

Aalto University

MS-E2177 - Seminar on Case Studies in Operations Research

Modelling of Paved Road Deterioration

Interim Report

Team Members:

Jonni Järvinen (Project manager)

Markus Junttila

Veera Wilkki

Oskar Heikkilä

Konsta Lahtinen

Returned: June 22, 2026

Contents

1	Changes in objectives/scope	3
2	Project status	4
3	Changes in project plan	4
4	Risk management plan	4

1 Changes in objectives/scope

Previously, it was unclear how the time frame for the prediction was to be implemented. The initial idea was to have a constant time frame such as 4 years for each prediction. Currently, the objective is to implement the time frame as an input parameter.

The initial idea was to predict both roughness (IRI) and rutting (URA). However, during preprocessing the roughness index was found to behave quite unpredictably both with and without logged maintenance events. After discussing the topic with the FTIA, it was decided that we should prioritize building a model that predicts rutting accurately.

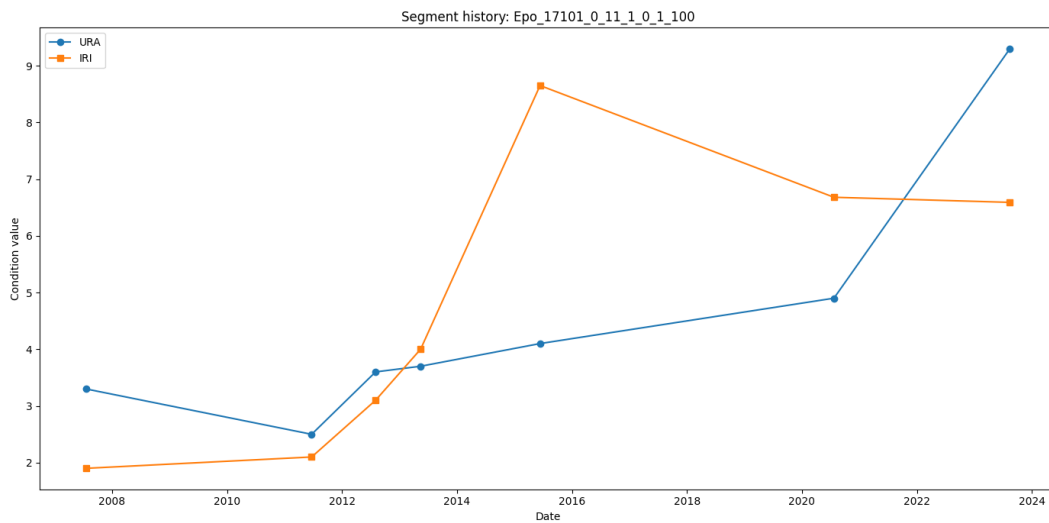


Figure 1: Picture illustrating the inconsistency of IRI. Smaller values of IRI indicate less roughness.

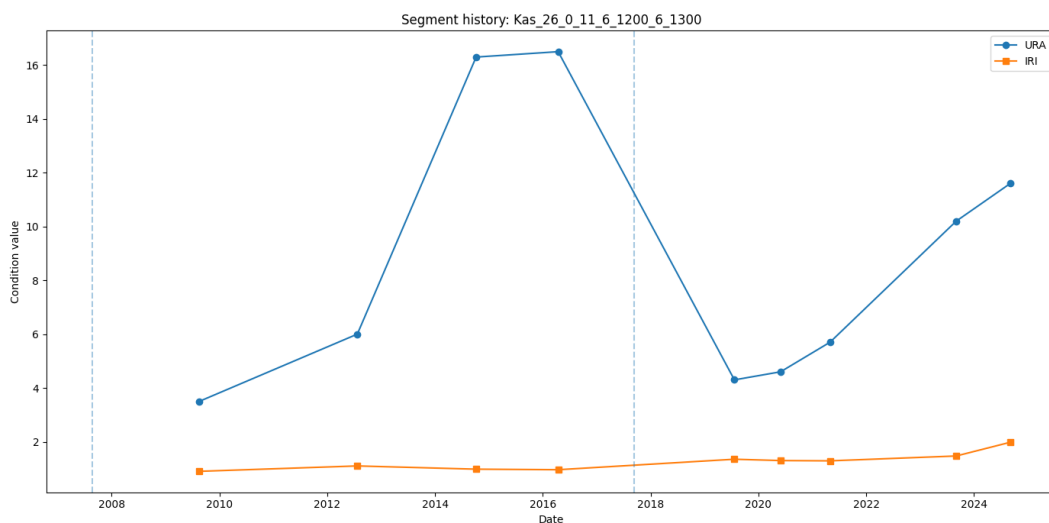


Figure 2: Second picture illustrating the inconsistency of IRI

Figures 1 and 2 show the evolution of URA (in blue) and IRI (in orange) between the years 2008 and 2024 of two road segments. The vertical lines represent maintenance events. From the figures it is evident that roughness cannot be predicted from maintenance events alone. The figure illustrates IRI decreasing without a maintenance event and increasing when there has been a maintenance event between the measurements.

2 Project status

During the project we have had multiple discussions with FTIA about the scope and objectives. We have also conducted a literature review as well as cleaned and preprocessed the raw data after performing exploratory data analysis.

Currently we are testing different machine learning methods to discover the most fitting one. We are also currently in the feature engineering phase to suit the model. We are experimenting with different feature sets to see, for example, whether to use static features, lifecycle features, or both. We are also discussing different ways to fine-tune the data preprocessing such as the URA cut-off value, i.e., the amount URA has to increase for the model to consider that an unrecorded maintenance event has taken place between the measurements.

3 Changes in project plan

The most significant change to the project plan is that climate data will no longer be included in the modelling work. In the initial plan, climate-related variables were considered as one possible group of explanatory features. However, as the project has progressed, it has become clear that adding climate data would increase the complexity of the work considerably while offering limited benefit for the main prediction task.

Related to this, we will no longer aim to assess whether climate change has had an observable effect on road deterioration. Although this was mentioned in the original objectives, we now consider that such an analysis would not offer reliable results within the scope of this project. The available measurement period of roughly twenty years is likely too short for this kind of analysis.

4 Risk management plan

The main risks assessed for this project are in Table 1. The risks and their effects are defined, and their probabilities and impacts are estimated on a discrete scale [Low, Medium, High]. The last column shows how to mitigate the risks. Updates made for the interim report are shown in green, while the original risk levels are indicated in parentheses. The probability of data quality issues was reduced from High to Low due to improved data preprocessing and validation. The other risks were not revised, as their underlying conditions remain unchanged.

Table 1: Main risks related to the project. Initial values are in parenthesis and updated values are highlighted in green.

Risk	Effect	Probability	Impact	Mitigation
Data quality inconsistencies	Measurement errors or missing maintenance records may bias model training and reduce reliability of results.	Low (High)	High	Preprocessing; detect unrealistic jumps; different data validation methods.
Time limitations	Incomplete implementation or limited experimentation.	Medium	High	Divide tasks clearly; set weekly checkpoints.
Model performance remains insufficient with available data	Despite preprocessing and model tuning, the available data may not contain enough signal to build a reliable predictive model.	Medium	High	Perform exploratory data analysis early; test baseline models first; document limitations clearly; consider redefining target or narrowing scope if needed.
Incorrect problem formulation	Prediction target or time horizon may not align with project objectives.	Low	High	Define target and horizon clearly.
Team coordination issues	Scheduling conflicts, illness, or uneven workload distribution may delay progress and reduce overall project quality.	Medium	Medium	Agree on clear roles and responsibilities; maintain regular communication and backup plans for key tasks.
Overfitting of machine learning models	Model performs well on training data but poorly on unseen data.	Medium	Medium	Use cross-validation; start from basic models before complex models; avoid "black box" models.