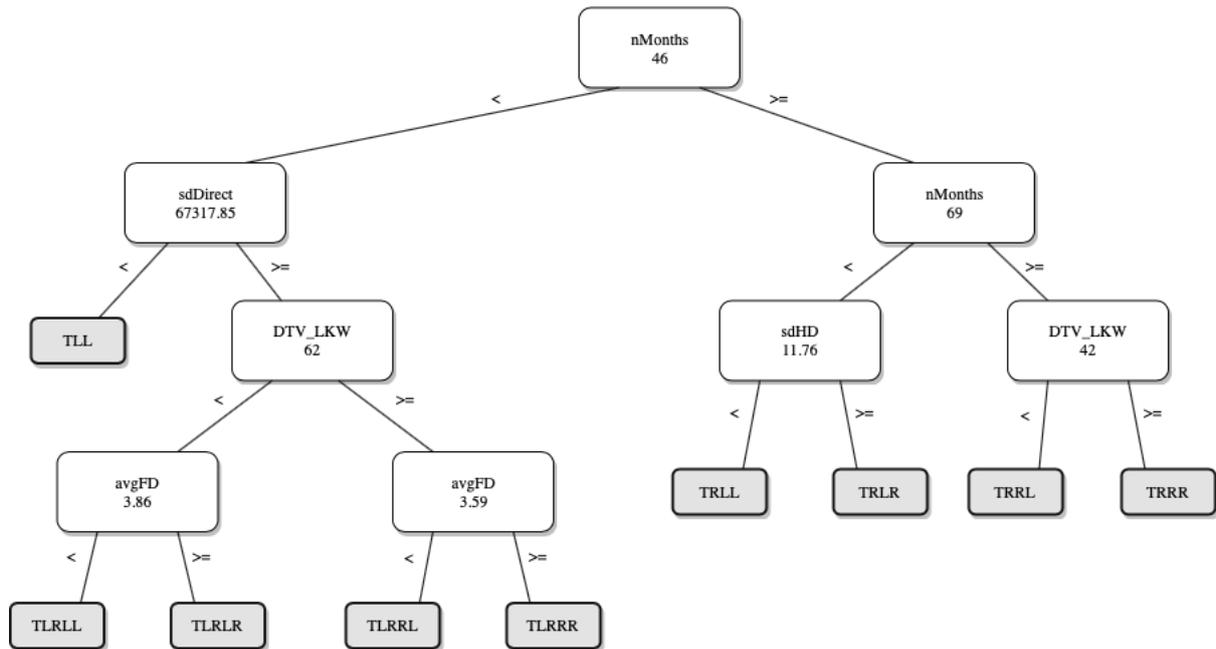
Figure 5.1.2: Tree structure for the final LNA model created by the LMT algorithm

Age and climate. The age of an LNA is the variable that contributes most to explaining the loss of the noise reducing effect of an LNA. This is, because the first split at top-node occurs through the nMonths variable, which was also the second variable selected by the stepwise forward variable selection. In addition, other split variables probably also measure age. The nMonths Boxplot (Appendix A.5.1.4) shows that TLL is the leaf with the youngest LNAs. This happens because the split is induced by the sdDirect variable. When looking at the sdDirect variable in data quality report (Table A.2.1), the first quantile is much closer to the mean than the third quantile, which indicates a right-skewed distribution. This right-skewness occurs because the sdDirect variable has to develop over time at the beginning; in the first year of a road surface, direct solar irradiation values are considered between the months from the end of March to the end of October at most. Frequently, however, it the timeframe is only a few summer months. Therefore, the standard deviation of direct solar radiation of young road surfaces is small. The split value of the sdDirect Variable is also smaller than its first quantile and thus successfully forces the presence of very young LNAs in TLL. In this sense, TL can

also be interpreted as measuring age. With the splits of the avgFD variables TLRL and TLRR it is not directly apparent whether the splits occur due to age or climate. The avgFD variable must also develop over time; the older a road surface is, the more winter and frost days will have occurred. The nMonths boxplot shows that the leaves that were formed on the left by an avgFD split (namely TRLL and TLRRL) are younger than the leaves on the right side of the same split (TLRLR and TLRRR). Additionally, the nMonths variable has a right-skewed younger than leaves on the right side of the same split (TLRLR and TLRRR). Additionally, the nMonths variable has a right skew distribution in the younger leaves and the older leaves have a left-skewed distribution. This means that the left leaves TRLL and TLRRL have young LNAs or older (age is approximately between the first and third quantile of the TLRLR and TLRRR nMonths boxplots) LNAs with few frost days, i.e. warmer climates. The opposite is true for the right leaves TLRLR and TLRRR. AvgFD splits are therefore probably due to climatic and aging criteria.

Heavy vehicles (HV). If it is supposed that the sdDirect-split is at TL measures, age, then it can be expected that the share of heavy vehicles is the second most explanatory variable. Each split on the second level (either TL or TR) is followed by a HV-split (TLR or TRR). It is noticeable that both splits are smaller than the mean of the DTV_LKW cluster with the lowest values (see Table A.5.1.1), but this cluster also accounts for approximately 80 % of the data.

5.1.3 Categorical analysis

The tree structure shows that the leaves can be categorized according to age, HVs share (DTV_LKW), and climate. These categories can be further subdivided as shown in Table 5.1.3. Age can be divided into five subcategories (boxplot nMonths; Appendix A.5.1.4). The heavy vehicles share can be divided into high (everything right of an HV split) and low (everything left of an HV split). The HVs boxplots (= DTV_LKW boxplots; Appendix A.5.1.4) show that the two leaves of the second oldest age group (TRLL, TRLR) also have an unequal distribution of the HV share. It is plausible to consider TRLL as a leaf with rather low HV share and TRLR as a leaf with rather high HV share.

As discussed in Section 5.1.2, the avgFD splits probably also explain climatic conditions. This means that all leaves to the left of an avgFD split are likely to have data that was exposed to relatively warmer climates than those to the right of an avgFD split. The sdHD boxplot (Appendix A.5.1.4) shows that the leaf left of the split has on average more frost days than the right one. The same argument as in Section 5.1.2 applies here as well. The sdHD split probably measures age and climatic conditions. The younger leaf, which is the one with less avgFD, has a right-aligned distribution in the nMonths variable and therefore probably has rather young segments or segments in warmer climates than the opposite leaf with more avgFDs and a left-aligned distribution of the nMonths variable.

The end result of this subdivision is summarized in Table 5.1.3.

Leaf	Age in years	HV-share	Climate
TLL	1.3	-	-
TLRLL	1.6	low	rather warm
TLRLR	2.9	low	rather cold
TLRRL	1.6	high	rather warm
TLRRR	2.9	high	rather cold
TRLL	4.8	rather low	rather cold
TRLR	4.8	rather high	rather warm
TRRL	8.6	low	-
TRRR	8.6	high	-

Table 5.1.3: Overview of categorical classification of LNA regressions

These categories can now be examined for similarities or differences by examining the coefficients of their regressions shown in Appendix A.5.1.3 in order to find general similarities.

Cold climate. The three leaves TLRLR, TLRRR and TRLL belong to the group with rather cold climates. Similarities are that avgFD, DTV_LKW, and sdDirect foster the deterioration of the noise reducing effect of the LNAs and the increasing age (nMonths) inhibits the loss of the noise reducing effect. The harmful effect of frost days and HV share on the noise-reducing effect of an LNA is in agreement with the literature (e.g. frost days: Bühlmann et al. 2015, Licitra et al. 2018; HV: van Blokland et al. 2014). The sdDirect variable is difficult to interpret. It is illogical that nMonths has a positive effect on the noise-reducing effect of a road surface. Unclear as well is the strongly different behavior of the different road surface types.

Warm climate. The three leaves TLRLL, TLRRL and TRLR belong to the group with warmer climates. Similarities are that nMonths damages the noise reducing effect of the LNAs and sdDirect has a beneficial impact on the noise reducing effect of the LNAs. If the leaves of warm climates are additionally subdivided according to HV share (many: TLRRL, TRLR; little: TLRLL), then it appears that in warm climates with a high HV share, frost has a favorable influence on the noise-reducing effect of an LNA and sdHD a harmful one. It is to be expected that with increasing age (nMonths) the noise reducing effect of an LNA decreases. SdDirect is difficult to interpret. It is interesting to note that frost in warm climates with high HV share seems to have a positive effect on noise reduction. This is in contradiction with the literature. SdHD is difficult to interpret. However, a high value of the sdHD variable requires a large number of heat days. This means that the standard deviation will be greater since there are scarcely any heat days in the non-summer months and for those months the values will usually be zero. This indicates that the larger the sdHD variable, the more heat days an LNA will have experienced. This would imply that in warm climates with high traffic loads, frost

is preferable over heat for an LNA. This hypothesis is explored further in Section 5.2. The strongly different behavior of the different road surface types is again unclear.

Low HV share. The four leaves TLRLL, TLRLR, TRLL and TRRL belong to the category with a low HV share. Similarities are that avgFD and DTV_LKW have a harmful impact on the noise reduction effect of an LNA. This is to be expected as already discussed in the previous section "cold climate". It remains unclear why the nMonths variable has a positive influence on the noise reduction effect of LNAs in the two leaves with rather cold climate conditions (TLRLR, TRLL). The strongly different behavior of the different road surface types is again unclear.

High HV share. The four leaves TLRRL TLRRR, TRLR and TRRR belong to the category with a high HV share. The only similarity is that the 4mmHV group is always worse than the 4mmLV reference defined by this model. This is to be expected because under high mechanical stress, road surfaces with a high void content are more fragile.

Age. TLL is the leaf with the highest uncertainty (see boxplot stnl1 in Appendix A.5.1.4 and Section 5.1.1, subgroup discovery) and should therefore not be considered in the interpretation. Considering the other age groups, there are always exactly two leaves for each, except the first age group (TLL). If the residual and the stnl1 (performance) boxplots (Appendix A.5.1.4) are included in the interpretation, the following similarities can be observed: At each age group the types of road surface behave almost exactly the opposite way than the other leaf of the same age group, since the HV share is also always inverted (low vs high). It further shows that those leaves with a higher HV share perform worse than those with a lower HV share: TLRLL < TLRRL, TLRLR < TRRR, TRLL = TRLR (where the traffic volume is almost the same), and TRRL < TRRR.

5.2 The heat hypothesis

The heat hypothesis is as follows: "Heat has a significant damaging effect on the noise reduction effect of a low-noise road surface."

5.2.1 Method

To test the heat hypothesis, the following procedure was applied. All results for the below Steps 4 and 5 can be found in Appendix A.5.2.

1. Two data sets are created. The first data set "low HV" contains segments of the leaves which in Section 5.1 were shown to belong to the category with low HV share (TLRLL, TLRLR, TRLL, TRRL). The second data set "warm, high HV" contains segments of leaves that belong to the category with high HV share and warm climates in Section 5.1 (TLRRL, TRLR).

2. All segments are aggregated to individual road surfaces by an average function.

3. All non-dummy coded, independent variables are standardized by z-transformation.

4. For both data sets, ordinary least square regression models are generated. Both models are subjected to a stepwise variable selection according to Venables and Ripley (2002). Neter et al. (1985) estimated that if the largest variation inflation factor (VIF) among all independent variables exceeds 10, multicollinearity is problematic. Since VIF is always $<= 5$ for all variables (VIF calculation according to Paradis and Schliep 2018), however, multicollinearity is not problematic for those models.

Both models show significant spatial autocorrelation. The possible variables are the same as those selected in variable selection (Section 4.5.1), except that sdHD is replaced by avgHD to simplify interpretability. Since spatial autocorrelation occurs due to unequal spatial distribution (i.e. spatial heteroskedasticity), the spatial error model is a suitable choice to ensure that the p-values of the coefficients are distorted as little as possible.

5. For both data sets a spatial error model (SEM) was developed using the method proposed by Bivand and Piras (2015). The neighborhood relationships are described by a Voronoi diagram.

5.2.2 Results and discussion

Both OLS models (Appendices A.5.2.1 and A.5.2.3) show a significant spatial autocorrelation of approximately 0.33. Therefore, these models are not examined in detail. However, it is important to note that in the warm, high HV data set the variable avgHD in the variable selection part was not selected and therefore has no significant influence on the performance of an LNA.

Both SEM models (Appendices A.5.2.2 and A.5.2.4) have a low spatial autocorrelation of 0.08 which is nevertheless significant. Compared to the OLS models, the fit is also significantly better. In the warm, high HV data set R^2 improves from 32.3 % to 51.3 %, and in the low HV data set from 44.4 % to 59.3 %.

The significant reduction of the spatial autocorrelation and the associated improvement of the fit shows that only the SEM models should be considered when interpreting the coefficients.

Since the avgHD variable was not considered in the variable selection, the warm, high HV data set, the hypothesis "Heat has a significant damaging influence on the noise reduction effect of a low-noise road surface" must be rejected given the experimental setup used in this thesis. It should also be noted that the direction of the avgFD coefficient in the warm, high HV data set is negative but not statistically significant.

This does not suggest that the assumption that heat has a harmful effect on LNAs must generally be rejected. In the low HV SEM the avgFD coefficient is the non-physical parameter coefficient with by far the most harmful effect. It is therefore surprising that avgFD has no significant effect on the warm, high HV data set. It is possible that if an exact control group had been created, e.g. by clustering LNAs that clearly experienced hot climate and high traffic, then variables such as frost or heat might have been statistically significant. Creating such a model was, however, not in the scope of this thesis.

5.3 Addressing the first research question

Which factors lead to the loss of the noise reducing effect of the LNAs? The first research objective of this work, to find out which factors lead to a loss of the noise reducing effect of LNAs, can be partially fulfilled.

Due to the structure of the tree (first splits are due to the variable nMonths), and the decreasing noise reduction effect of the LNAs in older leaves, age is the most important factor leading to a decrease in the noise reduction effect. The second most important factor is HV share. This is evident because for each age group (except the first) there is always a subgroup with a high HV share and a subgroup with a low HV share, whereby the subgroup with a low HV share performs acoustically better than the subgroup with a high HV share. It should be emphasized that between HV share and DTV a Pearson correlation coefficient of 0.7 is observed. HV share therefore probably stands not only for the load of HVs but also for the general direct mechanical stress of a road surface. However, HV share was preferred by the variable selection algorithm over the DTV variable, which indicates that HVs have a more detrimental effect on LNAs than passenger cars. Thus, the observation of van Blokland et al. (2014), namely that HVs are more harmful than passenger cars, is supported by the results of this thesis. Consequently, it can be claimed with certainty that time (nMonths) and mechanical stress (HV share) have a harmful effect on LNAs.

The contribution of environmental factors is less obvious. In general, this work shows that environmental factors have an influence on LNAs. The consideration of environmental factors

in the model led to a strong improvement in performance (Section 4.4). Furthermore, frost has the most damaging effect of all variables on road surfaces exposed to low HV share (Appendix A.5.2.4). These results support the observations of Licitra et al. (2018) and Bühlmann et al. (2015) that frost has a harmful effect on LNAs, while challenging the result of Irali et al. (2015) that temperature has a minor effect on LNAs. However, these results are not unambiguous. In this thesis, neither frost nor heat had a statistically significant effect on road surfaces with a high HV share in warmer climates. This finding is therefore more in line with the result of Irali et al. (2015) that temperature has a minor influence on LNAs. This is also interesting because the sample of Irali et al. (2015) originates from Spain, which suggests that Irali's et al. routes would probably also be classified into the warmer climates group of this data set.

The cause of these contradictory results does not mean that the environmental factors generally have no statistically significant influence. A reason might be that an LMT allocates only datapoints in a certain leaf, respectively regression, that meet certain criteria. The sampling of these regressions is therefore not random but has a sampling bias in this context. Thus, it might occur that with warmer leaves, the proportion of frost is so low that it has no significant influence in this specific regression model.

A further explanation might be that there may have been not enough data in this work. If a clearer division of the samples into hot areas with very high HV share had been made, certain coefficients might be statistically significant under such circumstances.

In conclusion, the results of this work on environmental factors are not precise enough to determine precise threshold values for climatic groupings where, for example, frost is harmful or no longer harmful. However, the results indicate that the relationship between environmental factors and LNAs is not linear. Assuming that the relationship between frost and the performance of an LNA is not linear would be a plausible explanation why frost has a statistically significant detrimental effect in the low HV model and no significant effect in the warm, high HV model.

In order to obtain more clarity, more data is needed from regions that differ strongly in terms of environmental factors. This could lead to more unambiguous groups, whose statistical analysis is likely to then produce more conclusive results.

The results regarding physical parameters are almost completely unclear. The results of the variable selection show that the physical parameters have an influence on the noise reducing effect of an LNA, since all types of LNAs were considered in the variable selection and contributed to a strong performance improvement. However, the behavior of the physical parameters is unexpected. No pattern can be recognized from the categorical analysis. Only 4mmHV performs worse than the reference 4mmLV in regressions with a high HV share. This is to be expected because with high void degree the road surface becomes more fragile to mechanical stress. Due to the subgroup analysis, however, it is also clear that exactly the

4mmHV road surfaces have a systematic error and should therefore not be weighted too heavily in the interpretation.

A reason for the unexpected behavior of the road surface types could be the broad Swiss standardization mentioned in Section 2.2.2. As a result, products with the same label still vary greatly and results with the same label are therefore difficult to interpret.

5.4 CG

As the focus of this work is on LNAs, the analysis and discussion of the CG conventional group will remain short in comparison.

5.4.1 Model evaluation

Performance metrics. The performance metrics for the final CG model are summarized in Table 5.4.1.

Performance metrics	Training Error	Test Error
MAE	0.888	0.851
RMSE	1.199	1.125
MSE	1.438	1.267
R^2	0.265	0.230

Table 5.4.1: Overview of the performance metrics of the final CG model in dB(A)

The MAE, RMSE and MSE suggest that the CG model performs significantly better than the LNA model. However, the R^2 performance metric, which is poor at approximately 25 %, shows that this conclusion is wrong. The reason why MAE, RMSE and MSE are significantly lower than in the LNA model is that the range of possible stnl1 values for the CG road surfaces is significantly smaller than that of the LNAs. Since the measurement error of the target variable (Section 3.1.1) is almost the same size as the MAE of the model, it is virtually impossible to obtain a high R^2 value. This further implies that it is hardly useful to examine non-noise-abating surfaces for their noise reducing effect as they seem to be rather constant.

5.4.2 Tree structure

As with the LNA tree structure, the letters in the leaves stand for the allocation path that a data point has taken. T stands for top, L for left and R for right.
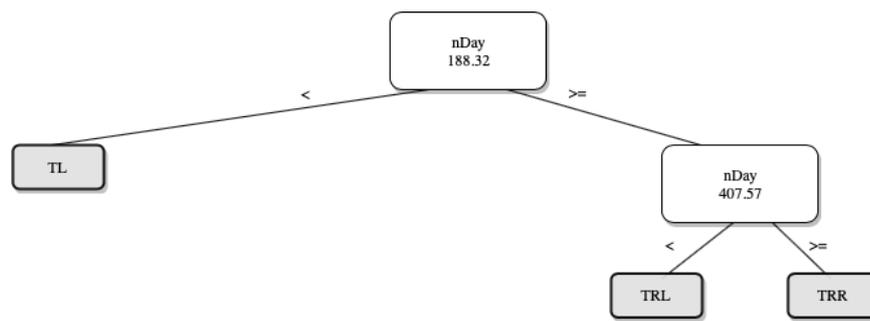
Figure 5.3: Tree structure for the final CG model created by the LMT algorithm

The CG tree structure shows that nDay, i.e. the daily traffic volume, is the most explanatory variable for CG road surfaces. In addition, the nDay variable alone would lead to an MSE of 1,382 (see Table A.4.2.1). Since the R^2 is very low, it is not useful to analyze the variables as precisely as the LNAs, since generally valid statements cannot be made. However, the result leads to the conclusion that for non-LNA road surfaces only the mechanical load of the traffic has a significant impact.

5.5 Addressing the second research question

How can a generalizable model to predict the performance of LNAs be created? The results of this thesis show that it is possible to create a generalizable model for predicting the performance of LNAs. An R^2 of the 5-fold cross-validation of 62.6 % (± 2.4 % in the 95 % Confidence Interval) shows that the model design is robust. By cleaning the outliers of the training and test data the R^2 of the final model can be improved by 10 % to 75-79 %, which is close to the maximum possible R^2 of approximately 87 % (maximum R^2 due to the measurement error of the target variable). This shows that the model yields precise estimates. The results of Section 5.4 also show that it is not advisable to use a model to estimate the acoustic performance of conventional road surfaces, since the range of possible values of the stnl1 variable is only marginally larger than that of the measurement error of the stnl1 variable itself.

When creating a model to estimate the performance of an LNA, consideration must be given to the variable selection, modelling technique and test design.

The discussion in Section 5.3 illustrates that variables that measure age and traffic load are indispensable for an LNA model. Variables representing the physical parameters of an LNA and variables representing the environmental factors should also be considered. Regarding environmental factors, it is recommended to choose climatic variables. Very specific variables that have small-scale effects, such as intersections, bus stops, proximity to construction sites

or farmland, did not score in the feature selection process in this work. This does not mean that Gradziejczyk's (2016) interpretation that proximity to farmland leads to higher dirt input in the pores of LNAs is wrong. Rather, it is difficult to create variables that can determine exactly where the vehicle routes of farmers are located that ultimately lead to dirt entry. Therefore, variables that represent small-area effects should not be used if it is unlikely that the variable represents precisely the desired small-area effect.

The boxplot in Appendix A.5.1.4 of the nMonths variable shows that the age of left-sided regressions increases to right-sided regressions in the tree. The boxplot of the stnl1 variable shows that the LNAs of the youngest age group perform worse than the second youngest, while the second youngest performs better than the third youngest. Hence, there is a nonlinear relationship between the target variable and the nMonths variable. Furthermore, the contradictory behavior of frost depending on the regression (positive effect on LNA for TLL, TLRRL, TRLR and TRRR, negative effect for the rest of the regression) illustrates the nonlinear relationship between the target variable and the avgFD variable.

Therefore, it is indispensable that a technique is chosen that can represent non-linear, non-monotonously rising relationships when generating a generalizable model. Attention should also be paid to the spatial autocorrelation of the data. Since LNAs are mainly installed in urban and suburban areas, there are fewer LNAs in low population areas such as the Alps. Consequently, the samples are spatially unbalanced, resulting in spatial autocorrelation. The results of this thesis, similar to Breuning's (2005) results, show that statistically correct models are not necessarily essential to produce a well performing model. Nevertheless, the results in presented in Section 5.3 clearly show that the consideration of spatial autocorrelation in a model leads to a strong performance improvement and also allows the interpretation of the coefficients of a model. It is therefore desirable to choose a function that takes spatial autocorrelation into account, as this allows the correct interpretation of the model.

In order for the selected model to fit the data well, it is worth excluding road surfaces whose age is less than approximately 1.5 years. The results in Section 5.1.1 show that the youngest leaf has significantly more outliers than the remaining leaves (this can also be seen in the residuals boxplot; Appendix A.5.1.4. It should also be considered to ignore segments that have been installed for less than 1.5 years. This conclusion is based on the fact that in this thesis spatially fine-grained variables such as intersections, bus stops or slope were never selected during the feature selection process. Therefore, it seems that the outliers in the data could be explained rather by the measurement error of the target variable itself than by the environmental factors. Therefore, it might be beneficial to choose a threshold value, for example to ignore all segments that are outside the standard deviation of the whole track.

6 Conclusion

6.1 Achievements

This thesis contributes to the area of research on the aging of LNAs. In this thesis a novel methodological approach was chosen to improve the prediction and understanding of acoustic aging of LNAs. The large database of Grolimund + Partner was used, which contains nearly every LNA in Switzerland at the time of 2018. Using GIS methods, a total of 49 environmental variables were produced and taken into the research questions hypothesis spaces. Through variable selection, the relevant variables for the acoustic aging of the LNAs could be identified. By comparing different machine learning techniques, a suitable model, the LMT, was selected, which allows to provide precise estimates and interpret the relationship between dependent and independent variables. The interpretation of the LMT model enables to understand what to be aware of when creating a generalizable acoustic LNA aging model in terms of obtaining the most accurate estimates and correctly interpreting the relationships between the dependent variable and the independent variable.

6.2 Insights

The first research question, "which factors lead to the loss of the noise reducing effect of the LNAs", can only partially be answered. The results agree with the literature. The most important explanatory variable is age, followed by mechanical stress, which is better described by the average number of heavy vehicles than the average number of vehicles. The results concerning environmental factors, on the other hand, are not conclusive, and further research will be necessary. It is evident from laboratory studies that environmental factors have an impact on the performance of an LNA. Similarly, our results strongly suggest that frost in areas with a low mechanical stress has a strong damaging effect on an LNA. It appears, however, that frost has no harmful effect in warmer areas. The exact reasons why frost has no statistically significant effect in warmer areas cannot be explained by the results obtained from the data available for this thesis. There are indications that there is a climatic range in which frost is harmful or no longer harmful to an LNA.

The second research question, "How can a generalizable model to predict the performance of LNAs be created?", can be answered. An R^2 of 62.6 %± 2.4 % in a 95 % confidence interval of the 5-fold cross-validation shows that the model design is robust. By cleaning of outliers in the training and test data, an improvement of the R^2 of the final model by 10 % to 75-79 % could be gained, which is close to the maximum possible R^2 of approximately 87 % due to the

measurement error of the target variables. This demonstrates that the model makes precise estimates. In order to produce a robust, precise model, variables measuring age, traffic load (preferably heavy vehicles), environmental factors (preferably climatic) and the physical parameters of the LNAs are needed. A non-linear, non-monotonous function should be chosen to describe the relationship between the variables. Since LNAs are mainly located in urban areas, attention should be directed to spatial autocorrelation when sampling the data or selecting the function. This should contribute to an improvement of the model and is indispensable for interpretation. LNAs also show strongly varying noise reduction effects in the first 1.5 years. Thus, it is recommended not to consider data younger than about 1.5 years. In general, the data used in this thesis seem to have several outliers. It is recommended to choose a threshold, for example certain percentile values, and to not consider all segments of a track falling outside these percentiles.

6.3 Limitations

The results and the conclusions that can be drawn from these are limited to the study area, that is, Switzerland. Consequently, the categorical classification of warm/cold climates, or high/low traffic refer to Swiss conditions. The physical parameters of an LNA are also subject to Swiss norms. Furthermore, the data sources used are of varying precision. The traffic load is taken from a data set from the year 2015. The traffic fleet is dynamic and may change significantly locally. Furthermore, the climatic variables avgFD and sdHD do not have a high spatial resolution compared to other variables used in the model.

6.4 Future Work

The results of this thesis show that while environmental factors have an influence on LNAs, they are only partly understood and should thus be researched in more detail. Specifically, it is advisable to investigate in a laboratory set-up how an LNA behaves at certain temperatures and traffic conditions. This would allow to identify more distinct climatic thresholds, which in turn would allow to have relevant variables in the ageing model when constructing or selecting a model.

The thesis also demonstrates that it is possible to create robust and accurate models for predicting the noise reduction effect of an LNA at a given location and provide concrete recommendations from variable selection, choice of mathematical function and data preprocessing. Future work should also look into the application of such prediction models, which is especially important at the political level. Such models would facilitate the decision-

making of measures against road noise in order to achieve the goals of the NAO for communes
and cantons in Switzerland.

7 References

Anselin, L., 2012. New directions in spatial econometrics. Springer, Berlin.

BAFU, B. für U., n.d. Massnahmen gegen Strassenlärm [www Document]. URL https://www.bafu.admin.ch/bafu/de/home/themen/thema-laerm/laerm--fachinformationen/massnahmen-gegen-laerm/massnahmen-gegen-strassenlaerm.html (accessed 31.7.19).

Basu, S., Meckesheimer, M., 2007. Automatic outlier detection for time series: an application to sensor data. Knowledge and Information Systems 11, 137–154.

Bendtsen, H., Lu, Q., Kohler, E., 2010. Acoustic Aging of Acoustic Aging of Asphalt Pavements: Asphalt Pavements: A Californian/Danish A Californian/Danish Comparison Comparison (Reprint Report). California Department of Transportation (Caltrans).

Bennert, T., Hanson, D., Maher, A., Vitillo, N., 2005. Influence of Pavement Surface Type on Tire/Pavement Generated Noise. Journal of Testing and Evaluation 33, 94–100.

Bernhard, R., Wayson, R.L., 2005. An Introduction to Tire/Pavement Noise of Asphalt Pavement. Institute if Safe, Quiet and Durabel Highways, Purdue University.

Bivand, R., Piras, G., 2015. Comparing Implementations of Estimation Methods for Spatial Econometrics. ournal of Statistical Software 63.

Brenning, A., 2005. Spatial prediction models for landslide hazards: review, comparison and evaluation. European Geosciences Union 5, 853–862.

Bühlmann, E., 2019a. Improvements in the CPX method and its ability to predict traffic noise emissions. Inter Noise Madrid.

Bühlmann, E., 2019b. Options for Cities to reduce Noise Levels while promoting good Quality of Life thorugh effective Management of Noise and Spatial planning. ICSV 26 Montréal.

Bühlmann, E., Dias, M., Steiner, S., 2015. Influence of environment- and traffic-related factors on acoustic ageing of low-noise road surfaces in Switzerland. EuroNoise 2015, Maastricht.

Bühlmann, E., Hammer, E., Bueche, N., Perret, J., 2017. Ausführungsbestimmungen für semidichte Asphalte Auswertung phyischer Parameter (No. 1).

Catillaz, A., Fischer, D., 2018. Lärmbelastung in der Schweiz-Ergebnisse des nationalen Lärmmonitorings sonBASE, Stand 2015 (No. 1820), Umwelt-Zustand. BAFU, Bern.

Cracknell, M., Reading, A., 2014. Geological mapping using remote sensing data: A comparison of five machine learning algorithms, their response to variations in the spatial distribution of training data and the use of explicit spatial information. Computer & Geosciences 63, 22–33.

Deng, K., 1998. OMEGA: On-Line Memory-Based General Purpose System Classifier (PhD thesis). Robotics Institute, Carnegie Mellon University, Pittsburgh, PA.

Descornet, G., 2000. Low-Noise Road Surface Techniques And Materials. Inter Noise Nice.

Dillard, L., 2018. The Best of Both Worlds: Linear Model Trees [www Document]. Medium. URL https://medium.com/convoy-tech/the-best-of-both-worlds-linear-model-trees-7c9ce139767d (accessed 25.7.19).

Dawes, J., 2012. Do Data Characteristics Change According to the Number of Scale Points Used? An Experiment Using 5 Point, 7 Point and 10 Point Scales. International Journal of Market Research 50.

Elvik, R., Transportøkonomisk institutt (Norway), 2003. Safety aspects related to low noise road surfaces.

EN/ISO, n.d. ISO 11819-2:2017 [www Document]. ISO. URL http://www.iso.org/cms/render/live/en/sites/isoorg/contents/data/standard/03/96/39675.html (accessed 10.15.19).

Environment Series No. 60, 1987. Computermodell zur Berechnung von Strassenlärm.

ESRI, 2016. ArcGIS Desktop (Version 10.5). Environmental Systems Research Institute, Redlands, CA.

European Asphalt Pavement Association (EAPA), 2007. Abatement of Traffic Noise—The Arguments for Asphalt.

Gahegan, M., 2000. On the Application of Inductive Machine Learning Tools to Geographical Analysis. Geographical Analysis 32, 113–139.

Gardziejczyk, W., 2016. The effect of time on acoustic durability of low noise pavements – The case studies in Poland. Bialystok, Poland, Transportation Research Part D 44, 93–104.

Gauch, H.G., 2003. Scientific method in practice. Cambridge University Press, New York.

Gill, P.Hall.N., 2018. Introduction to Machine Learning Interpretability. O'Reilly Media, Inc., S.l.

Görz, G., Schneeberger, J., Rollinger, C., 2012. Handbuch der Künstlichen Intelligenz, 5., korrigierte Auflage. ed. Oldenbourg, R, München.

Hammer, E., Bühlmann, E., n.d. A guidline for sustainable low-noise pavements. Conference Proceeding ICSV 2017.

Hammer, E., Steiner, S., Dias, M., Bühlmann, E., n.d. Long-term acoustical performance of low noise road surfaces in urban areas in Switzerland. EuroNoise 2015, Maastricht.

Hanson, D., James, R., NeSmith, C., 2004. Tire-Pavement Noise Study. National Center for Asphalt Technology, Auburn, AL, USA.

Harrell, F.E., 2015. Regression modeling strategies: with applications to linear models, logistic and ordinal regression, and survival analysis, Second edition. ed, Springer series in statistics. Springer, Cham Heidelberg New York.

Herrera, F., Carmona, C., Gonzalez, P., Del Jesus, M.J., 2011. An overview on subgroup discovery: Foundations and applications. Knowledge and Information Systems.

Hsu, C.-W., Chang, C.-C., Lin, C.-J., 2003. A Practical Guide to Support Vector Classification 16.

Irali, F., Gonzalel, M., Tighe, S., Simone, A., 2015. Temperature and aging effects on tire/pavement noise generation in ontratian road pavements. In Proceedings of the Transportation Research Board 94th Annual Meeting, Washington, DC.

Ising, H., Kruppa, B., 2004. Health effects caused by noise : Evidence in the literature from the past 25 years. Noise Health 6, 5–13.

ISO/IEC, n.d. ISO 11819-2:2017 [www Document]. ISO. URL http://www.iso.org/cms/render/live/en/sites/isoorg/contents/data/standard/03/96/39675.html (accessed 10.15.19).

Jung, J., Kim, S.-K., Kim, J.Y., Jeong, M.-J., Ryu, C.-M., 2018. Beyond Chemical Triggers: Evidence for Sound-Evoked Physiological Reactions in Plants. Front. Plant Sci. 9.

Kanevski, M., Pozdnoukhov, A., Timonin, V., 2009. Machine learning for spatial environmental data: theory, applications and software, Environmental sciences, environmental engineering. EPFL, Lausanne.

Kohavi, R., 1995. A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection. International Joint Conference on Artifical Intellience (CAI).

Lagemaat, R. van de, 2005. Theory of knowledge for the IB Diploma. Cambridge University Press, Cambridge ; New York.

Lärmliga Schweiz, 2018. Warum ein Klagepool? [www Document]. laermliga.ch. URL https://www.laermliga.ch/1-warum-ein-klagepool.html (accessed 31.7.19).

law of nature, 2005. . Oxford English Dictionary.

Legendre, P., 1993. Spatial Autocorrelation: Trouble or New Paradigm? Ecology 74, 1659–1673.

Licitra, G., Moro, A., Teti, L., Del Pizzo, A., Bianco, F., 2018. Modelling of acoustic ageing of rubberized pavements. Applied Acoustics 146.

Mahesh, P., Surinder, D., 2009. M5 model tree based modelling of reference evapotranspiration. Hydrological Process 23, 1437–1443.

Mierswa, I., & Klinkenberg, R. (2018). RapidMiner Studio (9.2) [Data science, machine learning, predictive analytics]. Retrieved from https://rapidminer.com/, n.d.

Mierswa, I., Klinkenberg, R., 2019. RapidMiner Studio (version 9.2).

NAO - Noise Abatement Ordinance of 15 December 1986, SR 814.41, n.d.

National plan of measures to reduce noise emissions, 2015.

Neter, J., Wasserman, W., Kutner, M., 1985. Applied linear statistical models. Illinois: Richard Irwin, Inc.

Noise Abatement Ordinance, 1986.

Ongel, A., Harvey, J., 2010. Pavement characteristics affecting the frequency content of tire/pavement noise. Institute of Noise Control Engineering. 58.

Ongel, A., Kohler, E., Nelson, J., 2007. Acoustical Absorption of Open-Graded, Gap-Graded, and Dense-Graded Asphalt Pavements (No. UCPRC-RR-2007-12). California Department of Transportation: Sacramento, CA, USA

.

Paradis, E., Schliep, K., 2018. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R, Bioinformatics.

Poole, M., O'Farrell, P., 1971. The Assumptions of the Linear Regression Model. Transactions of the Institute of British Geographers 52, 145–158.

Ronan, C., A., n.d. The Meaning Of Sound. Hart Publishing Co., New York.

Sandberg, U., 1999. Low noise road surfaces - A state-of-the-art review. J. Acoust. Soc. Jpn 20.

Sandberg, U., Descornet, G., 1980. Road surface influence on tire/road noise - Part I. Proc. Inter-Noise 80, Miami.

Sandberg, U., Ejsmont, J.A., 2002. Tyre/road noise: reference book, 1. ed. ed. INFORMEX Ejsmont & Sandberg Handelsbolag, Kisa, Sweden.

Schguanin, G., Ziegler, T., 2006. Leitfaden Strassenlärm. Bundesmat für Umwelt BAFU, Bundesamt für Strassen ASTRA, Bern.

Schweizerischer Verband der Strassen- und Verkehrsfachleute VSS Zürich, 2013. SNR 640 425 Lärmmindernde Decken, Grundlagen.

Sirin, O., 2016. State-of-the-Art Review on Sustainable Design and Construction of Quieter Pavements—Part 2: Factors Affecting Tire-Pavement Noise and Prediction Models. Sustainability 8, 692.

sonBase, 2018. sonBASE 2015 – Potential verschiedener Strassenlärmmassnahmen sound, n.d.

sonBASE GIS noise database [www Document], n.d. URL https://www.bafu.admin.ch/bafu/en/home/topics/noise/state/gis-laermdatenbank-sonbase.html (accessed 1.3.20).

TemaNord, 1996. Road Traffic Noise-Nordic Prediction Method.

Tikhonov, A.N., 1943. On the stability of inverse problems. Doklady Akademii Nauk SSSR 39, 195–198.

van Blokland, G., Tollenaar, C., van Loon, R., 2014. Report on Acoustic Aging of Road Surfaces. CEDR Transnational Road Research Programme.

Venables, W.N., Ripley, B.D., Venables, W.N., 2002. Modern applied statistics with S, 4th ed. ed, Statistics and computing. Springer, New York.

Wayson, R.L., 1998. Relationship between pavement surface texture and highway traffic noise, Synthesis of highway practice. National Academy Press, Washington, D.C.

WHO Regional Office for Europe., 2011. Burden of Disease from Environmental Noise—Quantification of Healthy Life Years Lost in Europe.

A Appendix

A.1 Data description report

Name	Type	Meaning	Source	Group	Variable selection	Unit
SID	Numeric	the unique ID of a measurement on a 20m segment	G+P	ID	IDness	label
PavID	Numeric	the unique ID of a measurement on a road surface	G+P/Const.	ID	IDness	label
MID	Numeric	indicates to which order the measurement belonged	G+P	ID	IDness	label
Coordinate ID	Numeric	the unique ID of a location of a 20m segment	G+P/Const.	Location	IDness	label
Xmid	Numeric	x coordinate between x-start and x-end coordinate	G+P/Const.	Location	IDness	easting in m, CH LV03
Ymid	Numeric	y coordinate between y-start and y-end coordinate	G+P/Const.	Location	IDness	northing in m, CH LV03
Canton	Polynominal	in which canton segment is located	G+P	Location	LNA yes, for CG too stable	label
vDriven	Numeric	speed of the measuring vehicle	G+P	Measurement condition	not of interest	km/h
STNL1	Numeric	target variable, deviation from Stl+86 value, PW tire,	G+P	Target variable	target variable	dB(A)
PavType	Polynominal	type of road surface	G+P/Const.	physical parameters	yes	label
PavYear	Numeric	year of installation	G+P/Const.	Time	yes	yyyy
Datum	Date	date and measuring time of the measuring point	G+P	Time	not of interest	dd.mm.yyyy
nMonths	Numeric	number of months elapsed between installation and measurement	G+P/Const.	Time	yes	#months
vecAngle	Numeric	The 2D orientation of a 20m segment	G+P/Const.	Geometry	yes	degree
deltaAngle	Numeric	curvature	G+P/Const.	Geometry	yes	degree
Busstop	Boolean	if a bus station is at the segment	BAV/Const.	Traffic	yes	boolean
Intersection	Numeric	distance to the nearest intersection	sonBase	Traffic	yes	meter
Spuren	Numeric	number of lanes	sonBase	Traffic	too Stable	label
DTV	Numeric	daily average of vehicles	sonBase	Traffic	yes	#vehicels

nDay	Numeric	daily average of vehicles during daytime	sonBase	Traffic	yes	#vehicels
nEvening	Numeric	daily average of vehicles during evening	sonBase	Traffic	yes	#vehicels
nNight	Numeric	daily average of vehicles during nighttime	sonBase	Traffic	yes	#vehicels
VMaxPv	Numeric	maximal speed limit	sonBase	Traffic	yes	km/h
DTV_LKW	Numeric	daily average of heavy vehicles	sonBase	Traffic	yes	#vehicels
DTV_LKW_Ni	Numeric	daily average of vehicles during nighttime	sonBase	Traffic	yes	#vehicels
DTV_total	Numeric	amount of vehicels rolled over road surface since installation	sonBase/Const.	Traffic	yes	#vehicels
tempTotal	Numeric	sum of average monthly temperatures	Meteo/Const.	Climate	yes	celsius
avgT	Numeric	average monthly temperature	Meteo/Const.	Climate	yes	celsius
sdT	Numeric	standard deviation of monthly temperatures	Meteo/Const.	Climate	yes	celsius
totalPerc	Numeric	sum of precipitation per month	Meteo/Const.	Climate	yes	mm
avgPerc	Numeric	average monthly precipitation	Meteo/Const.	Climate	yes	mm
sdPerc	Numeric	standard deviation of monthly precipitation	Meteo/Const.	Climate	yes	mm
forstdays	Numeric	number of frost days since installation	BFS/Meteo/Const.	Climate	yes	#frostdays
avgFD	Numeric	average monthly frost days	BFS/Meteo/Const.	Climate	yes	#frostdays
sdFD	Numeric	standard deviation of monthly frost days	BFS/Meteo/Const.	Climate	yes	#frostdays
icedays	Numeric	number of ice days since installation	BFS/Meteo/Const.	Climate	yes	#icedays
avgID	Numeric	average monthly ice days	BFS/Meteo/Const.	Climate	yes	#icedays
sdID	Numeric	standard deviation of monthly ice days	BFS/Meteo/Const.	Climate	yes	#icedays
heatdays	Numeric	number of heat days since installation	BFS/Meteo/Const.	Climate	yes	#heatdays
avgHD	Numeric	average monthly heat days	BFS/Meteo/Const.	Climate	yes	#heatdays
sdSD	Numeric	standard deviation of monthly heat days	BFS/Meteo/Const.	Climate	yes	#heatdays
summerdays	Numeric	number of summer days since installation	BFS/Meteo/Const.	Climate	yes	#summerdays
avgSD	Numeric	average monthly summer days	BFS/Meteo/Const.	Climate	yes	#summerdays
sdSD	Numeric	standard deviation of monthly summer days	BFS/Meteo/Const.	Climate	yes	#summerdays
tropnights	Numeric	number of tropical nights since installation	BFS/Meteo/Const.	Climate	yes	#tropnights
avgTN	Numeric	average monthly tropical nights	BFS/Meteo/Const.	Climate	yes	#tropnights

sdTN	Numeric	standard deviation of monthly tropical nights	BFS/Meteo/Const.	Climate	yes	#tropnights
totalDirect	Numeric	amount of direct solar radiation since installation	DSM/Const.	Climate	yes	WH/m²
avgDirect	Numeric	average monthly direct solar radiation	DSM/Const.	Climate	yes	WH/m²
sdDirect	Numeric	standard deviation of monthly direct solar radiation	DSM/Const.	Climate	yes	WH/m²
totalDiffuse	Numeric	amount of diffuse solar radiation since installation	DSM/Const.	Climate	yes	WH/m²
avgDiffuse	Numeric	average monthly diffuse solar radiation	DSM/Const.	Climate	yes	WH/m²
sdDiffuse	Numeric	standard deviation of monthly diffusesolar radiation	DSM/Const.	Climate	yes	WH/m²
totalDurationDIrect	Numeric	amount of time of direct solar radiation since installation	DSM/Const.	Climate	yes	WH/m²
avgDurationDirect	Numeric	average monthly duration of direct solar radiation	DSM/Const.	Climate	yes	WH/m²
sdDurationDirect	Numeric	standard deviation of monthly duration of direct solar radiation	DSM/Const.	Climate	yes	WH/m²
distnaceAgrar	Numeric	distance to the nearest farmland	Area statistics/Const.	Dirt infill	yes	meter
distanceConstruction	Numeric	distance to the nearest construction site	Area statistics/Const.	Dirt infill	yes	meter
slope	Numeric	slope inclination	Alti3D/Const.	Topographie	yes	degree
aspect	Numeric	aspect	Alti3D/Const.	Topographie	yes	degree
tri	Numeric	terrain roughness index	Alti3D/Const.	Topographie	yes	%
Zmid	Numeric	height	Alti3D/Const.	Topographie	yes	m.a.s.l

*Const. means that the author constructed/transformed the original data in order to create the variable.

A.2 Data quality report

A.2.1 LNA continuous variables

Name	N	Mean	StDev	Min	Pctl 25	Pctl 75	Max	% Miss
PavYear	70'071	2'010.70	3.98	1'980.00	2'009.00	2'013.00	2'018.00	0.00
Xmid	70'071	609'683.60	62'116.61	488'305.80	572'821.90	654'734.20	789'376.60	0.00
Ymid	70'071	204'217.70	56'493.18	96'317.30	144'441.30	250'860.20	279'399.20	0.00
STNL1	70'071	-2.26	3.02	-10.96	-4.57	0.14	10.05	0.00
vDriven	70'071	49.70	1.59	2.94	49.16	50.56	60.58	0.00
vecAngle	70'071	182.35	104.22	0.03	87.45	271.26	360.00	0.00
deltaAngle	70'071	1.03	31.80	-359.97	-0.44	0.49	359.89	0.00
nMonths	70'071	52.84	42.29	3.00	22.00	72.00	451.00	0.00
tempTotal	70'071	375.70	294.91	-48.38	98.07	573.72	1'499.52	0.00
avgT	70'071	7.99	3.76	-16.13	6.70	9.29	25.80	0.00
sdT	70'071	7.62	2.25	0.00	6.93	7.81	27.79	0.00
totalPerc	70'071	3'684.66	3'850.68	27.52	1'033.74	5'211.47	43'468.28	0.00
avgPerc	70'071	80.86	33.30	13.76	62.74	88.85	325.26	0.00
sdPerc	70'071	49.62	23.84	0.00	34.98	58.44	386.30	0.00
frostdays	70'071	265.33	257.56	0.00	88.00	381.00	2'602.00	0.00
avgFD	70'071	3.68	1.82	0.00	2.89	5.10	6.42	0.00
sdFD	70'071	46.34	33.54	0.00	0.00	73.13	95.19	0.00
icedays	70'071	126.41	116.85	0.00	49.00	183.00	1'267.00	0.00
avgID	70'071	1.78	0.91	0.00	1.33	2.43	3.49	0.00
sdID	70'071	23.45	17.83	0.00	0.00	37.31	57.50	0.00
heatdays	70'071	15.86	16.69	0.00	3.00	26.00	57.00	0.00
avgHD	70'071	0.48	0.72	0.00	0.04	0.62	3.80	0.00
sdHD	70'071	12.22	15.36	0.00	0.00	15.81	56.49	0.00
summerdays	70'071	26.15	22.48	0.00	8.00	42.00	92.00	0.00
avgSD	70'071	0.79	1.04	0.00	0.10	1.00	5.75	0.00

sdSD	70'071	20.05	21.23	0.00	0.00	33.13	88.71	0.00
tropnights	70'071	24.96	19.02	0.00	6.00	48.00	57.00	0.00
avgTN	70'071	0.76	0.93	0.00	0.09	0.98	3.60	0.00
sdTN	70'071	19.00	18.40	0.00	0.00	31.63	56.62	0.00
totalDirect	70'071	6'730'141.00	5'678'669.00	0.00	2'995'509.00	9'159'654.00	65'229'748.00	0.00
avgDirect	70'071	124'470.30	25'251.19	0.00	113'251.00	138'709.80	198'606.90	0.00
sdDirect	70'071	122'875.20	70'991.99	0.00	75'094.55	201'322.90	321'064.10	0.00
totalDiffuse	70'071	1'056'081.00	861'396.50	0.00	439'660.80	1'469'904.00	9'802'992.00	0.00
avgDiffuse	70'071	14'492.21	5'648.15	0.00	11'052.54	19'545.49	34'258.63	0.00
sdDiffuse	70'071	19'907.55	4'133.23	0.00	18'430.56	22'094.77	34'258.63	0.00
totalDurationDirect	70'071	13'348.32	11'265.78	0.00	5'452.52	18'568.26	131'745.60	0.00
avgDurationDirect	70'071	251.98	69.13	0.00	216.84	297.99	520.55	0.00
sdDurationDirect	70'071	150.08	98.59	0.00	75.42	233.54	520.55	0.00
slope	70'071	0.51	4.59	-20.00	-1.86	2.52	19.99	0.00
aspect	70'071	187.15	104.42	0.00	96.96	278.19	359.99	0.00
tri	70'071	0.49	0.14	0.00	0.40	0.58	0.95	0.00
Zmid	70'071	463.50	129.31	255.84	401.87	485.12	1'809.90	0.00
distanceAgrar	70'071	349.19	1'170.35	1.50	88.90	298.50	19'178.99	0.00
distanceConstruction	70'071	839.72	629.15	8.75	378.60	1'136.98	3'898.10	0.00
Intersection	70'071	85.95	92.43	0.31	28.38	109.16	1'109.07	0.00
Spuren	70'071	1.87	0.46	0.00	2.00	2.00	2.00	0.00
Steigung	70'071	0.02	2.48	-10.00	-0.90	1.10	10.00	0.00
DTV	70'071	7'019.82	6'601.98	7.00	2'560.00	10'000.00	112'738.00	0.00
nDay	70'071	360.49	364.91	0.00	127.29	497.20	6'464.82	0.00
nEvening	70'071	360.49	364.91	0.00	127.29	497.20	6'464.82	0.00
nNight	70'071	60.16	64.07	0.00	19.16	84.21	1'162.61	0.00
VMaxPv	70'071	56.96	19.66	0.00	50.00	80.00	120.00	0.00
DTV_LKW	70'071	157.93	270.28	0.00	18.00	175.20	5'430.00	0.00
DTV_LKW_Ni	70'071	10.85	19.63	0.00	1.00	12.00	419.00	0.00
DTV_total	70'071	11'513'352.00	17'104'287.00	2'520.00	2'460'000.00	14'173'650.00	261'990'000.00	0.00

A.2.2 LNA categorical variables

Name	N	Cardinality	Mode	Mode Freq.	Mode %	% Miss
Canton	70'071	18	Aargau	28'433	41.48	0.00
Pavtype	70'071	5	8mmLV	33'787	48.21	0.00
Busstop	70'071	2	"No"	67'666	1.30	0.00

A.2.3 CG continuous variables

Name	N	Mean	StDev	Min	Pctl 25	Pctl 75	Max	% Miss
PavYear	170'083	1'996.15	21.92	1'969.00	1'989.00	2'004.00	2'017.00	0.00
Xmid	170'083	655'517.60	17'901.28	580'997.20	645'979.60	663'749.50	788'673.50	0.00
Ymid	170'083	249'163.10	15'853.23	113'942.70	239'921.40	260'594.30	274'028.80	0.00
STNL1	170'083	1.08	1.34	-9.19	0.25	1.94	18.69	0.00
vDriven	170'083	50.03	1.45	45.00	49.34	50.84	55.00	0.00
vecAngle	170'083	188.29	104.50	0.03	100.72	284.09	360.00	0.00
deltaAngle	170'083	0.04	29.05	-359.86	-0.86	0.76	359.97	0.00
nMonths	170'083	205.45	116.85	3.00	113.00	294.00	487.00	0.00
tempTotal	170'083	707.25	364.84	-19.08	417.99	1'023.46	1'561.26	0.00
avgT	170'083	4.28	3.30	-15.29	2.44	5.96	23.36	0.00
sdT	170'083	9.11	2.08	0.00	7.85	10.03	26.48	0.00
totalPerc	170'083	18'287.07	11'029.18	46.10	9'450.85	26'883.75	49'152.13	0.00
avgPerc	170'083	90.96	18.69	0-00	37.88	55.15	325.46	0.00
sdPerc	170'083	47.13	18.66	0.00	37.80	55.16	325.46	0.00
frostdays	170'083	1'183.32	676.35	0.00	647.00	1'714.00	2'843.00	0.00
avgFD	170'083	5.35	1.04	0.00	5.37	5.75	6.56	0.00
sdFD	170'083	60.71	24.75	0.00	67.20	71.60	95.00	0.00
icedays	170'083	523.29	310.64	0.00	282.00	761.00	1'546.00	0.00
avgID	170'083	2.35	0.53	0.00	2.19	2.59	4.46	0.00
sdID	170'083	29.01	12.37	0.00	28.90	35.70	64.00	0.00
heatdays	170'083	9.48	11.66	0.00	3.00	10.00	57.00	0.00
avgHD	170'083	0.09	0.23	0.00	0.01	0.10	4.00	0.00
sdHD	170'083	9.16	11.43	0.00	3.00	9.90	56.00	0.00
summerdays	170'083	16.93	13.33	0.00	7.00	26.00	92.00	0.00
avgSD	170'083	0.16	0.34	0.00	0.03	0.10	5.00	0.00
sdSD	170'083	16.33	13.04	0.00	7.00	24.90	91.00	0.00
tropnights	170'083	15.79	12.19	0.00	6.00	24.00	57.00	0.00
avgTN	170'083	0.15	0.32	0.00	0.03	0.10	4.00	0.00

sdTN	170'083	15.21	11.95	0.00	6.00	23.90	56.00	0.00
totalDirect	170'083	27'407'832.00	15'973'151.00	0.00	14'581'537.00	39'640'130.00	73'805'243.00	0.00
avgDirect	170'083	132'595.50	15'325.60	0.00	130'952.40	140'316.50	191'732.90	0.00
sdDirect	170'083	78'822.02	11'565.67	0.00	77'005.23	81'722.59	314'205.90	0.00
totalDiffuse	170'083	4'093'281.00	2'470'746.00	0.00	2'100'729.00	5'983'530.00	10'929'949.00	0.00
avgDiffuse	170'083	10'852.91	2'308.39	0.00	10'366.12	12'172.79	31'459.83	0.00
sdDiffuse	170'083	20'064.79	4'047.56	0.00	19'054.40	22'295.49	36'986.61	0.00
totalDurationDirect	170'083	53'235.18	34'214.73	0.00	25'773.76	77'877.59	163'142.60	0.00
avgDurationDirect	170'083	259.73	72.86	0.00	224.30	314.01	516.66	0.00
sdDurationDirect	170'083	83.57	31.36	0.00	67.71	95.20	516.66	0.00
slope	170'083	0.69	5.43	-20.00	-2.33	3.29	20.00	0.00
aspect	170'083	175.21	151.31	0.00	77.77	271.61	359.98	0.00
tri	170'083	0.49	0.13	0.00	0.42	0.57	0.94	0.00
Zmid	170'083	438.22	155.69	0.00	380.81	477.68	1'765.60	0.00
distanceAgrar	170'083	266.00	1'145.26	1.74	59.00	214.00	18'297.03	0.00
distanceConstruction	170'083	951.79	642.70	4.72	448.93	1'331.93	3'726.05	0.00
Intersection	170'083	173.62	198.84	0.18	39.86	231.27	1'369.84	0.00
Spuren	170'083	1.88	0.45	0.00	2.00	2.00	2.00	0.00
Steigung	170'083	0.09	3.24	-10.00	-1.20	1.40	10.00	0.00
DTV	170'083	6'841.24	5'801.98	1.00	2'550	9'850	65'623	0.00
nDay	170'083	334.12	301.07	0.00	112.65	502.55	7'038.63	0.00
nEvening	170'083	334.12	301.07	0.00	112.65	502.55	7'038.63	0.00
nNight	170'083	55.60	53.22	0.00	17.26	85.36	1'224.11	0.00
VMaxPv	170'083	66.12	20.91	0.00	50.00	80.00	120.00	0.00
DTV_LKW	170'083	180.45	256.94	0.00	17.00	263.00	5'751.00	0.00
DTV_LKW_Ni	170'083	12.30	18.28	0.00	1.00	18.00	455.00	0.00
DTV_total	170'083	42'343'194.00	48'845'564.00	5'160.00	10'910'340.00	56'151'000.00	529'577'610.00	0.00

A.2.4 CG categorical variables

Name	N	Cardinality	Mode	Mode Freq.	Mode %	% Miss
Canton	170'083	15	Aargau	162'506	95.54	0.00
Pavtype	170'083	6	AC	75'726	44.52	0.00
Busstop	170'083	2	"No"	168'906	99.99	0.00

A.3 Spatial distribution

A.3.1 LNA spatial distribution
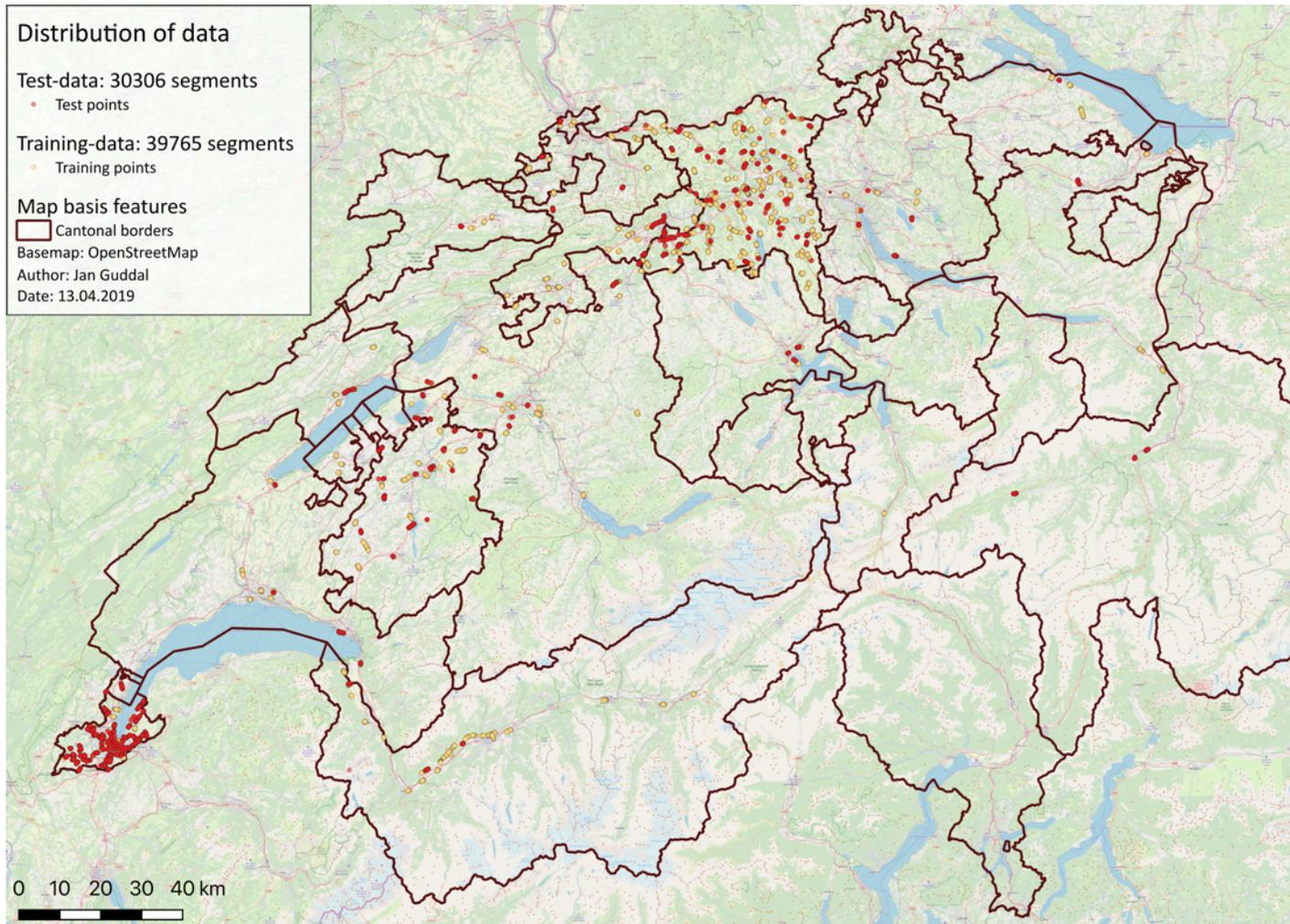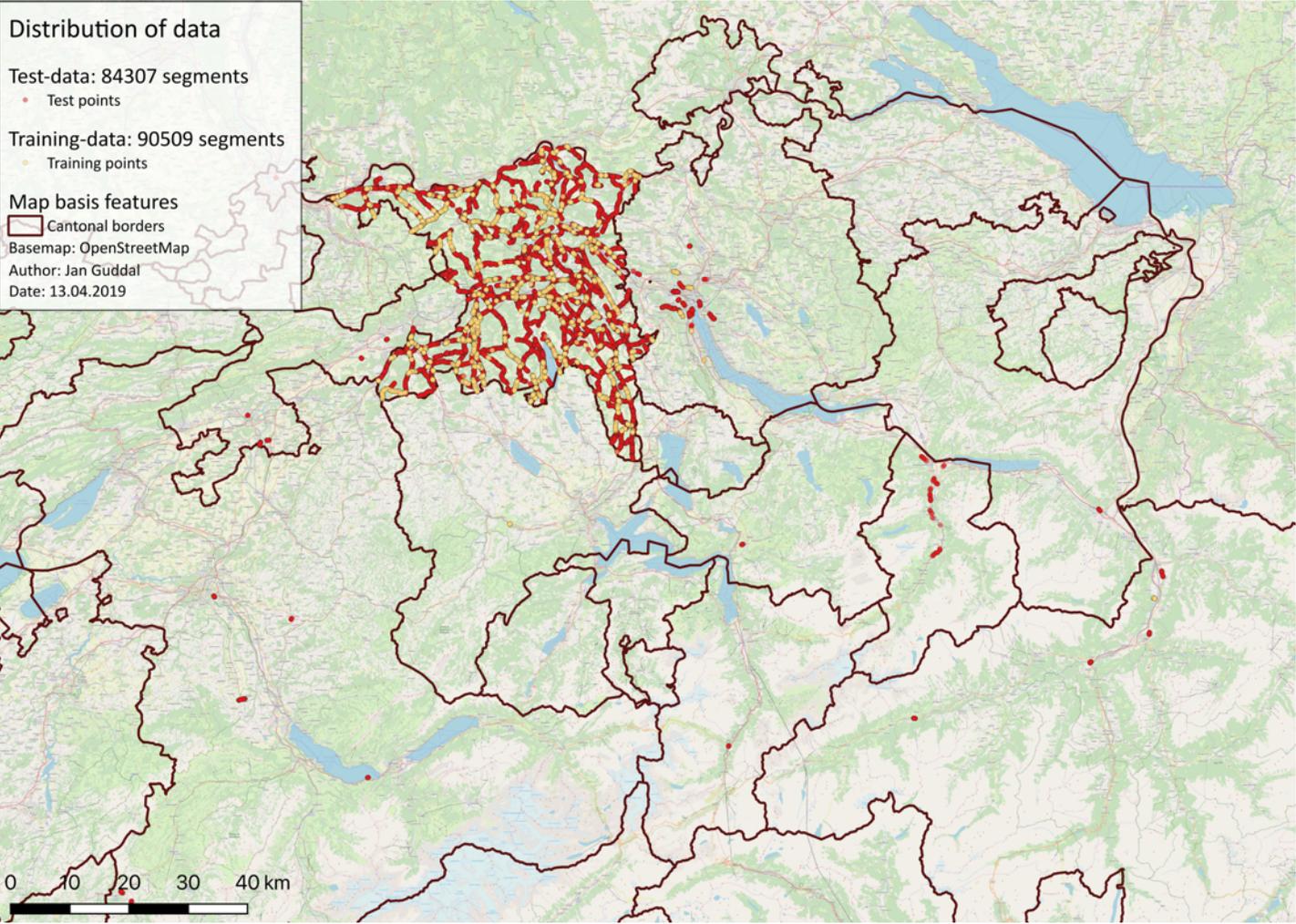
A.4 Variable selection results

A.4.1 LNA variable selection

A.4.1.1 Stepwise forward selection

Name	MSE	MSE improvement
8mmLV	6.118	...
nMonths	4.320	1.798
4mmHV	4.003	0.317
4mmMV	3.793	0.210
sdDirect	3.605	0.188
8mmMV	3.507	0.098
DTV_LKW	3.222	0.284
avgFD	3.156	0.067
ELBOWPOINT: sdHD	3.067	0.088
avgDiffuse	3.036	0.032
nNight	3.058	-0.022
sdSD	3.037	0.021
deltaAngle	3.033	0.004
avgDurationDirect	3.026	0.006
aspect	3.024	0.002
Busstop	3.024	0.000
slope	3.025	-0.001
vecAngle	3.027	-0.002
sdDurationDirect	3.047	-0.020
tri	3.068	-0.021
nDay	3.084	-0.016
nEvening	3.083	0.000
DTV	3.083	0.000
avgSD	3.096	-0.013
VMaxSv	3.061	0.035
...	...	...
all variables	4.992	...

A.4.1.2 Submodel comparison

Set of Variables	MAE training	MSE training	MAE test	MSE test	MAE difference	MSE difference
nMonths, 4mmMV, 4mmHV, 8mmLV, 8mmMV, sdDirect, DTV_LKW, avgFD, avgDiffuse	1.2884	2.8987	1.3041	3.1099	-0.0157	-0.2112
nMonths, 4mmMV, 4mmHV, 8mmLV, 8mmMV, sdDirect, DTV_LKW, avgFD, sdHD	1.2813	2.8798	1.3044	3.1060	-0.0231	-0.2262
nMonths, 4mmMV, 4mmHV, 8mmLV, 8mmMV, sdDirect, DTV_LKW, avgFD, sdHD, avgDiffuse	1.2755	2.8525	1.3023	3.1059	-0.0268	-0.2535
nMonths, 4mmMV, 4mmHV, 8mmLV, 8mmMV, sdDirect, DTV_LKW, avgFD, sdHD, nNight	1.2661	2.8243	1.2997	3.0852	-0.0336	-0.2610
nMonths, 4mmLV, 4mmMV, 4mmHV, 8mmLV, 8mmMV, DTV_LKW, sdDirect, avgFD, avgHD, DTV_LKW_Ni	1.2579	2.8015	1.2935	3.0778	-0.0356	-0.2764
nMonths, 4mmMV, 4mmHV, 8mmLV, 8mmMV, sdDirect, DTV_LKW	1.2994	2.9580	1.3330	3.2413	-0.0336	-0.2833
nMonths, 4mmLV, 4mmMV, 4mmHV, 8mmLV, 8mmMV, DTV_LKW, sdDirect, avgFD, avgHD	1.2714	2.8435	1.3088	3.1345	-0.0374	-0.2910
nMonths, 4mmLV, 4mmMV, 4mmHV, 8mmLV, 8mmMV, avgDirect, DTV_LKW_Ni, avgTN, DTV_LKW, sdDiffuse	1.2613	2.8449	1.3214	3.1817	-0.0601	-0.3368
nMonths, 4mmLV, 4mmMV, 4mmHV, 8mmLV, 8mmMV, avgDirect, DTV_LKW_Ni, avgTN, DTV_LKW, sdDiffuse	1.2613	2.8449	1.3214	3.1817	-0.0601	-0.3368

A.4.1.3 5-fold cross-validation comparison

Set of Variables	MAE train.	MSE train.	MAE test	MSE test	R² train.	R² test	MAE diff.	MSE diff.	R² diff.
nMonths, 4mmMV, 4mmHV, 8mmLV,8mmMV, sdDirect, DTV_LKW, avgFD,avgDiffuse	1.29	2.92	1.36	3.38	0.67	0.62	-0.07	-0.46	-0.05
nMonths, 4mmMV, 4mmHV, 8mmLV, 8mmMV, sdDirect, DTV_LKW, avgFD,sdHD	1.28	2.92	1.35	3.31	0.67	0.63	-0.07	-0.39	-0.04

A.4.2 CG variable selection

A.4.2.1 Stepwise forward selection

Name	MSE	MSE improvement
nDay	1.382	
VMaxSv	1.351	0.031
avgDiffuse	1.319	0.033
DSAK	1.302	0.017
avgDirect	1.294	0.008
avgTN	1.289	0.005
ELBOWPOINT: sdT	1.268	0.021
avgHD	1.264	0.004
SMA	1.262	0.002
DTV	1.260	0.001
vecAngle	1.259	0.001
slope	1.259	0.000
Busstop	1.259	0.000
nAbend	1.259	0.000
deltaAngle	1.259	0.000
avgSD	1.259	0.000
MA	1.261	-0.002
aspect	1.263	-0.002
avgDurationDirect	1.267	-0.003
sdHD	1.271	-0.004
Beton	1.276	-0.005
sdFD	1.275	0.001
sdTN	1.272	0.004
tri	1.272	0.000
AC	1.278	-0.006
...	...	...
all variables	1.885	...

A.4.2.2 Submodel comparison

Set of Variables	MAE training	MSE training	MAE test	MSE test	MAE difference	MSE difference
nTag, VMaxSv, avgDiffuse, DSAK, avgDirect, avgTN, sdT, avgHD, SMA, DTV, vecAngle, slope, Busstop	0.8809	1.4236	0.8458	1.2586	0.0351	0.1650
nTag, VMaxSv, avgDiffuse, DSAK, avgDirect, avgTN, sdT, avgHD, SMA, DTV, vecAngle, slope	0.8809	1.4236	0.8458	1.2587	0.0351	0.1649
nTag, VMaxSv, avgDiffuse, DSAK, avgDirect, avgTN, sdT, avgHD, SMA, DTV, vecAngle	0.8813	1.4252	0.8459	1.2590	0.0354	0.1662
nTag, VMaxSv, avgDiffuse, DSAK, avgDirect, avgTN, sdT, avgHD, SMA, DTV	0.8819	1.4263	0.8467	1.2602	0.0352	0.1661
nTag, VMaxSv, avgDiffuse, DSAK, avgDirect, avgTN, sdT, avgHD, SMA	0.8846	1.4324	0.8483	1.2615	0.0363	0.1709
nTag, VMaxSv, avgDiffuse, DSAK, avgDirect, avgTN, sdT, avgHD	0.8864	1.4348	0.8499	1.2639	0.0364	0.1709
nTag, VMaxSv, avgDiffuse, DSAK, avgDirect, avgTN, sdT	0.8887	1.4385	0.8511	1.2682	0.0376	0.1702
nTag, VMaxSv, avgDiffuse, DSAK, avgDirect, avgTN	0.8968	1.4696	0.8561	1.2890	0.0407	0.1805
nTag, VMaxSv, avgDiffuse, DSAK, avgDirect	0.9177	1.5789	0.8545	1.2937	0.0632	0.2851
nTag, VMaxSv, avgDiffuse, DSAK	0.9172	1.5914	0.8562	1.3020	0.0610	0.2894
nTag, VMaxSv, avgDiffuse	0.9405	1.6338	0.8693	1.3185	0.0712	0.3153
nTag, VMaxSv	0.9511	1.6529	0.8823	1.3515	0.0689	0.3014
nTag	0.9719	1.7236	0.8927	1.3822	0.0792	0.3414

A.5 Linear model tree analysis

A.5.1 LNA

A.5.1.1 k-Means clusters

Variable	--	-	=	+	++
avgFD	0.047	2.505	3.487	4.445	5.502
avgFD r.s.	0.151	0.110	0.193	0.271	0.274
DTV_LKW	64.697	454.574	1066.387	2491.577	5430.000
DTV_LKW r.s.	0.811	0.147	0.039	0.002	0.000
nMonths	19.769	53.786	95.811	244.754	376.114
nMonths r.s.	0.402	0.348	0.240	0.005	0.005
predError	-3.084	-1.203	0.169	1.530	4.008
predError r.s.	0.102	0.260	0.378	0.228	0.052
sdDirecht	19464.750	77153.323	160296.447	209469.182	243780.044
sdDirecht r.s.	0.032	0.625	0.054	0.139	0.149
sdHeatdays	1.916	12.497	25.534	34.727	50.851
sdHeatdays r.s.	0.563	0.198	0.104	0.050	0.085

A.5.1.2 Subgroup discovery models

ClusterPred --	WRAcc
Kanton=Geneva --> ClusterPred = --=true	0.012
Kanton=Geneva PavType2=Nanosoft4 --> ClusterPred = --=true	0.012
ClusterNMonths=-- RegressionLines=TLL --> ClusterPred = --=true	0.010
RegressionLines=TLL --> ClusterPred = --=true	0.010
PavType2=Nanosoft4 --> ClusterPred = --=true	0.010

ClusterPred -	WRAcc
ClusterDTV_LKW=-- ClusterNMonths=-- --> ClusterPred = =true	0.018
ClusterNMonths=-- --> ClusterPred = =true	0.016
ClusterNMonths=-- ClusterSDDirect=- --> ClusterPred = =true	0.013
ClusterDTV_LKW=-- ClusterNMonths=-- ClusterSDDirect=- --> ClusterPred = =true	0.013
ClusterDTV_LKW=-- ClusterNMonths=-- ClusterSDHD=-- --> ClusterPred = =true	0.010

ClusterPred =	WRAcc
PavType2=ACMR8 ClusterSDDirect=- --> ClusterPred = ==true [Pos=11916.0	0.032
PavType2=ACMR8 --> ClusterPred = ==true [Pos=12358.0	0.031
ClusterNMonths== --> ClusterPred = ==true [Pos=7796.0	0.031
PavType2=ACMR8 ClusterAVGFD=++ --> ClusterPred = ==true [Pos=6994.0	0.030
PavType2=ACMR8 ClusterAVGFD=++ ClusterSDDirect=- --> ClusterPred = ==true	0.030

ClusterPred +	WRAcc
ClusterNMonths== RegressionLines=TRRL --> ClusterPred = +=true	0.007
ClusterDTV_LKW=-- ClusterNMonths== RegressionLines=TRRL --> ClusterPred = +=true	0.007
RegressionLines=TRRL --> ClusterPred = +=true	0.007
ClusterDTV_LKW=-- RegressionLines=TRRL --> ClusterPred = +=true	0.007
PavType2=ACMR8 ClusterSDDirect=- --> ClusterPred = +=true	0.006

ClusterPred ++	WRAcc
ClusterNMonths=-- --> ClusterPred = ++=true	0.010
ClusterAVGFD=-- ClusterSDHD=-- --> ClusterPred = ++=true	0.010
ClusterAVGFD=-- --> ClusterPred = ++=true	0.009
ClusterNMonths=-- ClusterSDHD=-- --> ClusterPred = ++=true	0.009
ClusterAVGFD=-- ClusterNMonths=-- ClusterSDHD=-- --> ClusterPred = ++=true	0.009

A.5.1.3 Coefficients of LNA ridge-regressions
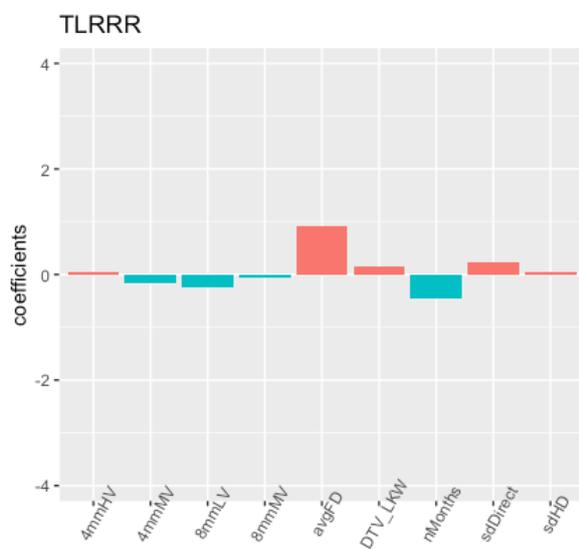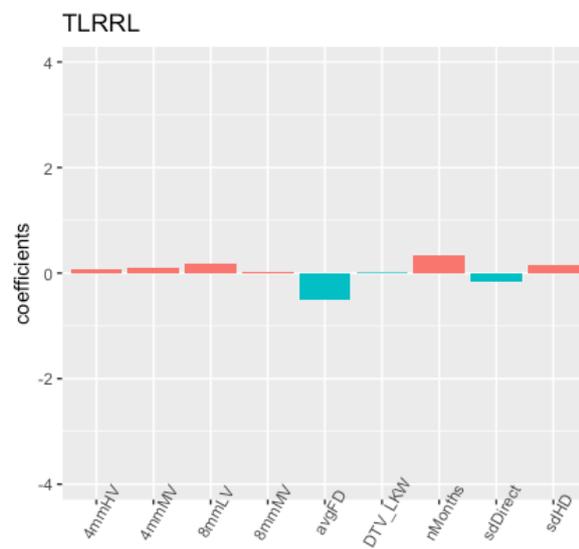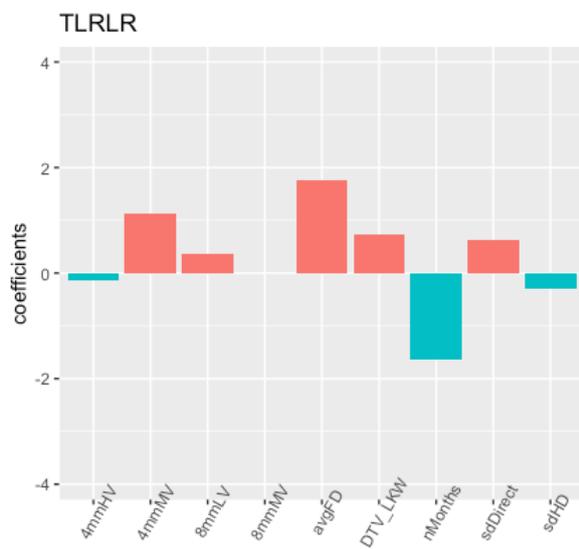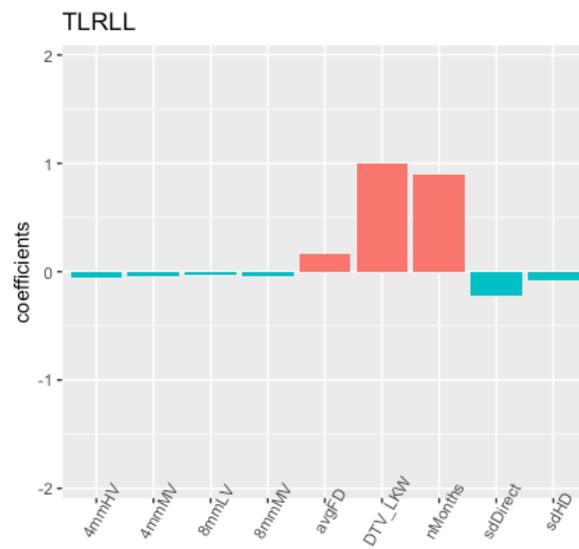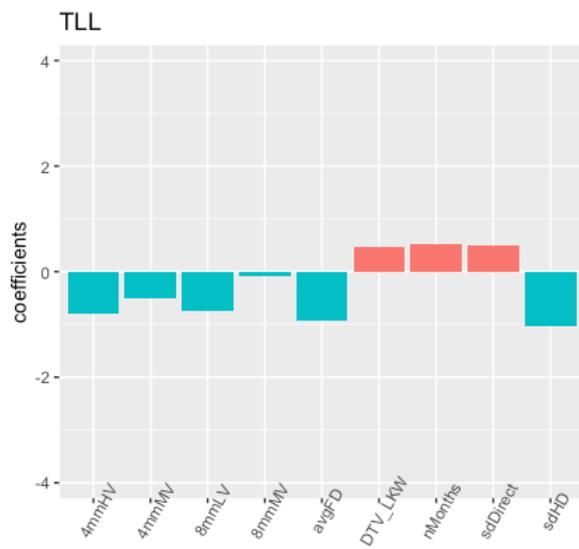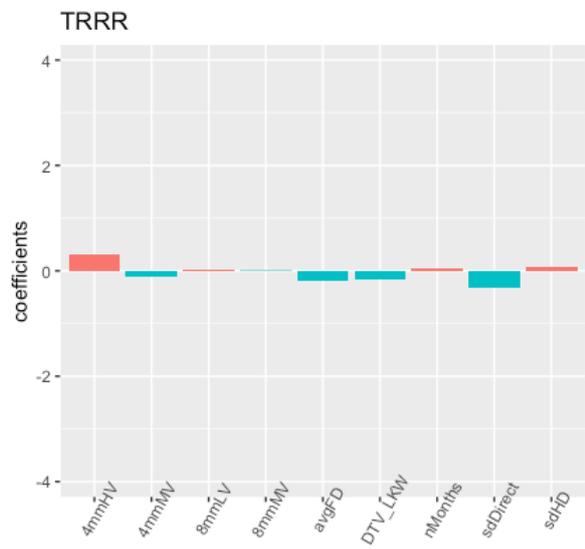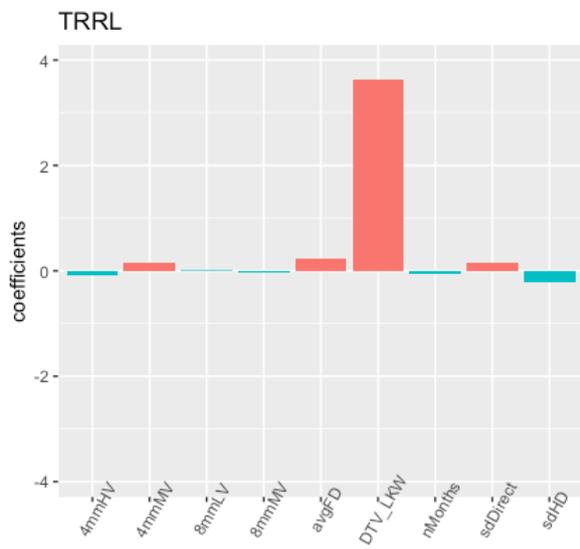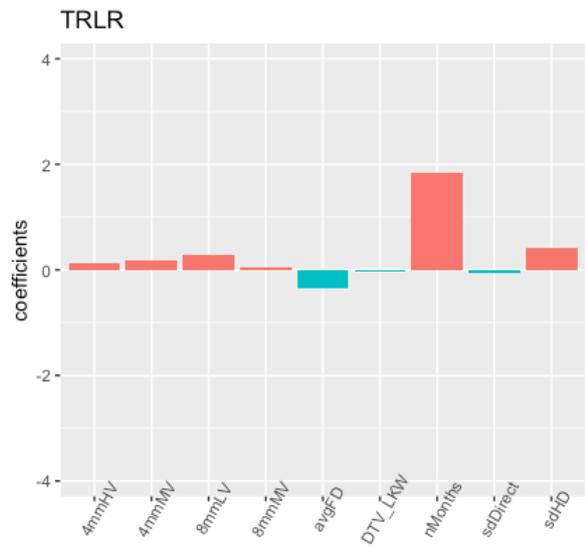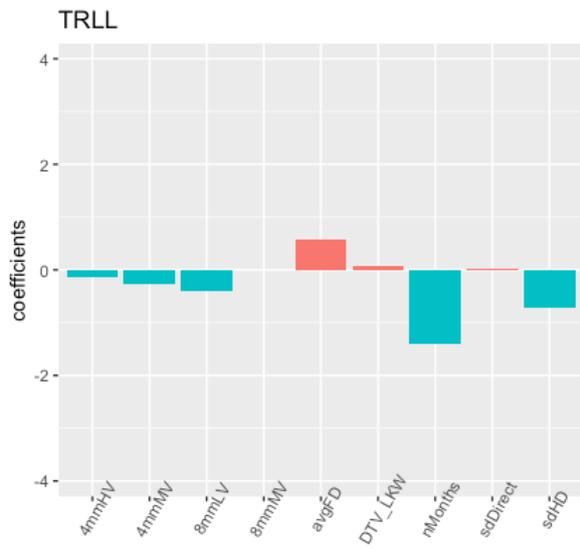
A.5.1.4 Boxplots of LNA variables

boxplots of avgFD variable
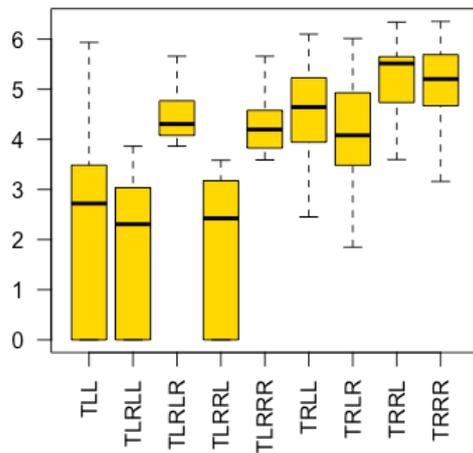
boxplots of DTV_LKW variable
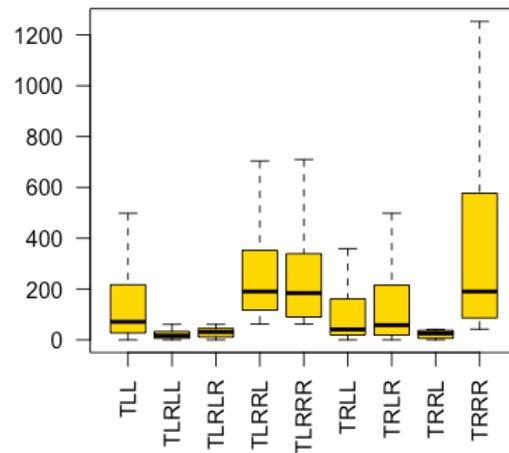
boxplots of nMonths variable
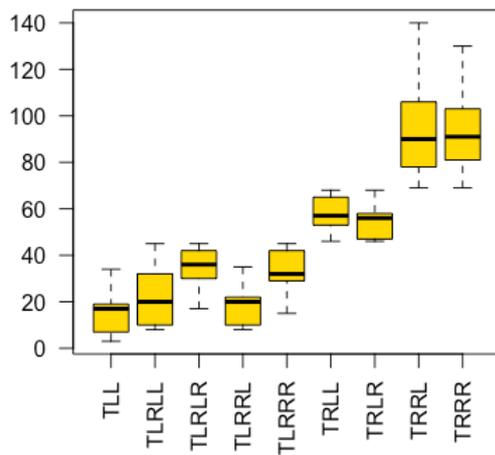
boxplots of sdDirect variable
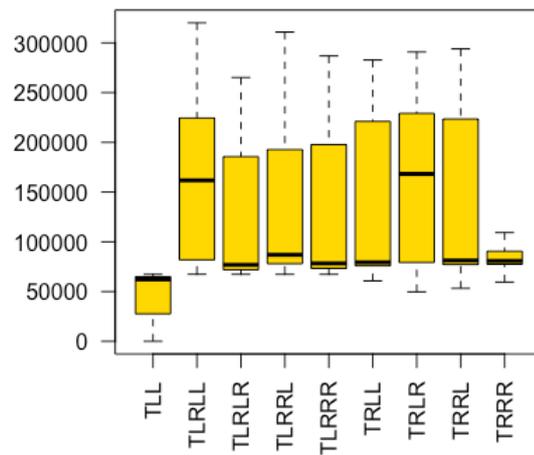
boxplots of sdHD variable
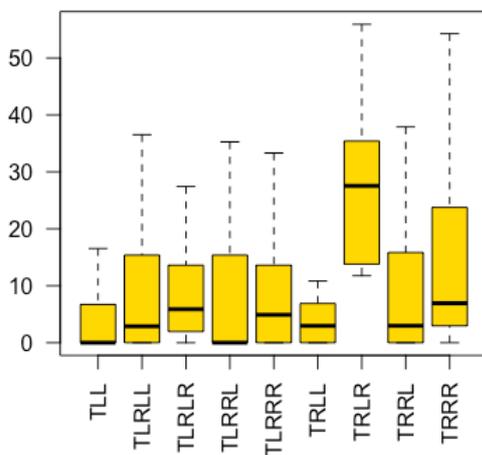
boxplots of stnl1 variable
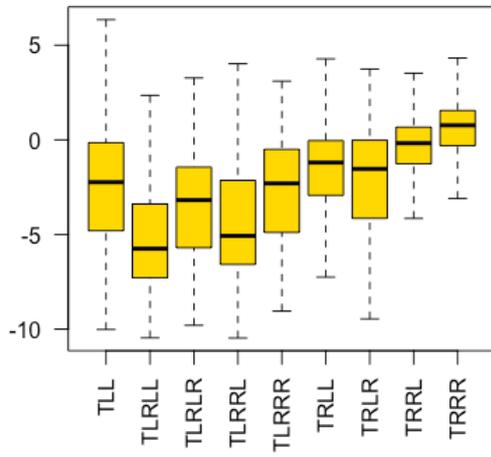
boxplots of predicted value variable
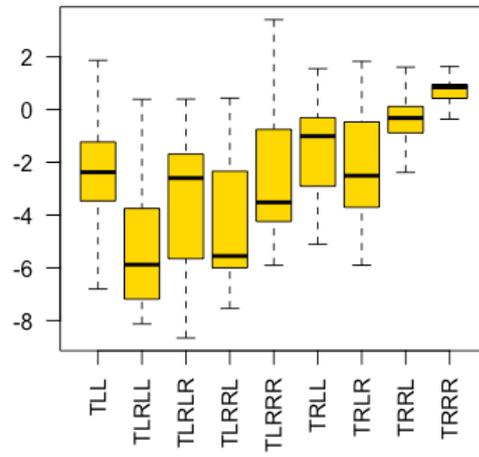
boxplots of residuals variable
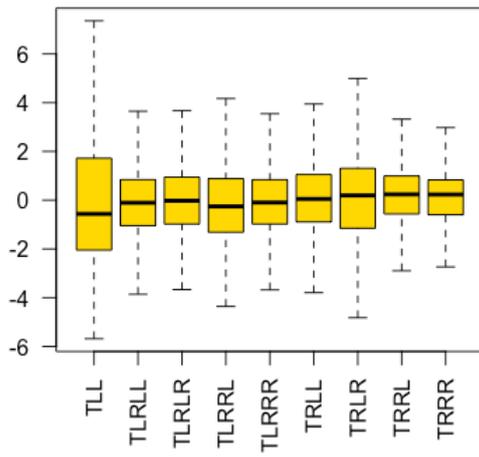
A.5.2 Heat hyptohesis

A.5.2.1 warm, high HV OLS regression model

Coefficients	Estimate	Std. Error	t-value	Pr(>|t|)	VIF
(Intercept)	-2.934	0.126	-23.228	0.000	
avgFD	-0.369	0.246	-1.497	0.135	5.327
nMonths*	1.339	0.249	5.389	0.000	5.199
DTV_LKW*	0.440	0.114	3.852	0.000	1.072
sdDirect*	-0.982	0.117	-8.400	0.000	1.126
SDA4-12*	-0.989	0.477	-2.076	0.038	1.122
SDA4-16*	-1.127	0.417	-2.703	0.007	1.049
SDA8-8*	1.635	0.756	2.164	0.031	1.021
SDA8-12*	1.554	0.517	3.007	0.003	1.029
R²	0.323				
Moran's I	0.319*				

* P < 0.05

A.5.2.2 warm, high Spatial Error Model

Coefficients	Estimate	Std. Error	z-value	Pr(>|z|)
(Intercept)	-2.854	0.233	-12.238	0.000
avgFD	-0.128	0.219	-0.584	0.559
nMonths*	0.974	0.222	4.378	0.000
DTV_LKW*	0.333	0.112	2.973	0.003
sdDirect*	-0.724	0.176	-4.114	0.000
SDA4-12*	-1.822	0.399	-4.568	0.000
SDA4-16*	-1.348	0.361	-3.731	0.000
SDA8-8	0.778	0.619	1.257	0.209
SDA8-12*	0.990	0.438	2.262	0.024
R²	0.513			
Moran's I	0.080*			

* P < 0.05

A.5.2.3 low HV OLS regression model

Coefficients	Estimate	Std. Error	t-value	Pr(>|t|)	VIF
(Intercept)	-2.493	0.104	-24.092	0.000	
avgFD*	0.977	0.167	5.860	0.000	3.018
nMonths*	0.487	0.123	3.956	0.000	1.545
DTV_LKW*	0.338	0.105	3.214	0.001	1.119
avgHD	0.230	0.147	1.566	0.118	2.170
sdDirect*	-1.051	0.106	-9.925	0.000	1.208
SDA4-16*	-1.017	0.456	-2.231	0.026	1.035
SDA8-8*	3.099	0.841	3.687	0.000	1.016
SDA8-16	4.178	2.200	1.899	0.058	1.005
R²	0.444				
Moran's I	0.337*				

* P < 0.05

A.5.2.4 low HV Spatial Error Model

Coefficients	Estimate	Std. Error	z-value	Pr(>|z|)
(Intercept)	-2.478	0.210	-11.799	0.000
avgFD*	1.080	0.155	6.987	0.000
nMonths*	0.291	0.101	2.885	0.004
DTV_LKW*	0.288	0.093	3.092	0.002
avgHD	0.047	0.136	0.350	0.726
sdDirect*	-0.537	0.154	-3.489	0.000
SDA4-16*	-1.524	0.397	-3.841	0.000
SDA8-8*	1.779	0.721	2.467	0.014
SDA8-16*	4.492	1.809	2.484	0.013
R²	0.593			
Moran's I	0.073*			

* P < 0.05

Declaration

I hereby declare that the submitted thesis is the result of my own, independent work. All external sources are explicitly acknowledged in the thesis.

_____
Jan Guddal

_____
Place/Date