CHALMERS
UNIVERSITY OF TECHNOLOGY

Machine Learning-Based Prediction of International Roughness Index for Road Maintenance

Master's thesis in Infrastructure and Environmental Engineering

ELMUIZ IBRAHIM

DEPARTMENT OF ARCHITECT AND CIVIL ENGINEERING
CHALMERS UNIVERSITY OF TECHNOLOGY

Gothenburg, Sweden 2026
www.chalmers.se

Master’s Thesis 2026

Machine Learning-Based Prediction of International Roughness Index for Road Maintenance

ELMUIZ IBRAHIM

Department Of Architect and Civil Engineering
Chalmers University of Technology
SE-412 96 Göteborg
Sweden
Telephone + 46 (0)31-772 1000
Göteborg, Sweden 2026

Machine Learning-Based Prediction of International Roughness Index for Road Maintenance
ELMUIZ IBRAHIM

Examiner: Kun Gao, Department of Architecture and Civil Engineering
Supervisor: Erik Oscarsson, Swedish Transport Administration.
Co. Supervisor: Kun Gao, Department of Architecture and Civil Engineering

Department Of Architect and Civil Engineering
Chalmers University of Technology
SE-412 96 Göteborg
Sweden
Telephone + 46 (0)31-772 1000
Göteborg, Sweden 2026
Chalmers University of Technology

Abstract

Road roughness is a key indicator of pavement condition and significantly affects transportation safety, ride quality, vehicle operating costs, and maintenance planning. The International Roughness Index (IRI) is widely used by road administrations to assess pavement performance and support maintenance decision-making. IRI accurate prediction is important to optimise the maintenance strategies by upgrading long term sustainability of road infrastructure management. This thesis investigates the application of machine learning techniques for predicting IRI changes on the Swedish E4 highway, using data obtained from Trafikverket's Pavement Management System (PMS4) and the Road Weather Information System (VVIS). The study integrates pavement condition, traffic, maintenance, structural, and climatic variables to predict IRI changes based on data from approximately 130 km of the E4 corridor. Several predictive models were evaluated, including Random Forest and Ridge Regression. Following feature selection and correlation analysis, a final set of twelve explanatory variables was retained. Model performance was evaluated using K-fold cross-validation together with commonly used regression performance indicator, namely the coefficient of determination (R^2), Root Mean Square Error (RMSE), and Mean Absolute Error (MAE).

The results demonstrate that the Random Forest model outperformed Ridge Regression, achieving an R^2 of 0.70, an RMSE of 0.13, and an MAE of 0.078. Previous IRI values, maintenance history, rut depth, and traffic loading were identified as the most influential predictors. Weather-related predictors likewise demonstrated a meaningful contribution to overall model accuracy, reinforcing the significance of incorporating environmental exposure factors within pavement degradation forecasting framework.

The findings indicate that Random Forest provides a reliable framework for pavement roughness prediction and can support data-driven maintenance planning and long-term asset management within the Swedish road network.

Keywords: International Roughness Index (IRI); Random Forest; PMS4; pavement performance prediction.

Acknowledgements

The research presented in this master's thesis was conducted in partnership wit Trafikverket and the Department of Architecture and Civil Engineering at Chalmers University of Technology.

First and foremost, sincere gratitude is expressed to Professor Kun Gao, examiner of this thesis, for his valuable feedback, guidance, and support throughout the research process. His thoughtful observations and detailed feedback have played a considerable role in elevating the overall standard of this work.

Deepest appreciation is extended to Dr. Erik Oscarsson, Pavement Specialist at Trafikverket, for providing the opportunity to conduct this study in collaboration with Trafikverket, and for his continuous support, professional guidance, and expertise in pavement management throughout the project.

Special thanks are extended to Omkar Parishwad for his assistance in developing the machine learning models and for sharing his knowledge and guidance throughout the modelling process. His support greatly contributed to overcoming technical challenges and improving the quality of the analysis.

Thanks are also extended to all colleagues at Trafikverket who contributed data, technical support, and valuable discussions during the course of this research.

Finally, heartfelt gratitude is expressed to Patricia Roffel, for her unwavering encouragement, patience, and support throughout this journey. Her understanding and motivation were invaluable during both the challenging and rewarding stages of this work.

Göteborg, Sweden

June 2026

Elmuiz Ibrahim

List of Acronyms

AADT	Annual Average Daily Traffic
ANN	Artificial Neural Network
AI	Artificial Intelligence
CV	Cross-Validation
GPS	Global Positioning System
HMA	Hot Mix Asphalt
IRI	International Roughness Index
LiDAR	Light Detection and Ranging
LTPP	Long-Term Pavement Performance
LVR	Low-Volume Road
MAE	Mean Absolute Error
MDI	Mean Decrease in Impurity
ML	Machine Learning
MSE	Mean Square Error
OOB	Out-of-Bag
PCA	Principal Component Analysis
PCI	Pavement Condition Index
PDP	Partial Dependence Plot
PMS4	Pavement Management System (version 4)
R^2	Coefficient of Determination
RF	Random Forest
RMSE	Root Mean Square Error
SEK	Swedish Krona (Svenska kronor)
VIF	Variance Inflation Factor

VVIS	VägVäderInformationsSystem (Road Weather Information System)
XGBoost	Extreme Gradient Boosting

Table of Contents

List of Figures

List of Tables

1. Introduction

Road infrastructure forms a fundamental component of modern society, enabling the efficient movement of people, goods, and services while supporting economic development, accessibility, and social cohesion. Pavement structures are continuously subjected to repeated traffic loading, environmental and climatic influences, and material aging, causing them to deteriorate over time. This deterioration manifests as surface irregularities that adversely affect ride quality, vehicle operating costs, fuel consumption, exhaust emissions, and traffic safety, with consequent implications for the broader socioeconomic performance of transport systems (Lin, Yau, & Hsiao, 2003; Rita, Ferreira, & Flintsch, 2021).

Road surface unevenness are commonly quantified using the International Roughness Index (IRI), a standraised indicator developed to enable consistent assessment of pavement condition across road network (Sayers, Gillespie, & Queiroz, 1986). The importance of IRI extend beyond condition monitoring, as its strongly associated with user comfort and operational efficiency. Poor pavement smoothness may lead to greater fuel requirement, increased vehicle maintenance needs, and higher environmental impacts, thereby imposing additional economic costs on both road users and infrastructure agencies (Tamagusko & Ferreira, 2023).

Accurate prediction of IRI progression is therefore essential for optimising maintenance planning, reducing lifecycle costs, and enabling the efficient allocation of limited infrastructure budgets. Recent advances in machine learning have created new opportunities for modelling the complex, nonlinear deterioration processes that govern pavement performance, offering the potential to improve prediction accuracy and support more proactive maintenance decision-making.

1.1 Problem Statement

Despite the availability of extensive pavement condition datasets and continued advances in machine learning, accurately forecasting pavement roughness remains a significant challenge. Predicting pavement deterioration remains as challenging task due to the complex interaction among traffic demand, pavement properties, environmental influences, and previous maintenance interventions. While machine learning techniques have shown promising result in pavement performance forecasting, several challenges remains, including limited model transparency, difficult in comparing result across studies, and effective integration of heterogenous data sources. In addition, research based on comprehensive Swedish pavement datasets remain relatively limited, particularly for the prediction of International Roughness Index development. Consequently, there is a clear need to develop and evaluate robust predictive models capable of forecasting IRI progression and identifying the key factors governing pavement unevenness within the Swedish road network.

1.2 Objectives

This study seeks to build and assess predictive machine learning framework capable of estimating future pavement roughness condition along the E4 corridor. Roughness is quantified through the International Roughness Index, with input data sourced from Pavement Management System maintained by Trafikverket, The specific objectives are as follows.

- Examination and interpretation of IRI measurements and related pavement performance indicator extracted from PMS4 database.
- To develop and evaluate machine learning models capable of forecasting IRI progression for the period 2026–2030 using a Python-based implementation.
- To identify and assess the key pavement, traffic, and environmental variables influencing pavement unevenness development.
- To compare the predictive performance of the developed models and determine the most suitable approach for IRI forecasting.
- To demonstrate the potential of machine learning-based prediction for supporting proactive, data-driven pavement maintenance planning.

1.3 Research Questions

The study seeks to address the following research questions:

1. How reliably machine learning algorithms forecast upcoming IRI values when trained longitudinal records encompassing pavement condition history, traffic characteristic, structural properties, and climatic exposure?
2. Which machine learning model provides the strongest predictive performance for forecasting IRI progression on the E4 highway?
3. Which pavement, traffic, and environmental variables exert the greatest influence on pavement unevenness development?

2. Background and Literature Review

2.1 Swedish Road Infrastructure and Pavement Condition Monitoring

The Swedish national road network is managed and maintained by Trafikverket, encompassing a total of roughly 98,000 km of roads, 17,000 bridges, and approximately 100 tunnels. Beyond the state-owned network, around 75,000 km of privately operated roads are supported through government subsides. To support the long-term functionality and sustainability of the transportation system, Trafikverket employs systematic asset management practices and standardised performance indicators to evaluate pavement conditions and prioritise maintenance interventions.

The strategic importance of infrastructure maintenance has been reinforced by the Swedish government's National Transport Infrastructure Plan 2026–2037, which represents a historically high investment exceeding SEK 1,200 billion. Of the total SEK 1,171 billion to be invested over the planning period, approximately SEK 343 billion is allocated specifically to road maintenance, reflecting the long-term commitment to ensuring transport efficiency, socioeconomic benefit, and infrastructure sustainability (Trafikverket, 2025).

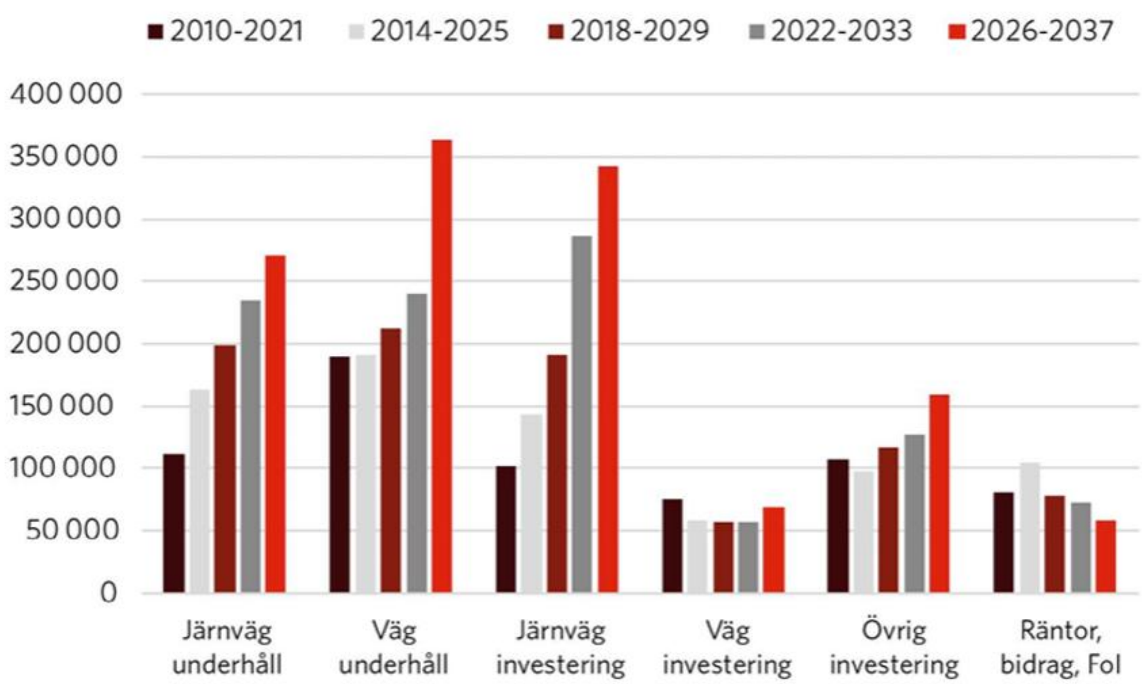

Figure 1. Funds for maintenance and development (investments) 2026–2037 (Trafikverket, 2025).

2.2 International Roughness Index and Pavement Condition Monitoring

The development of International Roughness index can be traced to the Road Roughness Experiment carried out in Brasília, Brazil, in the early 1980s. the primary objective was to create a common reference measures that would facilitate consistent evaluation of comparison of pavement roughness across different road networks and measurement technologies. Two principal categories of measurement methods were evaluated: profilometric methods, which directly measure the longitudinal road profile, and response-type systems, which assess vehicle response to surface irregularities. Based on these findings, IRI was defined as a single stable index derived from a mathematical model simulating vehicle response to road profiles at a reference speed of 80 km/h. It has since become a fundamental parameter in pavement management systems worldwide, closely linked to ride quality, vehicle operating costs, and pavement deterioration assessment (Sayers, Gillespie, & Queiroz, 1986).

In Sweden, systematic pavement condition monitoring has been conducted since 1987 using specialised measurement vehicles equipped with laser-based profilographs and advanced sensing technologies. Collected data are geo-referenced using the SWEREF 99 TM coordinate system, enabling consistent spatial positioning and longitudinal analysis of pavement condition over time. These vehicles acquire high-resolution measurements of IRI, rut depth, and surface distress indicators, providing a comprehensive basis for network-level condition assessment and maintenance planning (Andrén, Eriksson, & Lundberg, 2014).

More recently, vehicles equipped with 360-degree dual LiDAR scanners, imaging systems, and GPS positioning technologies have enabled the creation of detailed digital representations of road infrastructure and surroundings, capturing pavement parameters at 20-metre intervals along both wheel tracks as well as the road environment up to 100 meters on each side (VTI, n.d.). Trafikverket classifies the development of pavement condition monitoring into three phases: foundation (1987–2005), standardisation (2005–2009), and digital integration (2010–present). At present, pavement condition assessment is primarily based on four performance indicator which are: IRI, rut depth, edge deformation, and surface cracking. Together, these measures provide a comprehensive evaluation of pavement functionality and structural condition. These metrics underpin network-level planning, guide maintenance decisions and contractor control, support functional contract verification, and provide data for research and development (Lundberg & Thomas, 2010).

Trafikverket's maintenance standards address both functional condition, encompassing ride quality and safety, and technical condition, covering structural integrity. Threshold values trigger maintenance actions to ensure that interventions are technically justified and costefficient, with prioritisation based on traffic volume and socioeconomic impact (Trafikverket Underhållsstandard, 2011).

2.3 Pavement Management System (PMS4)

To manage the extensive condition data collected across the national road network, Trafikverket maintains the Pavement Management System (PMS4), which stores historical pavement condition measurements alongside information on pavement structure, maintenance history, traffic characteristics, and environmental conditions (PMS4, n.d.). PMS4 serves as a key decision-support tool for maintenance planning and infrastructure management (Lundberg & Thomas, 2010). The availability of long-term performance data within PMS4, combined with the integration of traffic, material, and weather variables, provides a valuable foundation for developing machine learning-based predictive models.

2.4 Machine Learning for Pavement Deterioration Prediction

Historically, IRI prediction models relied on linear or nonlinear regression techniques. However, the increasing availability of large infrastructure datasets and advances in computational tools have encouraged the adoption of machine learning approaches, including Artificial Neural Networks (ANNs) (Hossai, Gopisetti, & Miah., 2018). ANNs, inspired by the structure and function of the human brain, are widely applied in engineering and environmental modelling due to their capacity to capture complex, nonlinear relationships that are difficult to represent using traditional statistical methods. Comprising an input layer, one or more hidden layers, and an output layer, ANNs are commonly regarded as black-box models in which interconnected neurons are linked through weighted connections adjusted iteratively during training (Braspening, F.Thujisman, & Weijters, 1995). Several researchers have applied ANNs to pavement engineering problems, demonstrating their utility for IRI modelling and pavement performance assessment (Abdelaziz, El-Hakim, & El-Badawy, 2020; Meshram & Kisi, 2020).

The Random Forest algorithm, first proposed by (Breiman, 2001), the method generates a large collection of decision trees by repeatedly sampling the training datasets and selecting random subsets of predictor variable during the construction. Instead on relying on a single decision tree, the algorithm combine the prediction from all generated trees to produce more stable and accurate estimate. This ensemble strategies helps improve predictive performance while reducing sensitivity to noise and overfitting. An additional advantages of RF is its use of out-of-bag (OOB) data — observations excluded from the training of individual trees — which provides an unbiased internal estimate of model performance and enables assessment of variable importance, making RF a robust and widely applicable method for complex, high-dimensional datasets.

2.5 Machine Learning Applications in Pavement Performance Prediction

The prediction of pavement performance has undergone a significant shift over recent decades, transitional away from classical statistical techniques such as multiple and non linear regression to estimate future pavement condition (El-Badawy, Elhadidy & Elbeltagi, 2021; Mactutis, Alavi, & Ott, 2000). Although these methods have contributed to understanding deterioration patterns, their ability to represent the complex interaction among traffic demand, environmental factors, and pavement structure often limited.

Recent developments in machine learning have introduced alternative modelling approaches capable of capturing non linear relationships and processing large volume of heterogeneous data. In an investigation by Gong, Sun, Shu, and Huang (2018), RF was deployed to estimated IRI values across Hot Mix Asphalt (HMA) road sections, yielding notably greater accuracy and stability relative to standard linear regression benchmarks. The analysis was based on more than 11,000 pavement observations from the Long Term Pavement Performance (LTPP), with the datasets divided into 80/20 training and testing subsets. The result demonstrated that RF model achieved a high level of predictive accuracy and consistently outperformed linear regression approach, with coefficient of determination values exceeding 0.95 for both model development and validation datasets.

Marcelino et al. (2019) proposed a machine learning framework for long-term IRI prediction over a 5–10 year forecasting horizon using the LTPP database, evaluating multiple RF variants and emphasising model generalisation capability as a key criterion for reliable long-term forecasting. Karballaeezadeh et al. (2020) applied several machine learning techniques to predict both IRI and the Pavement Condition Index (PCI), concluding that Random Forest provided the highest prediction accuracy among all evaluated models. Regression-based approaches have also been applied to low-volume road networks; Albuquerque and Núñez (2011) developed multiple linear regression models for asphalt pavements in northeastern Brazil, finding that IRI progression follows an exponential trend over time, with R^2 values of 0.87 and 0.94.

In the Swedish context, IRI and rut depth have served as core pavement condition parameters since the late 1980s. Trafikverket has progressively refined its measurement methods, incorporating advanced technologies and analytical models to improve accuracy and regional comparability (Lundberg & Eriksson, 2019; Lundberg & Thomas, 2010). Taken together, the body existing research consistently demonstrates that the data-driven machine learning approaches deliver considerably greater predictive precision than traditional based-method, with a particular advantages in representing the complex relationships among climatic exposure, traffic loading, and pavement structure factors that drive surface deterioration.

3. Methodology and Data Source

3.1 Study Area and Data Sources

This study utilises two primary data sources: pavement condition records are extracted from PMS4 database operated by Trafikverket, and meteorological data acquired from the Swedish Road Weather Information System (VVIS). Historical datasets spanning 2021 to 2025 were extracted from both systems and used as the primary input for model training and calibration. These datasets were implemented within a Python-based analytical framework to develop a predictive model for pavement roughness progression.

The modelling approach is applied to forecast pavement roughness development for the period 2026 to 2030. Specifically, the model is trained on observed data from 2021 to 2025 and subsequently applied to predict future pavement condition evolution from 2026 onwards. By grounding prediction in observed structural, traffic, and environmental inputs, this framework produces data-informed characterisation of how pavement degradation unfolds under real world operating conditions.

The analysis is conducted along selected sections of the E4 highway in Sweden, a strategically important transport corridor extending approximately 130 km from southern to central Sweden. The route is characterised by heterogeneous climatic exposure and varying traffic loading conditions, making it particularly suitable for analysing long-term pavement performance under Nordic environmental conditions. Integrating pavement condition records with weather related observations facilitate a through examination of the combined influence that structural properties, traffic loading, and climatic variables exert on the progression of road surface degradation.

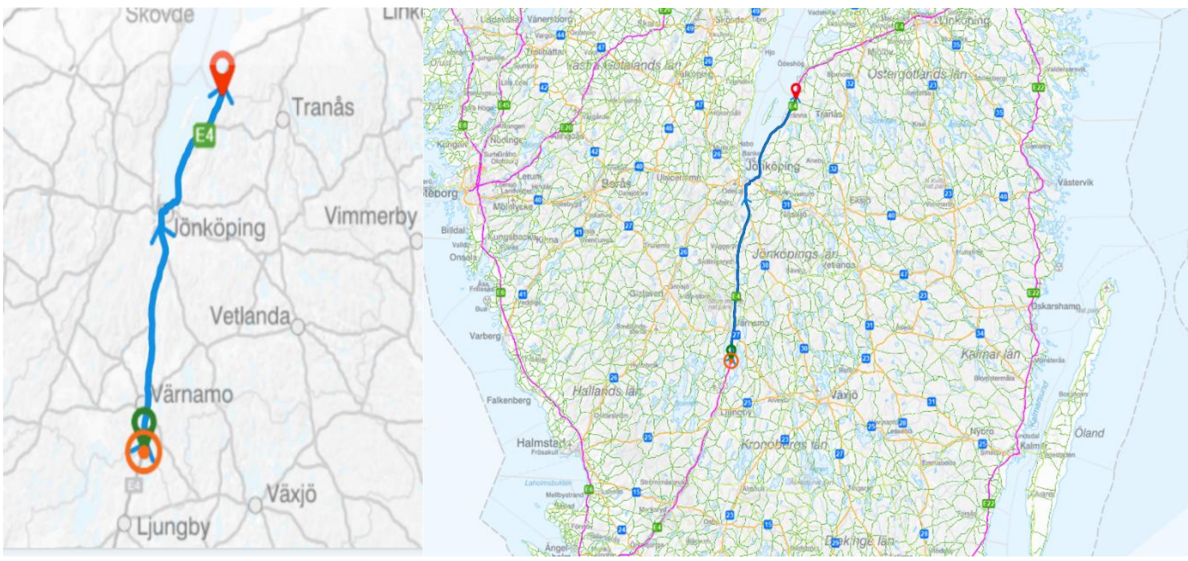

Figure 2. The E4 road is shown in blue as an extension of road corridor in purple.

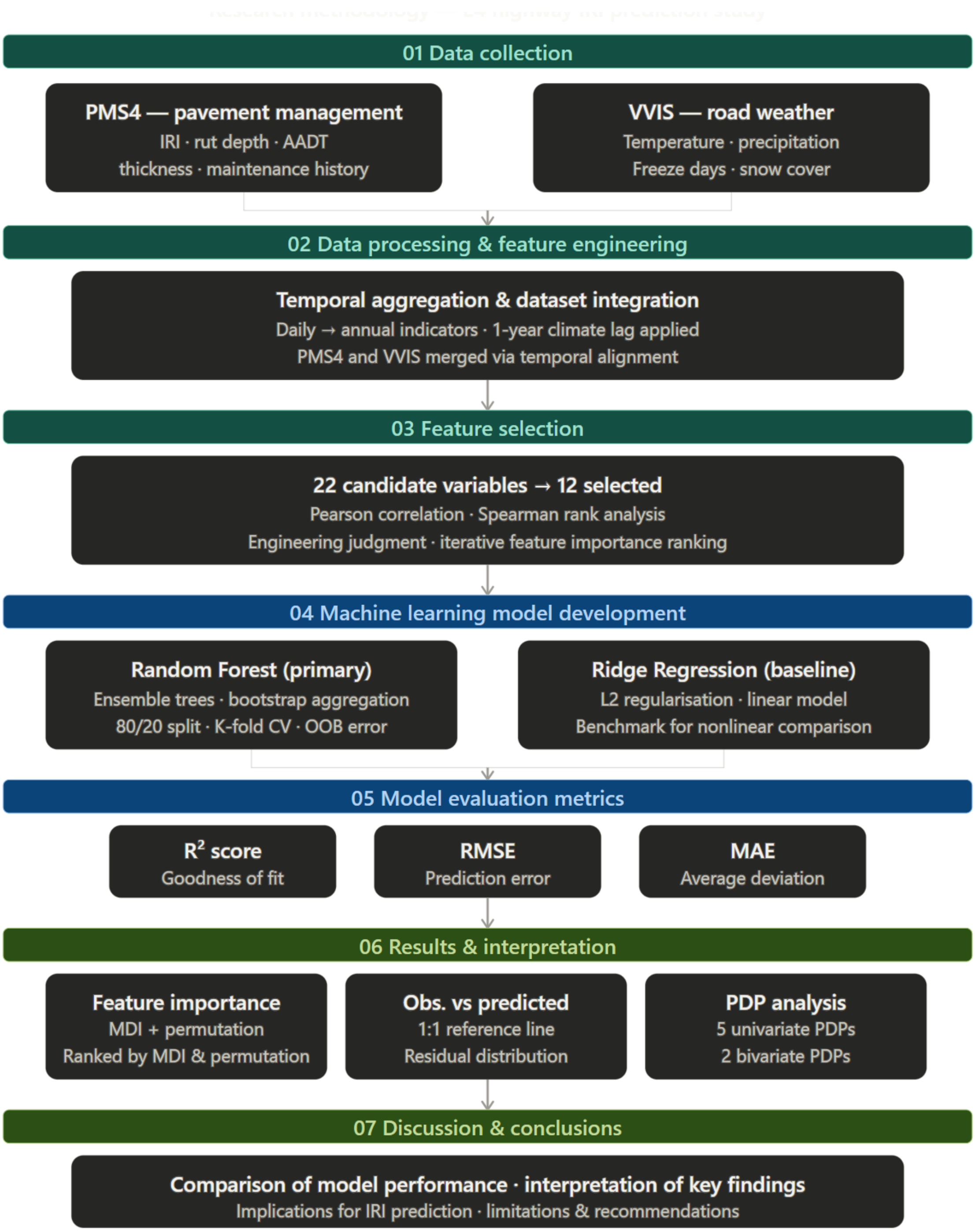

Figure 3. . Methodology workflow chart

Figure 3 presents the overall study workflow, commencing with data collection, followed by data preprocessing and feature engineering. Relevant predictors are subsequently selected and used to develop and compare machine learning models for pavement roughness prediction. Model performance is evaluated using standard regression metrics, and the best-performing model is further interpreted using feature importance and partial dependence analysis. Finally, the results are analysed and discussed to identify key drivers of pavement deterioration.

3.2 Pavement Management System (PMS4)

The Pavement Management System (PMS4) is the primary database used by Trafikverket for storing and managing pavement condition and road inventory data across the Swedish national road network. In this study, PMS4 served as the main data source following selection of the E4 highway as the study corridor and definition of the required pavement layers and attributes for analysis. The database provides detailed information for each road segment, typically aggregated at 20 m and 100 m intervals in both travel directions. The extracted dataset includes key variables such as IRI, rut depth, pavement thickness, road width, number of lanes, vehicle speed, Annual Average Daily Traffic (AADT), maintenance history, and other relevant road structure and inventory parameters.

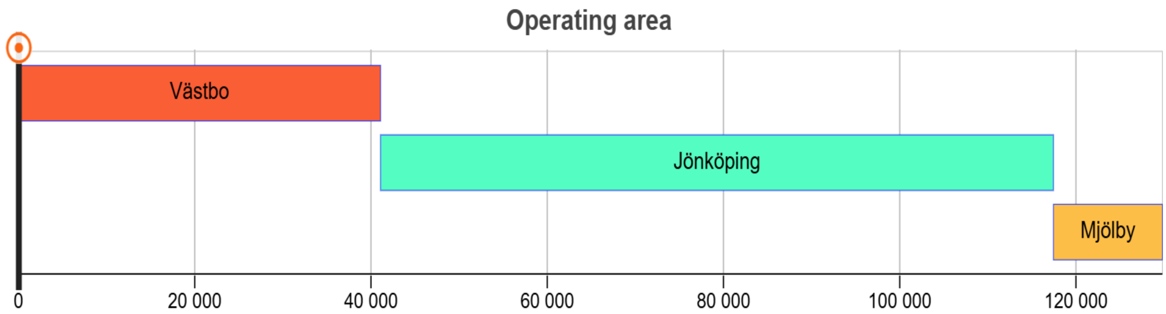

Figure 4. E4 Highway Corridor and Major Cities Along the Route in Sweden

The figure above illustrates the spatial extent of the selected E4 highway study corridor, which is divided across three administrative regions. The analysed road section has a total length of approximately 130,000 m, originating in the northern section within Västra Götaland (0–41,000 m), continuing through the Jönköping region (41,000–118,000 m), and terminating in the southern section over the remaining 12,000 m. The corridor is characterised by significant spatial variability in both traffic loading and environmental exposure, including variations in temperature, precipitation, and freeze–thaw cycles, which are known to influence pavement deterioration mechanisms.

3.3 Meteorological Data (VVIS)

Meteorological data were obtained from Trafikverket VVIS (VägVäderInformationsSystem) stations located in proximity to the study corridor and spatially matched to the corresponding road segments based on geographic proximity. The dataset comprised daily aggregated observations, including air temperature, precipitation, relative humidity, and snow-related variables. Since road surface degradation is substantially driven by environmental stressors most notably cyclic temperature variation, water penetration into pavement layers, and repeated freeze-thaw action incorporated weather derived indicators into the modelling framework is critical for adequately representing the role of climate in pavement performance. Accordingly, these meteorological variables were added to the dataset to strengthen the forecasting capacity of the developed models, and were temporally synchronized with pavement condition observations to maintain coherence between predictors input and measured outcomes. Prior to modelling, the data were subjected to preprocessing steps to address missing values and temporal variability.

3.4 Data Processing and Feature Selections

To ensure compatibility between datasets, daily meteorological observations were aggregated into annual indicators through feature engineering. Variable-specific aggregation methods were applied to preserve the physical meaning of each parameter, as follows:

- Annual mean temperature and relative humidity: arithmetic mean of daily values.
- Total precipitation: cumulative annual sum.
- Snow cover days: total number of days on which snow depth was above zero.
- Frost days: count of days of which minimum daily temperature fell below 0°C.
- Freeze–thaw cycles: number of transitions across 0°C between consecutive days.
- Temperature variation: mean daily diurnal temperature range ($T_{max} - T_{min}$), representing thermal stresses acting on pavement materials.

The resulting engineered variables collectively represent the principle climate-driven process known to contribute to road surface degradation, encompassing water penetration into pavement layers, cracking induced by repeated freezing and thawing and progressive loss of structural integrity in asphalt bound materials.

3.5 Data Integration and Feature Selection

To capture the delayed impact of environmental conditions on pavement performance, a lag-based approach was implemented. Climatic variables were shifted by one year to represent their influence on subsequent pavement condition development, reflecting the cumulative and progressive nature of pavement deterioration processes. The final analytical dataset was established by merging the aggregated annual weather indicators from VVIS with the corresponding pavement condition measurements from PMS4 using temporal alignment procedures. The resulting dataset comprises the following variable categories:

- Pavement condition variables: IRI (target variable), rut depth, pavement structure, and thickness.
- Traffic variables: AADT and the proportion of heavy vehicle within traffic stream.

- Climatic variables: ambient temperature, total precipitation, relative humidity, number of frost days, frequency of freeze thaw transition, and duration of snow cover.

This process enabled the identification of the most relevant variables contributing to pavement roughness progression and enhanced the robustness of the machine learning framework.

3.6 Machine Learning Model Development

The forecasting task was framed as a supervised regression problem and implemented using Python, whereby future IRI value were estimated as an output of combined traffic, structural, and climatic input variables. The developed model was additionally applied to forecast pavement roughness evolution over a five-year horizon (2026–2030). The analysis was based on 12,562 observations collected from both travel directions of the E4 highway, covering two lanes in each direction and four observation years (2022–2025).

A Random Forest regression model was selected as the primary predictive algorithm due to its ability to:

- Capture nonlinear relationships between explanatory variables.
- Model high-order interactions without explicit functional assumptions.
- Provide robust performance under noisy and correlated input data.
- Reduce overfitting through bootstrap aggregation (bagging).

Ridge Regression model was incorporated as a linear reference point against which Random Forest framework could be assessed, allowing the additional predictive benefit gained through adoption of a nonlinear ensemble approach to be objectively quantified.

3.7 Training Procedure and Validation Strategy

The cleaned dataset was partitioned into a training set (80%) and a test set (20%). Cross-validation was applied to ensure robust model evaluation and minimise sampling bias. All preprocessing steps, including feature scaling, were performed within the training pipeline to prevent data leakage. Standardisation was applied to continuous variables to ensure comparability across different measurement scales, particularly between traffic, structural, and climatic indicators.

3.8 Feature Selection

An iterative feature selection process was conducted based on predictive performance and feature importance ranking. A total of 12 explanatory variables were retained in the final model following multiple experimental iterations. Variables with consistently low predictive contribution were excluded to improve model efficiency and reduce multicollinearity. This process ensured that only the most informative and physically relevant predictors were included in the final modelling framework, achieving a balance between model complexity and generalisation capability.

Table 1: Random Forest model used predictor variables.

Variable	Unit	Description	Expected Influence on IRI
Previous IRI (prev_iri)	mm/m	Historical IRI value from the previous survey period, representing existing pavement condition.	Higher previous IRI is associated with higher future IRI, as pavement deterioration is path-dependent.
Previous Rut Depth (prev_rut_depth_mm)	mm	Rut depth measured during the previous observation period, indicating accumulated pavement distress.	Greater rut depth is associated with increased pavement deterioration and roughness progression.
AAT- Annual Average Daily Traffic	vehicles/day	Average number of vehicles per day, serving as a proxy for traffic loading.	Higher traffic volumes accelerate pavement wear and increase roughness over time.
Days Since Last Maintenance	days	Time elapsed since the most recent maintenance intervention.	Longer periods without maintenance generally result in greater deterioration and higher IRI.
Pavement Thickness (thickness_mm)	mm	Thickness of asphalt layer in total	Thicker pavements are more resistant to traffic and environmental damage, reducing IRI growth.
Speed Limit (speed_limit_kmh)	km/h	Posted maximum vehicle speed on the road segment.	Higher speeds may increase dynamic loading and influence deterioration rates.
Road Width (road_width_m)	m	Width of road which reflect traffic distribution and road classification	Wider roads may distribute traffic loads more effectively, exhibiting different deterioration patterns.

Variable	Unit	Description	Expected Influence on IRI
Total Rainfall (rain_mm_total)	mm	Total accumulated precipitation during the observation period.	Rainfall contributes to moisture infiltration, weakening pavement materials and accelerating deterioration.
Freeze–Thaw Days (freeze_day)	days	Freeze-thaw cycle days during the observation	Repeated freezing and thawing induces cracking and material degradation, increasing roughness.
Thermal Stress (thermal_stress)	°C	Indicator of temperature variation and thermal loading experienced by the pavement.	Large temperature fluctuations cause expansion, contraction, and cracking.
Days Since Previous IRI Observation	days	Time interval between consecutive IRI measurements.	Longer intervals allow more deterioration to accumulate between observations.
Maintenance Count Before Observation	count	Total number of maintenance activities recorded prior to the current observation.	Higher counts reflect historically deteriorated sections requiring repeated interventions.

The chosen input variables were selected to collectively represent the principle determinations of road surfaces degradation were existing pavement condition, traffic loading, maintenance history, structural characteristics, and environmental conditions. Previous IRI and rut depth describe the current pavement state, while AADT, speed limit, pavement thickness, and road width capture traffic and structural effects. Rainfall, freeze–thaw cycles, and thermal stress account for climate-related deterioration. And Maintenance variables provide information on past interventions.

3.9 Model Evaluation and Performance Metrics

Model performance was assessed using standard regression metrics widely applied in pavement performance prediction studies. Together, these indicators offer a through assessment of how well the model perform in term of predictive precision, consistency of output, and its capacity to explain variance in the observed data.

The Root Mean Square Error (RMSE) measures the overall magnitude of prediction errors, penalising larger deviations more heavily due to the squared error term:

Equation 1:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2} \qquad (1)$$

The Mean Absolute Error (MAE) measures the mean of the absolute deviations between model output and actual observed values, applying equal weight of all prediction errors of their magnitude:

Equation 2:

$$MAE = \frac{1}{n}\sum_{i=1}^{n} | \hat{y}_i - y_i | \qquad (2)$$

The Mean Square Error (MSE) evaluates the average squared prediction error and is particularly useful during model optimisation:

Equation 3:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2 \qquad (3)$$

The Coefficient of Determination (R^2) assesses the explanatory power of the model, representing the proportion of variance in the observed IRI values explained by the predictive model:

Equation 4:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \qquad (4)$$

Higher R^2 values indicate stronger predictive capability and improved model fit. When considered together, these evaluation measures for judging how well the developed model predicts unseen data, generalizes beyond the training set, and produces dependable outcomes across varying pavement conditions.

3.10 Feature importance and model interpretation

To enhance model interpretability, feature importance analysis was conducted using the internal Random Forest importance metric, specifically Mean Decrease in Impurity (MDI). This analysis quantifies the relative contribution of each variable to prediction accuracy and enables identification of the dominant drivers of pavement deterioration. Variables with consistently low importance were iteratively removed during model refinement to improve stability and reduce redundancy among correlated predictors.

In addition, PDP were generated to investigate the marginal effect of individual explanatory variables on predicted IRI, isolating each variable's contribution while averaging over the effects of all others. Both univariate and bivariate PDPs were produced to examine single-variable effects and pairwise interaction effects, respectively.

3.11 Summary of Methodology

The overall methodology integrates infrastructure, traffic, and environmental data within a data-driven analytical framework. The methodology is organised as an end-to-end analytical workflowdata acquisition from multiple sources, preprocessing and feature engineering, construction of a unified modelling dataset, model development, systematic validation, and interpretability analysis. Through the integration of physically grounded climatic indicators with ensemble machine learning methods this study attempt to enhance the precision of IRI forecasting while generating applicable engineering understanding of pavement degradation behavior under Nordic environmental conditions in Sweden.

4. Results and Discussion

4.1 Overview of Model Development

Various machine learning algorithms and predictor variables combinations were examined within a Python-based environmental to forecast future road roughness from datasets of pavement condition such as traffic, maintenance, and meteorological observations sourced from PMS4 and VVIS. The assessment included two modelling approaches which are Random Forest and Ridge Regression. Random Forest demonstrated the strongest predictive performance and generalisation capability, leading to its selection as the primary predictive framework in this study. Although Random Forest is more complex and less directly interpretable than linear models, its ability to capture nonlinear relationships and variable interactions proved particularly effective for modelling pavement deterioration processes.

K-fold cross-validation was employed to improve the robustness and reliability of model evaluation by ensuring that training and testing were performed across different road segments. This strategy minimized the likelihood of overfitting while yielding a more representative evaluation of how model performs when applied to road segments not seen during training. Furthermore, nonlinear association screening confirm that the progression of pavement degradation is shaped by intricate, nonlinear interdependence spanning structural, traffic, maintenance, and environmental factors. Consequently, the discussion in this chapter primarily focuses on the Random Forest model, while selected Ridge Regression results are presented for comparison purposes.

4.2 Feature Selection and Correlation Analysis

An initial candidate set of twenty-two explanatory variables was defined prior to model development, drawing on the pavement condition, traffic, maintenance, structural, and climatic data available from PMS4 and VVIS. Feature selection was conducted through a two-stage screening process combining statistical analysis with engineering judgement, with the dual objectives of retaining variables that are both statistically associated with IRI and physically interpretable in the context of pavement deterioration. A correlation matrix was generated to examine the relationships between all explanatory variables and their association with IRI, thereby identifying potential multicollinearity and guiding the subsequent feature selection process.

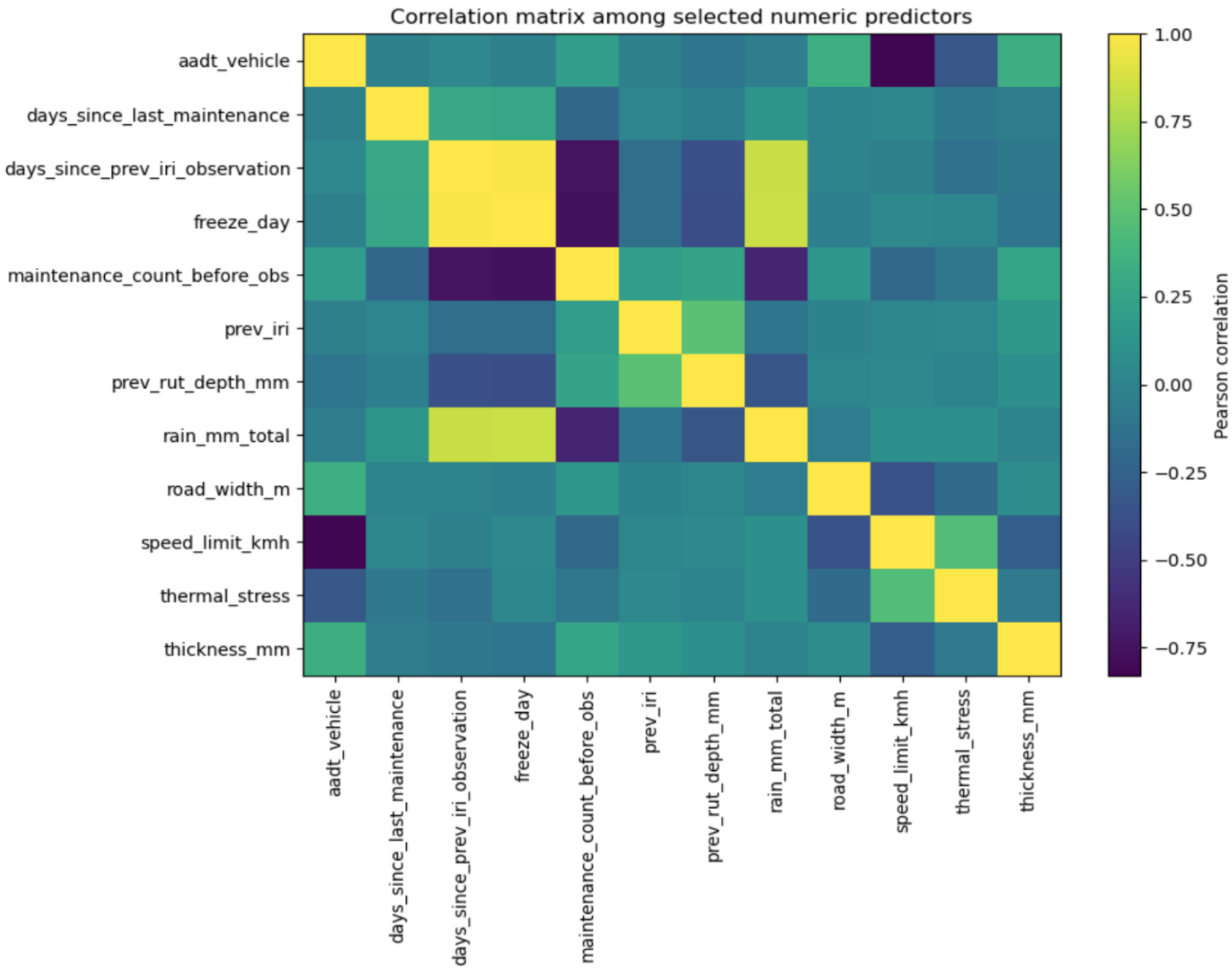

Figure 5. Correlation matrix of input variables

The correlation structure among the candidate input variables is illustrated in the figure above. Several variables exhibit moderate correlations; however, no severe multicollinearity was observed among the selected features. The correlation matrix was used primarily as a screening tool to identify redundant variables and support the feature selection process.

Figure 6 below indicates that previous IRI exhibits the strongest positive linear relationship with the target variable, while several traffic, structural, and maintenance-related variables show weaker linear associations. Linear relationships between each prospective variable predictor and the IRI target variables were assessed throughout the application of Pearson correlation coefficients. This provided an initial ranking of variables based on the strength of their direct linear relationship with future pavement roughness.

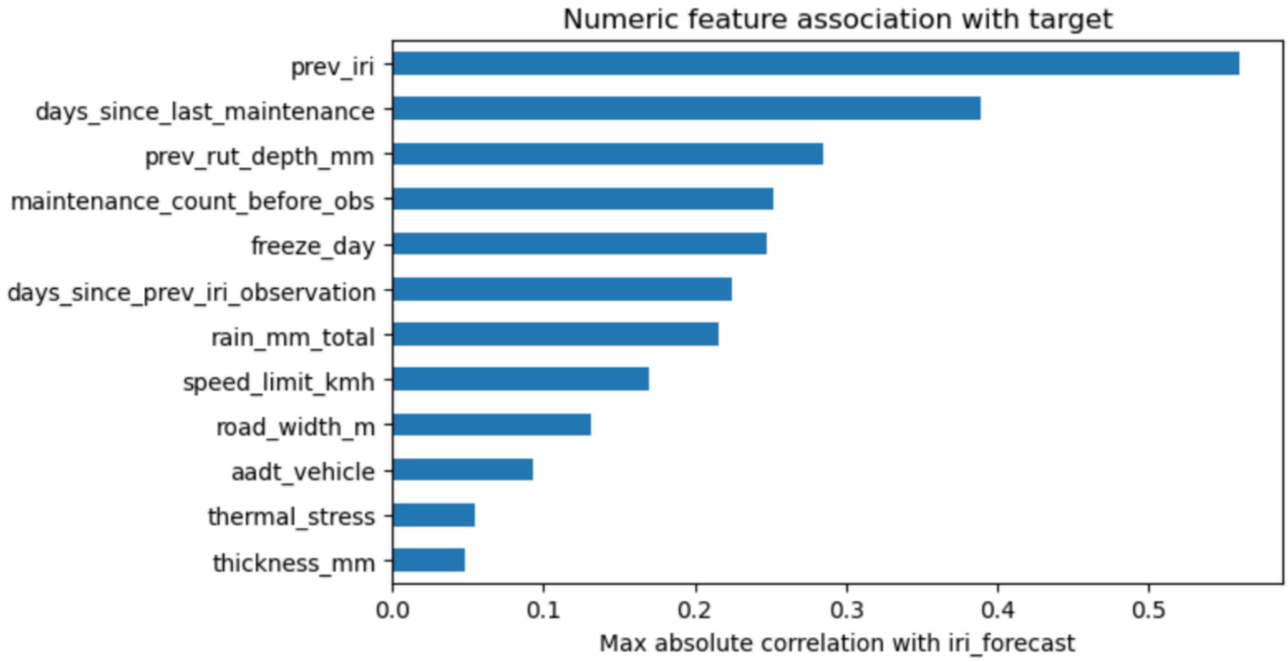

Figure 6. Numerical feature association with IRI.

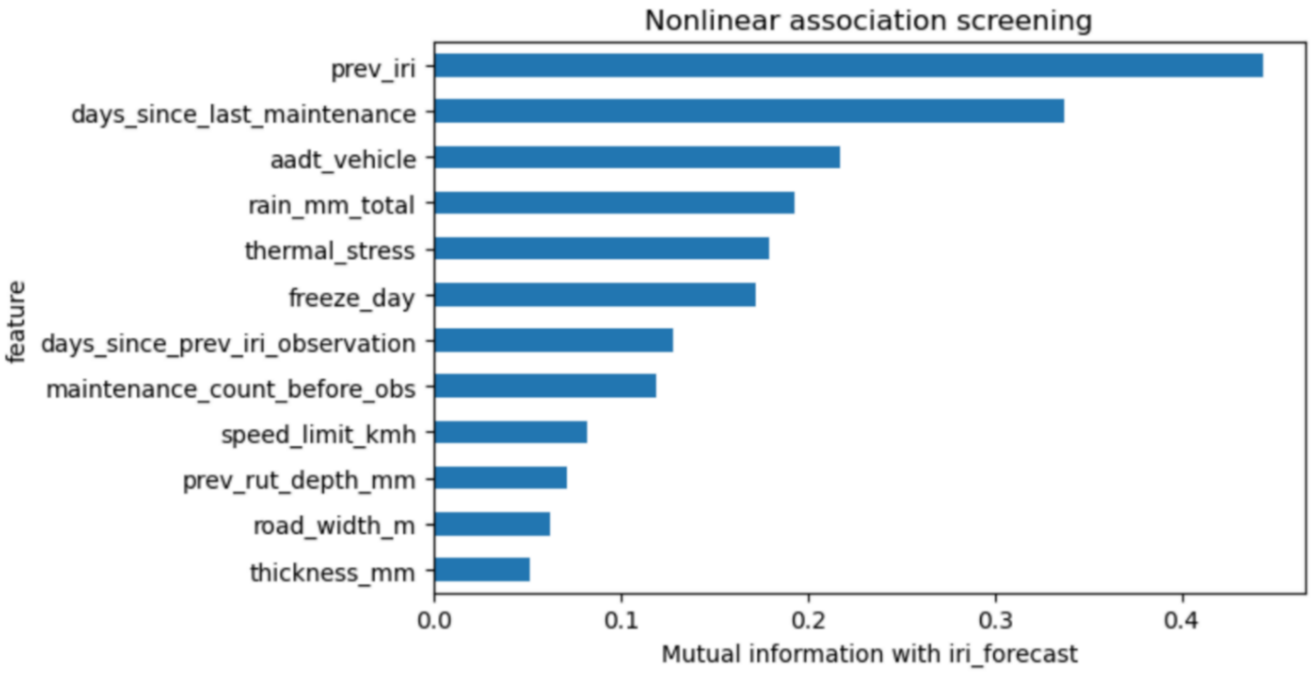

Figure 7. Nonlinear association screening of input features with IRI.

Figure 7 Spearman rank correlation and mutual information analysis were used to identify variables exhibiting nonlinear but consistent monotonic associations with IRI, capturing relationships that Pearson correlation may underestimate. Variables exhibiting consistently negligible association with IRI across both screening methods were excluded, as were variables identified as highly collinear with retained predictors based on variance inflation factor (VIF) assessment.

Based on this combined screening process, the feature space was reduced from twenty-two to twelve variables. This reduction improved model interpretability and reduced the risk of overfitting; however, it may also have excluded variables contributing to nonlinear interaction effects relevant to pavement deterioration mechanisms.

4.3 Model Performance Evaluation

Table 2: Cross-validated performance metrics for Random Forest and Ridge Regression on the held-out test set.

Model	R^2	RMSE (m/km)	MAE (m/km)
Random Forest (primary)	0.70	0.13	0.078
Ridge Regression (baseline)	0.47	0.18	0.121

4.3.1 Random Forest Model

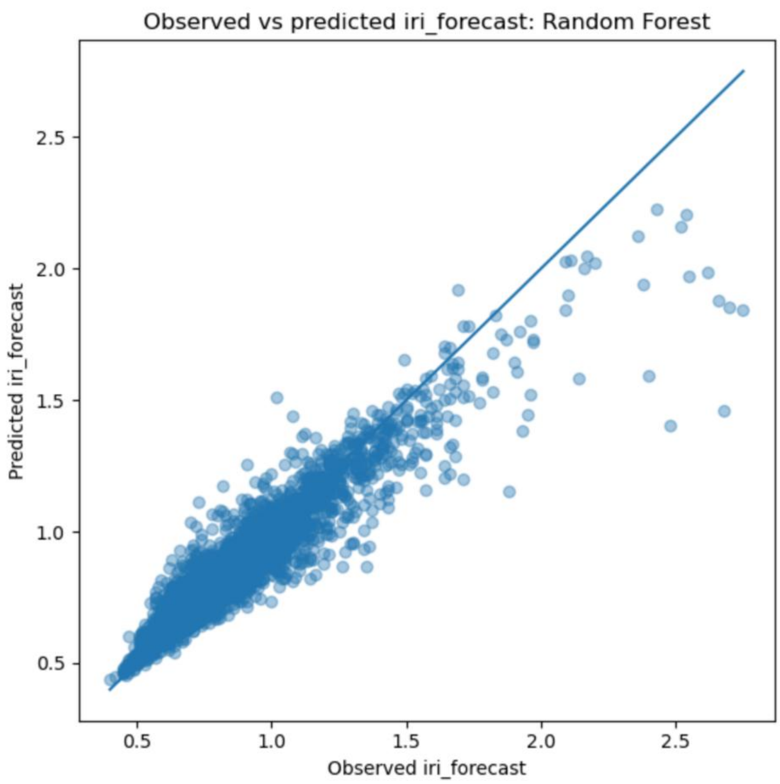

Figure 8. Comparison of predicted and observed IRI using the Random Forest model

The Random Forest model achieves a cross-validated R^2 of 0.70, indicating that approximately 70% of the observed variance in IRI is explained by the model across the heterogeneous pavement conditions represented in the E4 dataset. Error metrics are consistently low across validation folds, with RMSE ≈ 0.13 m/km and MAE ≈ 0.078 m/km, confirming stable and accurate predictions on unseen road segments. The close agreement between training and test set performance, together with the stability of metrics across cross-validation folds, demonstrates that the model generalises well and does not exhibit overfitting despite the inherent flexibility of the ensemble framework.

The assessment of the observed versus forecasted IRI scattered plot reveals that the large majority of predictions are distributed closely around the 1:1 reference line, with limited evidence of systematic directional bias. Increased scatter is observed at the upper end of the IRI range, corresponding to severely deteriorated sections; this pattern is consistent with findings in the broader pavement performance literature and reflects the influence of localised structural failures, heterogeneous or unrecorded maintenance interventions, and extreme climatic events that are not fully represented in the annually aggregated predictor set.

4.3.2 Ridge Regression Baseline

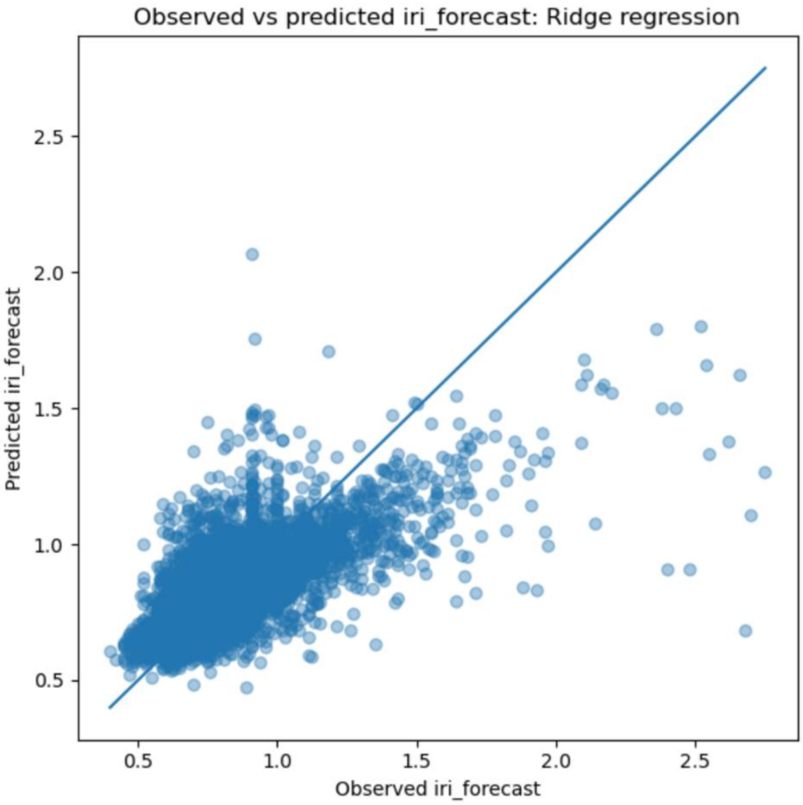

Figure 9. Comparison of predicted and observed IRI using the Ridge Regression model.

Ridge Regression achieves $R^2 = 0.47$, RMSE = 0.18 m/km, and MAE = 0.121 m/km, representing substantially weaker predictive performance than Random Forest across all evaluation criteria. The performance differential — an increase in R^2 of 0.23, a reduction in RMSE of 28%, and a reduction in MAE of 36% for Random Forest — confirms that linear modelling assumptions are inadequate for representing the nonlinear deterioration mechanisms that characterise pavement behavior on the E4 corridor. Although the application of L2 regularisation enhances the numerical robustness of Ridge Regression compared with ordinary least square and offers partial mitigation of multicollinearity among input variables, the model ultimately restricted by inherently linear structure and cannot represent curved predictor response relationship. Ridge Regression is retained as a transparent and reproducible baseline that enables objective quantification of the added predictive value delivered by the nonlinear ensemble approach.

4.3.3 Residual Distribution Analysis

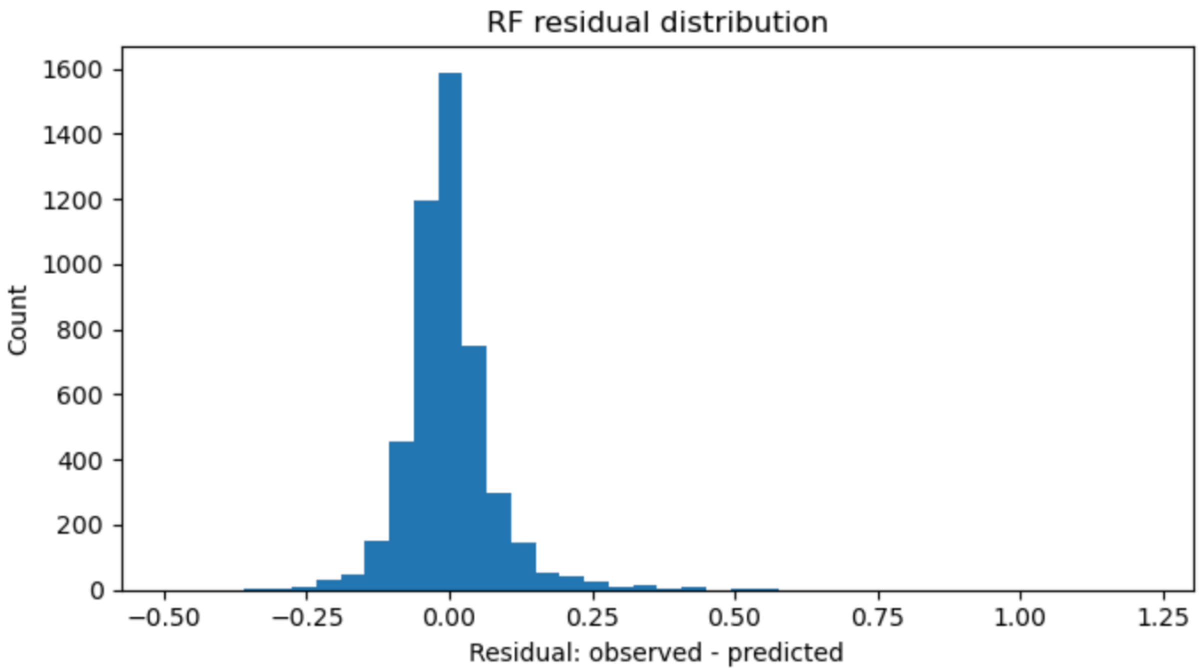

Figure 10. Residual Distribution of the Random Forest Model

The residual histogram illustrates the distribution of prediction errors, computed as the difference between observed and predicted IRI values. The residual distribution is approximately centred around zero, indicating that the model does not exhibit significant systematic overprediction or underprediction. Most residuals fall within a relatively narrow range of approximately −0.25 to +0.25 IRI units, with the highest frequency occurring close to zero. This pattern suggests stable predictive performance, low variance in errors, and satisfactory generalisation across the majority of pavement conditions.

A limited number of larger residuals are observed under extreme pavement conditions, likely associated with sudden pavement failures, unrecorded maintenance interventions, localised material degradation, or climatic effects not fully captured by annually aggregated explanatory variables. Residuals show no systematic pattern with respect to predicted values, confirming the absence of structural bias in the model and supporting its suitability for network-level pavement management applications along the E4 highway corridor.

4.4 Feature Importance Analysis

The relative contribution of each predictor variable to model performance was assessed through the internal node impurity reduction metric of the Random Forest algorithm, widely known as Mean Decrease in Impurity (MDI). The results indicate that previous pavement condition (prev_iri) was the dominant predictor of future IRI progression, with an importance value of approximately 0.464, confirming the cumulative and path-dependent nature of pavement deterioration. Other influential variables included the number of days since the last maintenance intervention, previous rut depth, traffic loading (AADT), and maintenance frequency. Environmental and structural variables, including total rainfall, pavement thickness, thermal stress, and freeze days, also contributed to the prediction process, albeit with comparatively lower importance values.

Table 3: Feature importance from the Random Forest model (MDI), ranked by contribution.

Variable	Importance
Previous IRI	0.464
Days since last maintenance	0.206
Previous rut depth	0.090
AADT	0.052
Maintenance count	0.044
Rainfall total	0.031
Pavement thickness	0.028
Thermal stress	0.020
Freeze days	0.019

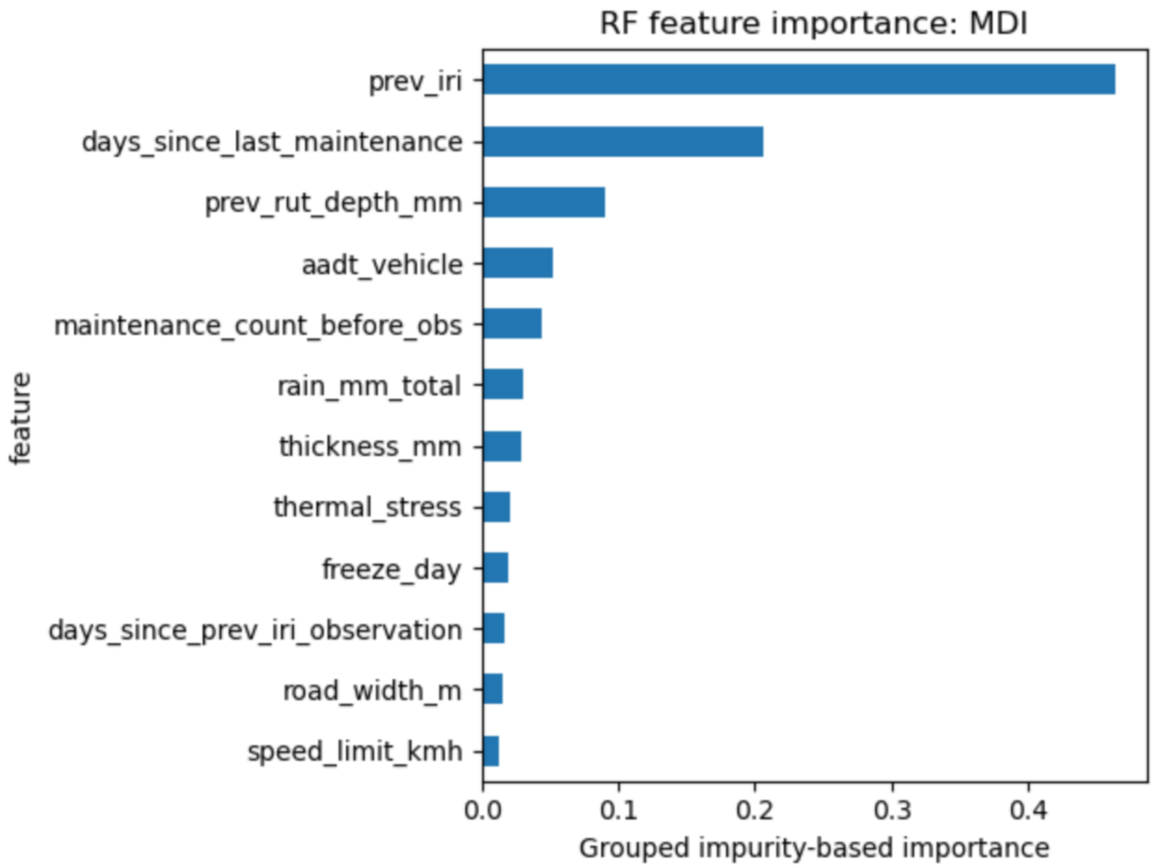

Figure 11. Random Forest feature importance.

To further validate feature relevance, permutation importance analysis was conducted using the holdout dataset. Unlike impurity-based measures, permutation importance evaluates the reduction in predictive performance caused by randomly shuffling each feature, providing a more robust estimate of variable importance based on the model's generalisation capability on unseen data.

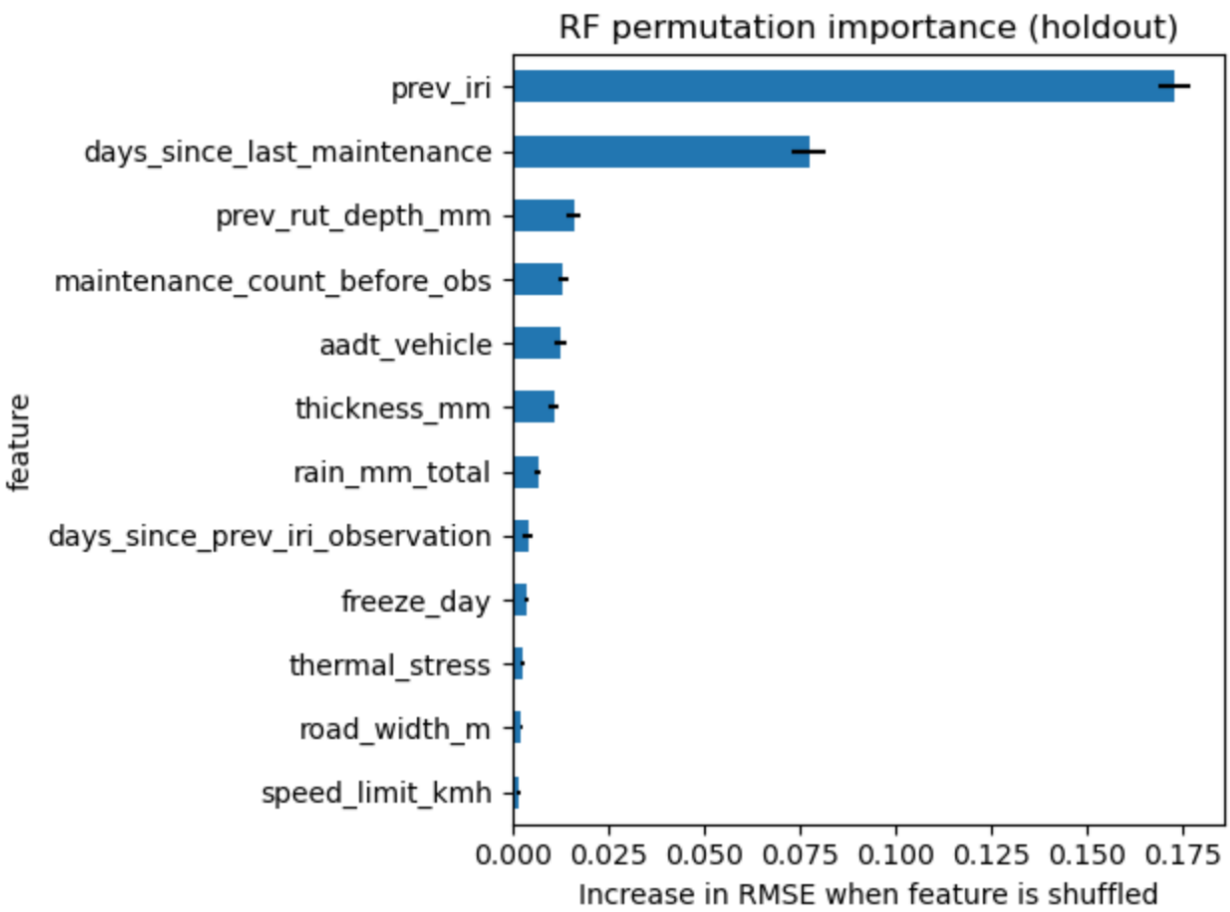

Figure 12. Random Forest feature importance based on the holdout dataset.

The dominance of previous IRI indicates that pavement deterioration is strongly path-dependent and cumulative in nature, where existing pavement condition remains the primary indicator of future deterioration progression. Maintenance-related variables also demonstrated substantial influence. environmental variables such as rainfall, thermal stress, and freeze-day exposure provided measurable explanatory capability. Although their relative importance was lower than historical and maintenance-related variables, their contribution confirms that climatic conditions significantly influence pavement degradation under Swedish environmental conditions. The inclusion of climatic variables improved overall model interpretability and demonstrated the importance of integrating environmental exposure factors into pavement deterioration forecasting models for long-term network-level pavement management.

4.5 Partial Dependence Plot (PDP) Interpretation

The PDPs were generated to investigate the marginal relationship between individual explanatory variables and predicted IRI behaviour, isolating the effect of each variable while averaging over the effects of all others.

4.5.1 PDP Analysis of Previous IRI

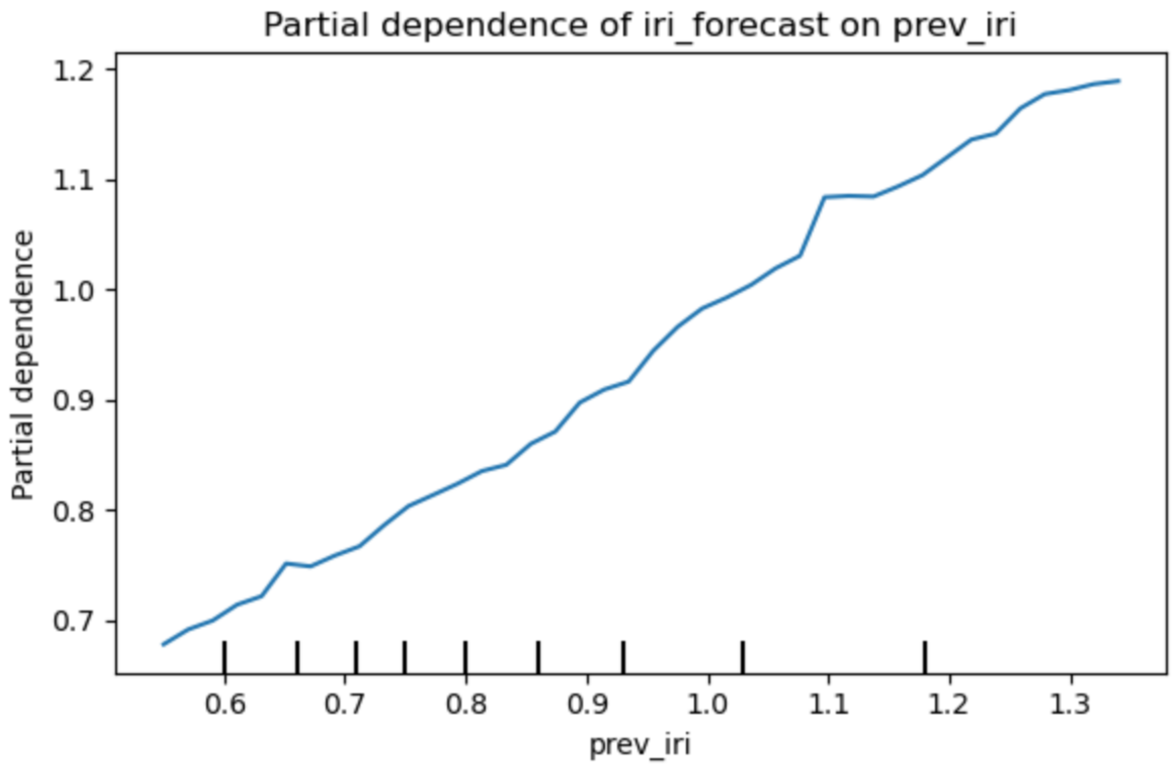

Figure 13. Partial Dependence Plot Analysis of previous IRI on predicted IRI.

The PDP for previous IRI demonstrated a strong positive relationship between previous pavement condition and predicted future roughness, confirming the cumulative nature of pavement deterioration. As previous IRI increased, the model predicted progressively higher future IRI values, indicating that rough pavements tend to remain rough over time due to the temporal continuity of pavement performance. The nonlinear shape of the plot indicates that pavement degradation accelerate at higher roughness which aligned with a widely recognised mechanisms of progressive pavement distress.

4.5.2 PDP Analysis of Pavement Thickness

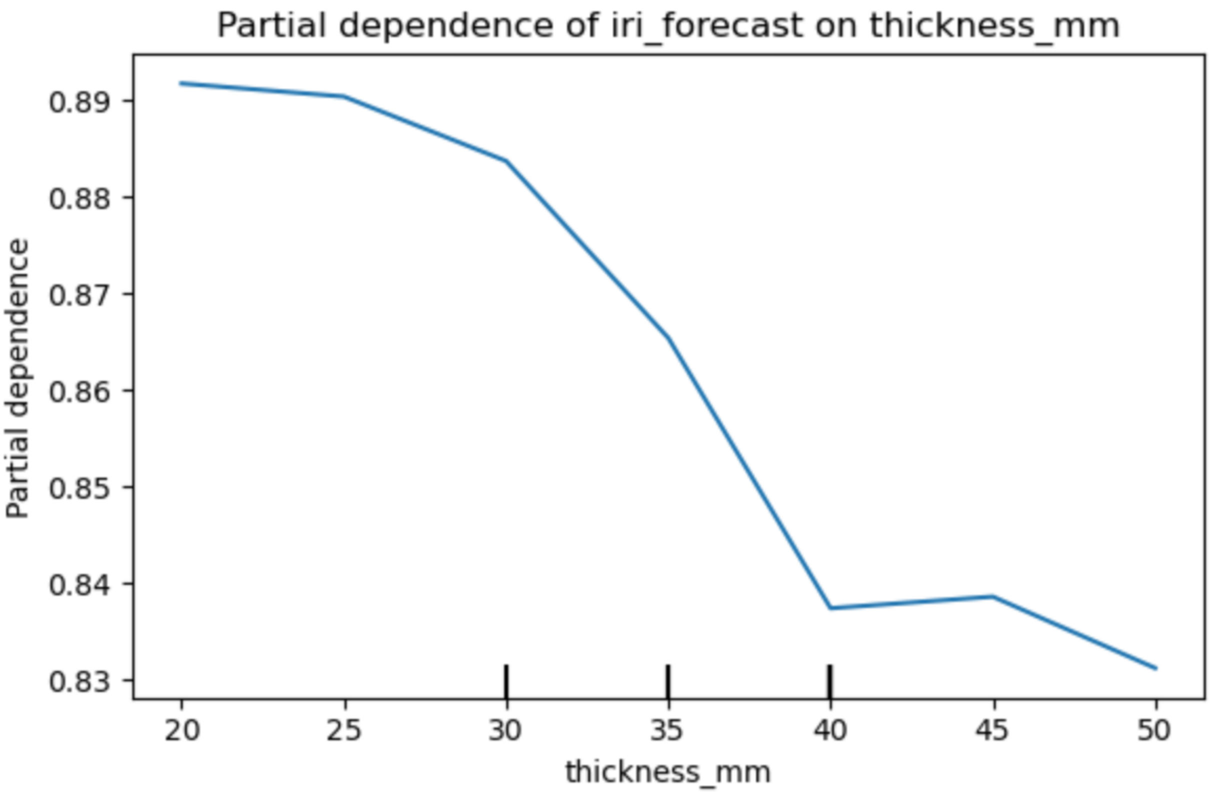

Figure 14. Partial Dependence Plot Analysis of Thickness on predicted IRI.

The PDP for pavement asphalt layers thickness demonstrated a negative relationship between pavement thickness and predicted IRI values. The total depth of asphalt layer explain the ability of pavement structure to resist traffic axel load generated from passing vehicle due to the contraction. The plot revealed that road section with relatively shallow asphalt depths as shown in 20 mm station produce high predicted IRI value therefore higher level of surface degradation, as the pavement increases in thickness the forecasted IRI decreases accordingly. This behaviour is consistent with pavement engineering principles, where thicker pavement structures provide greater load distribution capacity and improved resistance against cumulative traffic-induced damage, resulting in reduced pavement roughness over time.

4.5.3 PDP Analysis of AADT

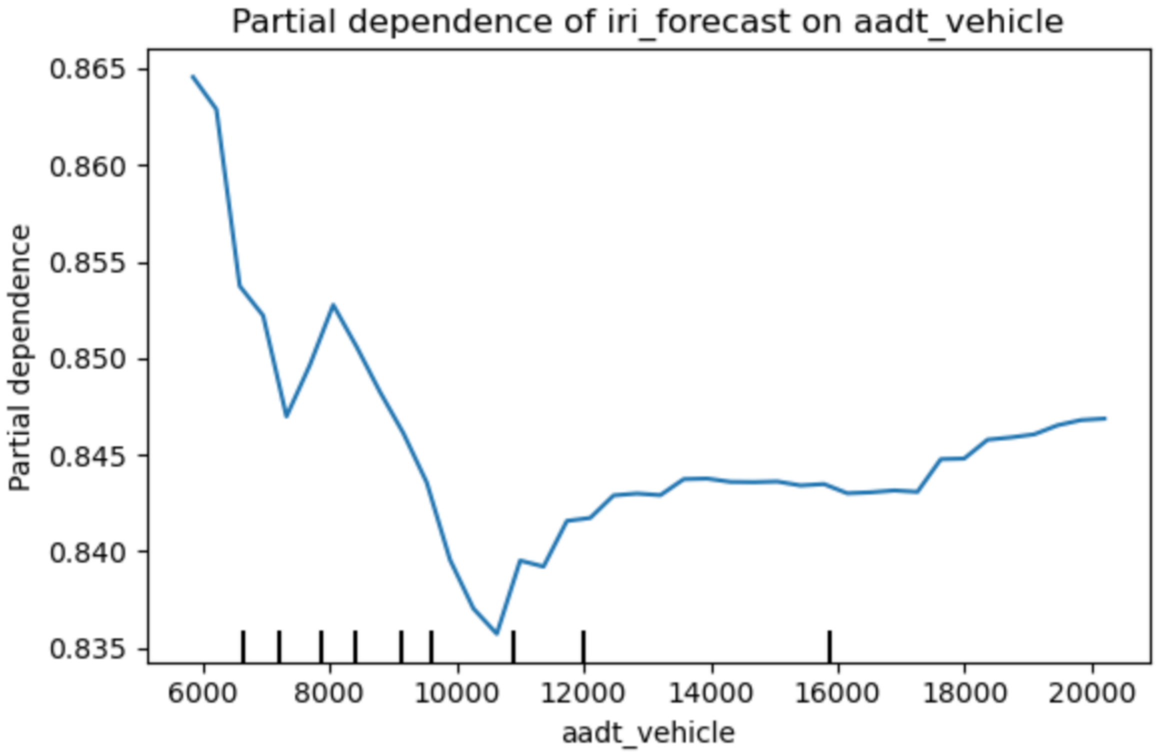

Figure 15. Partial Dependence Plot Analysis of AADT on predicted IRI.

PDP for AADT indicates a nonlinear relationship between traffic volume and predicted IRI values. Predicted IRI initially decreases with increasing AADT before increasing at higher traffic levels. The overall increase in IRI at higher traffic volumes is consistent with pavement engineering principles, as greater traffic loading contributes to fatigue damage, rutting, cracking, and progressive pavement deterioration. The decrease observed at the highest AADT levels is likely caused by overfitting due to the limited number of observations within this range. Random Forest models may fit local patterns in sparsely populated regions of the dataset that are not representative of the underlying physical relationship. Consequently, the apparent decrease in predicted IRI at high AADT values is unlikely to reflect actual pavement behaviour and should instead be interpreted as a modelling artefact resulting from insufficient data coverage. Consequently, the upward trend of the AADT distribution is a clear evident for the representation of physical relationship between cumulative traffic demand and pavement surface roughness.

4.5.4 PDP Analysis of Effect of Maintenance Frequency

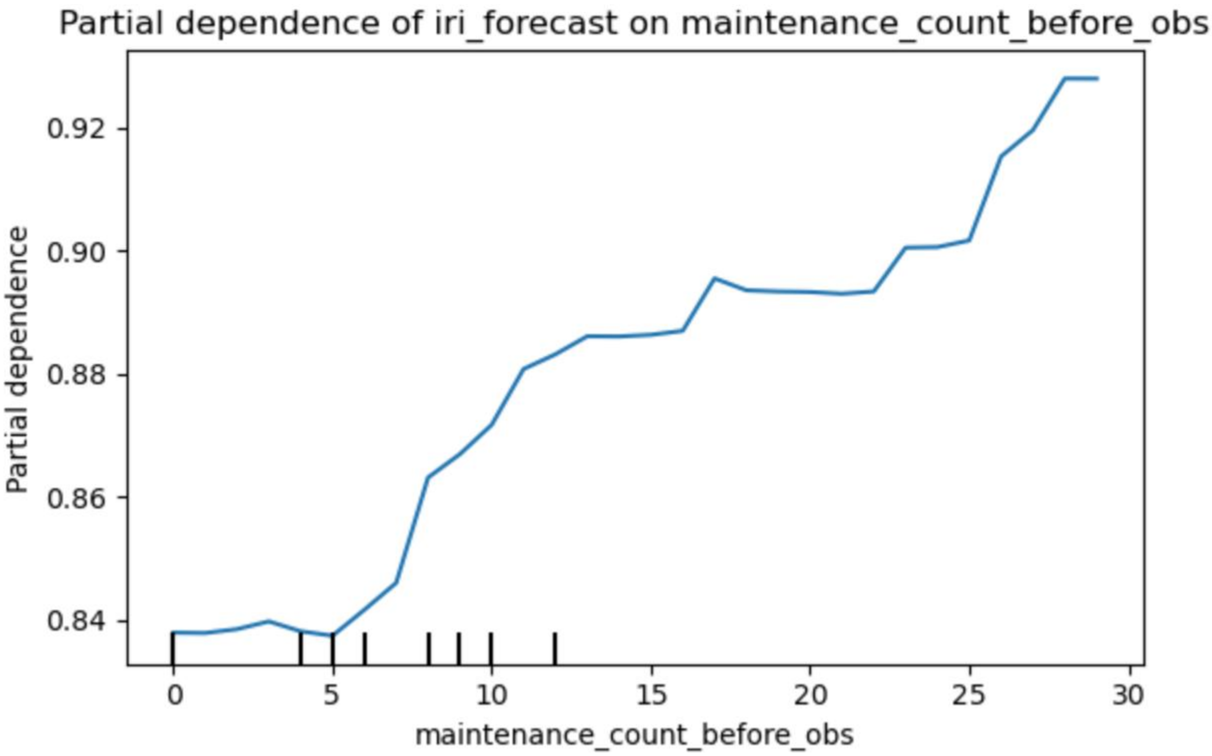

Figure 16. PDP Analysis of Maintenance count prior to observation on predicted IRI.

Maintenance is know as a reactive approach. The maintenance count prior to observation represents the number of maintenance activities carried out on a road section before the recorded IRI measurement. A higher maintenance count generally reflects sections that have experienced more frequent interventions, which is typically associated with poorer pavement condition, ageing infrastructure, or accumulated structural and functional deterioration. Road sections that deteriorate more rapidly tend to receive repeated interventions over time, resulting in higher maintenance counts. At the same time, delayed maintenance can accelerate pavement deterioration, leading to higher IRI values as distress accumulates over time. Therefore, both insufficient and delayed maintenance can contribute to increased roughness.

The observed positive association between maintenance intervention frequency and predicted IRI reflect the historical distress of the selected individual road sections, rather than implying maintenance activity itself drive road roughness progression. Instead, it indicates that more severely deteriorated road sections both require more frequent maintenance and exhibit higher roughness. The exact magnitude of this relationship may also depend on the type, timing, and effectiveness of maintenance actions applied.

4.5.5 PDP Analysis of Effect of Maintenance Frequency

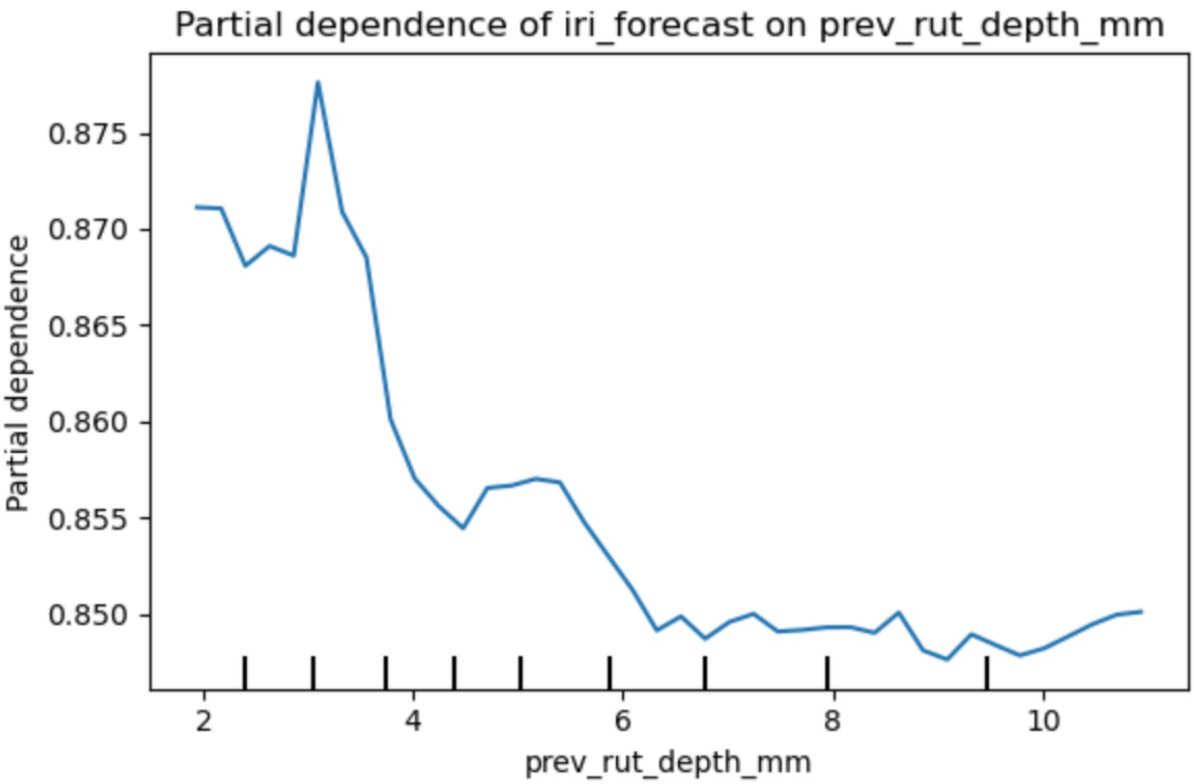

Figure 17. Partial Dependence Plot Analysis of Rut Depth on predicted IRI.

Although increasing rut depth is generally expected to increase pavement roughness, the Partial Dependence Plot exhibited an initially high predicted IRI followed by a decreasing trend. This behaviour is not fully consistent with direct engineering intuition, where higher rutting is typically associated with increased surface roughness. However, this does not indicate a modelling error; rather, it reflects that the Random Forest model captures complex interactions between rut depth and other explanatory variables in the dataset. In practice, IRI is influenced by multiple pavement distress types, including surface cracking, patching, structural deterioration, and longitudinal irregularities, which may not be directly correlated with rutting alone. As a result, sections with low rut depth may still exhibit high IRI due to other dominant distress mechanisms, while sections with higher rut depth may belong to pavement groups with different structural or maintenance characteristics. Therefore, the observed nonlinear and non-monotonic behaviour indicates that the model is not isolating rut depth as a single causal driver but instead capturing the combined effect of multiple interacting pavement condition variables. This underscore the inherent complexity of road surface degradation dynamics and illustrated capacity of the Random Forest algorithm to extract and represent the complex, multivariable nonlinear patterns embedded within the dataset.

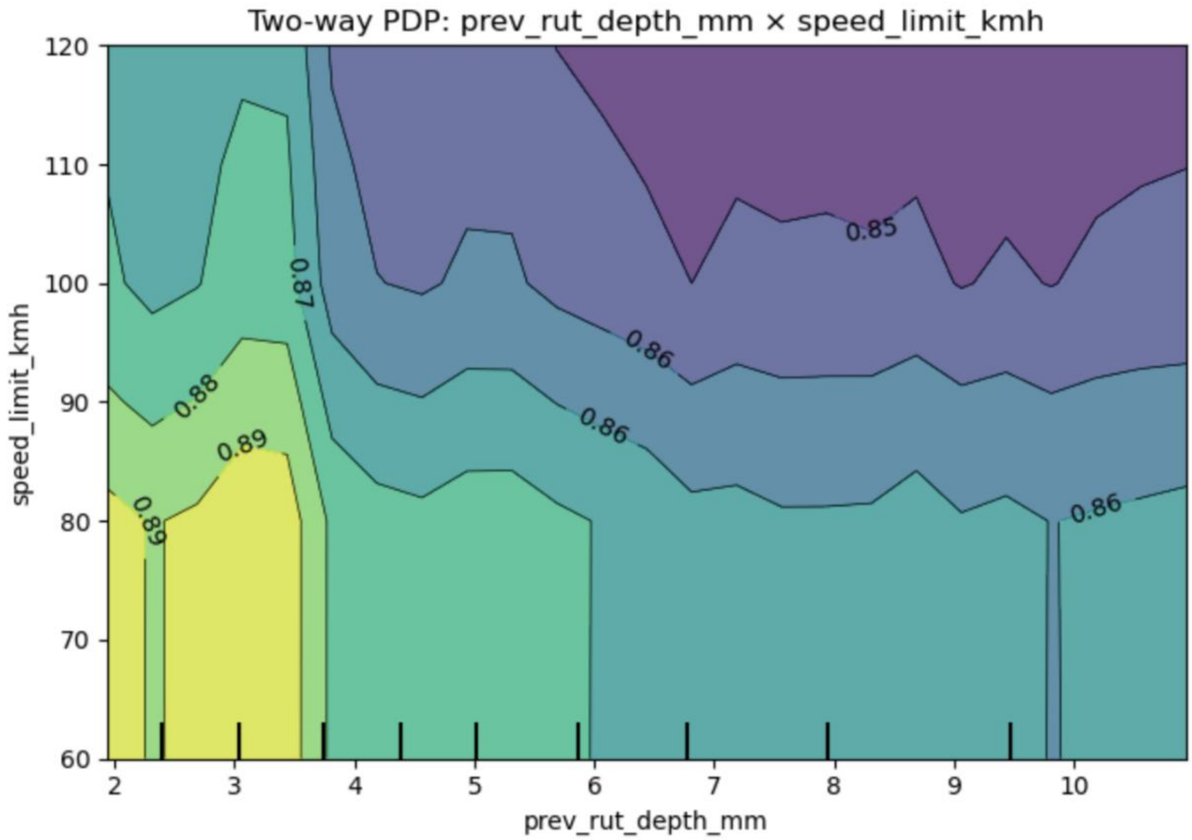

Figure 18. Two-Way PDP Showing the Combined Effect of Previous Rut Depth and Speed on Predicted IRI.

Rut depth values were also associated with increased predicted roughness, reflecting the interaction between surface deformation and ride quality deterioration. Traffic loading represented by AADT showed a positive relationship with IRI progression, although the relationship was moderately nonlinear. This suggests that heavy traffic exposure contributes to pavement deterioration, particularly when combined with existing structural weakness. Pavement thickness exhibited a stabilizing effect, where thicker pavement structures generally reduced predicted roughness progression. This observation is consistent with pavement engineering principles, since thicker structural layers improve load distribution and fatigue resistance.

4.6 Synthesis and Discussion

The The results presented in this chapter collectively demonstrate that pavement roughness progression on the E4 highway is governed by a hierarchy of interacting factors that cannot be adequately represented under linear modelling assumptions. The performance differential between Random Forest ($R^2 = 0.70$) and Ridge Regression ($R^2 = 0.47$) offers concrete empirical support for the importance of nonlinear modelling approaches in this context, a conclusion that aligns with results documented across the international pavement performance forecasting mentioned earlier in the literature chapter of this study (Gong et al., 2018; Marcelino et al., 2019; Karballaeezadeh et al., 2020).

The dominance of previous IRI in both feature importance analyses and the clearly nonlinear, accelerating PDP shape together provide a mechanistically coherent picture of pavement deterioration as a cumulative, self-reinforcing process. This finding has a direct and actionable implication for maintenance planning: network screening should prioritise segments exhibiting elevated recent IRI measurements, as these are most likely to exhibit continued and accelerating deterioration in the near term. The substantial importance of days since last maintenance reinforces the value of timely, proactive intervention: the longer a segment remains without maintenance, the greater the predicted roughness accumulation, consistent with the established relationship between maintenance deferral and accelerated deterioration costs.

The measurable contribution of climatic variables — collectively accounting for approximately 7% of MDI — confirms that IRI prediction models for the Swedish road network should incorporate environmental exposure indicators, particularly freeze–thaw cycle frequency, total precipitation, and thermal stress. The inclusion of these variables improves model accuracy and enhances the physical interpretability of the predictive framework, supporting its relevance for pavement management in Nordic cold-climate conditions where freeze–thaw cycling is a primary driver of surface cracking and structural degradation.

Finally, the comparison between MDI and permutation importance provides assurance that the identified feature hierarchy is not an artefact of the MDI estimation procedure but reflects genuine predictive relationships that generalise to unseen data. The consistency of rankings across both methods, combined with the physically interpretable PDP shapes, supports confidence in the engineering relevance of the model's learned representations and in its practical utility as a decision-support tool for pavement maintenance planning on the Swedish national road network.

5. Conclusions, Limitations, and Recommendations

5.1 Conclusions

The present research is designed and assess a series of machine learning frameworks aimed at forecasting the International Roughness Index on the E4 highway in Sweden, integrating pavement structural, traffic, maintenance history, and Nordic climatic variables from the PMS4 and VVIS databases. The following principal conclusions are drawn:

- Random Forest demonstrated significantly superior predictive performance over Ridge Regression, achieving cross-validated $R^2 = 0.70$, RMSE = 0.13 m/km, and MAE = 0.078 m/km, compared with $R^2 = 0.47$, RMSE = 0.18 m/km, and MAE = 0.121 m/km for the linear baseline. The performance gap confirms that pavement deterioration on the E4 corridor is characterised by complex nonlinear relationships that linear modelling frameworks cannot adequately represent.
- Previous IRI is the dominant predictor of future roughness, accounting for approximately 46% of total feature importance (MDI). This finding confirms that pavement roughness progression is fundamentally a cumulative, path-dependent process, in which existing condition is the primary determinant of future performance.
- Maintenance history — captured through days since last maintenance and cumulative maintenance count — is the second most influential factor group, underscoring the critical role of timely, proactive maintenance intervention in controlling roughness development on the network.
- Traffic loading (AADT), pavement structural characteristics (thickness), and Nordic climatic variables (total rainfall, thermal stress, freeze–thaw day frequency) all contribute meaningfully to model performance, confirming that IRI prediction models for Swedish conditions must integrate environmental exposure indicators alongside structural and traffic variables.
- The Random Forest model demonstrates stable and well-calibrated predictive performance across the observed IRI range, with symmetric, near-zero residuals and no evidence of systematic bias or heteroscedasticity, supporting its suitability for network-level pavement management applications.

Collectively, these findings demonstrate that Random Forest constitutes a robust, practically applicable, and physically interpretable framework for IRI prediction on the Swedish national road network, with direct potential to support data-driven maintenance prioritisation, infrastructure planning, and long-term asset management.

5.2 Limitations

The following limitations should be considered when interpreting the findings of this study:

- Temporal scope: the modelling dataset spans four observation years (2022–2025). This constrains the model's representation of long-term deterioration dynamics, particularly for deterioration mechanisms that operate over multi-year cycles, such as progressive subgrade weakening or slow-developing fatigue cracking.
- Data sparsity at distribution extremes: some predictor variables contain relatively few observations at the extremes of their observed ranges. This introduces uncertainty in the interpretation of PDP shapes at these extremes and may contribute to local overfitting artefacts — as observed in the AADT PDP at high traffic volumes — that do not reflect genuine physical relationships.
- Data-driven, non-mechanistic modelling: as a statistical machine learning method, Random Forest does not encode explicit physical deterioration mechanisms. The identified predictor–response relationships represent statistical associations observed in the training data and should not be interpreted as direct causal relationships without supporting mechanistic evidence.
- Long-range forecasting uncertainty: IRI forecasts for the period 2026–2030 are conditional on assumed future trajectories of traffic volumes, climatic conditions, and maintenance activities. Since these trajectories are inherently uncertain, long-range predictions should be treated as scenario-based estimates rather than deterministic forecasts and should be updated as new observational data become available.
- Single-corridor dataset: the scope of this investigation is confined to road section along E4 highway. While this corridor spans diverse administrative regions and climatic sub-zones, the generalisability of the model to other Swedish highway corridors — particularly those with different structural design standards, traffic compositions, or climatic regimes — has not been formally validated.

5.3 Recommendations for Future Research

The following directions are recommended for future research building on the findings of this study:

- Dimensionality management: the integration of additional candidate predictors — including subgrade bearing capacity indicators, pavement age, and surface texture measurements — is recommended to broaden the explanatory scope of the model. Principal Component Analysis (PCA) or regularised feature selection methods should be applied to manage the resulting increase in dimensionality and to control multicollinearity among expanded predictor sets.

- Long-range forecasting with scenario inputs: integration of projected traffic growth scenarios from Trafikverket's national traffic forecasts, regional climate projections, and planned maintenance schedules would substantially improve the credibility and operational utility of long-range IRI predictions for the 2026–2030 period and beyond.
- Algorithm benchmarking: intensive benchmarking exercise evaluating Random Forest alongside gradient boosting variants such as XGBoost and others, as well as the application of deep learning architectures, applied to similar datasets under concrete cross-validation protocols, could clarify weather additional acquisition in predictive accuracy are attainable and yield more definitive algorithm guidance for this category of forecasting challenges.
- Multi-corridor and multi-network validation: applying the developed methodology to additional Swedish highway corridors and, potentially, to road networks in other Nordic countries would test the generalisability of the findings and contribute to the development of transferable, climate-adapted IRI prediction frameworks for cold-climate environments.
- Integration with PMS4 decision support: embedding the Random Forest model within the existing PMS4 decision-support workflow — providing segment-level IRI forecasts alongside uncertainty bounds — would represent a practical pathway to operational deployment and would directly support evidence-based maintenance prioritisation across the Swedish national road network.

6. AI Use Disclosure

Artificial intelligence (AI) tools, including ChatGPT (OpenAI) and Claude (Anthropic), were used during the preparation of this thesis as research support and productivity tools. These tools assisted with language refinement, code debugging, explanation of machine learning concepts, generation of presentation materials, and review of draft text.

All research design decisions, data processing, model development, analysis, interpretation of results, and final conclusions were conducted and verified by the author. AI-generated content was critically reviewed, edited, and validated before inclusion in the thesis.

The author accepts full responsibility for the accuracy, originality, and integrity of the work presented in this thesis.

7. Bibliography

Abdelaziz, N., Abd El-Hakim, R. T., El-Badawy, S. M., & Afify, H. A. (2020). International Roughness Index prediction model for flexible pavements. International Journal of Pavement Engineering, 21(1), 88–99. https://doi.org/10.1080/10298436.2018.1441414

Albuquerque, F. S., & Núñez, W. P. (2011). Development of roughness prediction models for low-volume road networks in northeast Brazil. Transportation Research Record, 2205(1), 198–205. https://doi.org/10.3141/2205-25

Andrén, P., Eriksson, O., & Lundberg, T. (2014). Prognostic models for the road condition indices in the Swedish Transport Administration's pavement management system: IRI and rut depth (VTI Report 812). Swedish National Road and Transport Research Institute.

Braspenning, P., Thuijsman, F., & Weijters, A. (1995). Artificial Neural Networks: An Introduction to Artificial Neural Networks and Practice. Springer Verlag Berlin Heidelberg.

Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5–32.

Elhadidy, A. A., El-Badawy, S. M., & Elbeltagi, E. E. (2021). A simplified pavement condition index regression model for pavement evaluation. International Journal of Pavement Engineering, 22(5), 643–652. https://doi.org/10.1080/10298436.2019.1633579

Gong, H., Sun, Y., Shu, X., & Huang, B. (2018). Use of random forests regression for predicting IRI of asphalt pavements. Construction and Building Materials, 189, 890–897.

Hossain, I., Gopisetti, L. S. P., & Miah, M. S. (2018). International Roughness Index prediction of flexible pavements using neural networks. Journal of Transportation Engineering, Part B: Pavements, 145(1). https://ascelibrary.org/doi/10.1061/JPEODX.0000088

Karballaeezadeh, N., Mohammadzadeh, S., Moazemi, D., Band, S. S., Mosavi, A., & Reuter, U. (2020). Smart structural health monitoring of flexible pavements using machine learning methods. Coatings, 10(11), 1100. https://doi.org/10.3390/coatings10111100

Lin, J. D., Yau, J. T., & Hsiao, L. H. (2003). Correlation analysis between International Roughness Index and pavement distress by neural network. Proceedings of the 82nd Annual Meeting of the Transportation Research Board, Washington D.C.

Lundberg, T., & Eriksson, O. (2019). Development of pavement condition models for Swedish roads. VTI Report. Retrieved from https://vti.diva-portal.org/smash/record.jsf?pid=diva2%3A1313184

Lundberg, T., & Thomas, P. (2010). Pavement management in Sweden: Methods and models. VTI Report. Retrieved from https://vti.diva-portal.org/smash/record.jsf?pid=diva2%3A848534

Mactutis, J. A., Alavi, S. H., & Ott, W. C. (2000). Investigation of relationship between roughness and pavement surface distress based on WesTrack project. Transportation Research Record, 1699(1), 107–113. https://doi.org/10.3141/1699-15

Marcelino, P., Antunes, M. D. L., Fortunato, E., & Gomes, M. C. (2019). Machine learning approach for pavement performance prediction. International Journal of Pavement Engineering, 22(3), 341–354. https://doi.org/10.1080/10298436.2019.1609673

Meshram, S. G., Singh, V. P., & Kisi, O. (2020). Application of artificial neural networks, support vector machine and multiple model-ANN to sediment yield prediction. Water Resources Management, 34, 4561–4575. https://doi.org/10.1007/s11269-020-02672-8

PMS4. (n.d.). Trafikverket Pavement Management System. Retrieved from https://pmsv4.trafikverket.se

Rita, R., Ferreira, A., & Flintsch, G. (2021). Sustainability of road pavements. Sustainability, 13(9), 5248. https://www.mdpi.com/2071-1050/13/9/5248

Sayers, M. W., Gillespie, T. D., & Queiroz, C. A. (1986). The International Road Roughness Experiment (IRRE): Establishing correlation and a calibration standard for measurements. World Bank Technical Paper 45. Retrieved from https://documents.worldbank.org

Tamagusko, T., & Ferreira, A. (2023). Machine learning for prediction of the International Roughness Index on flexible pavements: A review, challenges, and future directions. Infrastructures, 8, 170. https://doi.org/10.3390/infrastructures8120170

Trafikverket. (2011). Underhållsstandard belagd väg (Publ. 2012:074). Retrieved from https://trafikverket.diva-portal.org

Trafikverket. (2025). National Transport Infrastructure Plan 2026–2037. Retrieved from https://bransch.trafikverket.se/en/startpage/planning/long-term-infrastructure-planning/

VTI. (n.d.). Road surface measurements. Retrieved from https://www.vti.se/en/services/road-measurements/measurement-of-road-surface

Zeiada, W., Dabous, S. A., Hamad, K., Al-Ruzouq, R., & Khalil, M. A. (2020). Machine learning for pavement performance modelling in warm climate regions. Arabian Journal for Science and Engineering, 45, 4091–4109.

A. APPENDIX A

A.1 Overview

This appendix documents the technical implementation of the predictive algrothims developed to forecast future road surface roughness in this study, expressed as IRI the evaluation contain two modelling approaches:

- Ridge Regression (baseline model)
- Random Forest Regression (primary model)

The models were developed using Python and the Scikit-learn library. Input variables were selected to capture the influence of road surface condition, vehicle loading, intervention history, pavement structure properties, and climatic exposure

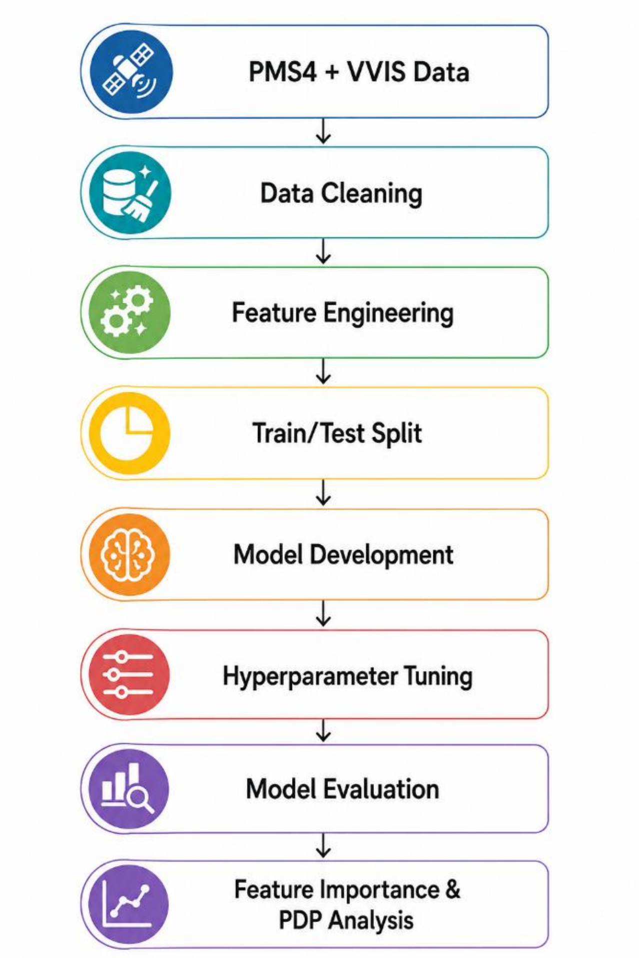

Figure 19. Overall modelling process.

A.2 Predictor Variables

Table 4: Predictor variables used for IRI prediction.

Variable	Unit
Previous IRI	mm/m
Previous Rut Depth	mm
AADT	vehicles/day
Days Since Last Maintenance	days
Pavement Thickness	mm
Speed Limit	km/h
Road Width	m
Total Rainfall	mm
Freeze–Thaw Days	days
Thermal Stress	°C
Days Since Previous IRI Observation	days
Maintenance Count Before Observation	count

A.3 Ridge Regression Model

Ridge Ridge Regression was selected as the baseline model due to its ability to provide a simple and interpretable representation of linear relationships between predictors and future pavement roughness.

$$\min \beta \ \Sigma_i (y_i - \hat{y}_i)^2 + \alpha \ \Sigma_j \beta_j^2$$

The Python implementation is as follows:

```
ridge_pipeline = Pipeline([("preprocessor", preprocessor), ("model", Ridge(alpha=1.0))])
```

A.4 Random Forest Model

Random Forest Model was adopted as the principle predictive framework on the basis of its demonstrated capacity to represent complex, nonlinear interdependence among traffic, structural, maintenance, and climatic variables. The implemented model parameters are specified below:

```
rf_model = RandomForestRegressor(n_estimators=500, max_depth=10,
min_samples_leaf=2, random_state=42)
```

A.5 Hyperparameter Optimization

Random Forest hyperparameters outlined in table 5 were optimised using Randomised Search Cross-Validation to improve predictive performance and reduce overfitting.

Table 5: Hyperparameter optimisation search space.

Parameter	Range
Number of Trees	250–700
Maximum Depth	6–20
Minimum Samples Split	2–20
Minimum Samples Leaf	1–8
Maximum Features	sqrt, 0.4–1.0

Five-fold cross-validation was used to identify the optimal model configuration using the following implementation:

rf_random_search = RandomizedSearchCV(estimator=rf_base_pipeline, param_distributions=param_distributions, n_iter=30, cv=5, scoring="neg_root_mean_squared_error")

A.6 Model Evaluation

Model performance was evaluated using R^2, RMSE, and MAE, implemented as follows:

r2 = r2_score(y_test, y_pred) rmse = np.sqrt(mean_squared_error(y_test, y_pred)) mae = mean_absolute_error(y_test, y_pred)

A.7 Feature Importance Analysis

Feature importance was computed using the Random Forest model's internal impurity-based metric (MDI) to identify the variables contributing most strongly to prediction performance:

feature_importance = pd.DataFrame({"Feature": feature_names, "Importance": rf_model.feature_importances_})

Feature importance results indicated that previous IRI was the most influential predictor, followed by maintenance history, rut depth, traffic loading, and pavement thickness.

A.8 Supplementary Material

The complete Python implementation, including data preprocessing, feature engineering, model development, etc is provided as supplementary material accompanying this thesis.

DEPARTMENT OF ARCHITECT AND CIVIL ENGINEERING
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2026
www.chalmers.se

CHALMERS
UNIVERSITY OF TECHNOLOGY