Data Fusion for Predicting Highway Maintenance and Deterioration Trends

Chandra Sekhar Katipalli
Sindura Reddy Challa
Sanjana Reddy Soma

University of Maryland, Baltimore County

DATA 606: Capstone in Data Science

Dr. Masoud Soroush

Data Fusion for Predicting Highway Maintenance and Deterioration Trends

Abstract: This study on Highway Maintenance and Deterioration Trends investigates the potential of data fusion to enhance highway deterioration prediction by integrating the Highway Performance Monitoring System (HPMS) and Freight Analysis Framework (FAF) datasets from the year 2013 to 2022. Highway deterioration is measured by the International Roughness Index (IRI) and this often poses significant challenges to maintenance planning and infrastructure management.

The central research question is: "What are the benefits of data fusion in predicting highway deterioration compared to traditional methods?" We hypothesize that integrating these datasets will provide a more comprehensive view of influencing factors ranging from traffic volume to freight movement and thus enhance prediction accuracy. The methodology includes rigorous data cleaning, normalization, and feature engineering followed by the application of machine learning algorithms, including regression analysis and forecasting, to predict future pavement conditions.

1. Introduction

The deteriorating condition of highway infrastructures poses severe economic and safety challenges. As transportation networks become increasingly burdened by higher traffic and freight demands. Conventional methods often rely on isolated datasets, limiting their predictive power. This research leverages data fusion between Highway Performance Monitoring System (HPMS) and Freight Analysis Framework (FAF) data to generate a robust deterioration model. Such an approach is timely and crucial, as underscored by emerging research trends [1][2].

2. Background

Historically, highway deterioration research has focused on single data streams, such as pavement condition indices derived from Highway Performance Monitoring System (HPMS). However, recent investigations suggest that incorporating pavement structures and environmental factors from Long-Term Pavement Performance (LTPP) can reveal additional insights into pavement performance [3].

Building on this foundation, recent research emphasizes the transformative role of predictive models in highway deterioration analysis. As highlighted in [2], predictive models offer a proactive approach that forecasts future pavement conditions by integrating diverse data streams. Unlike detection models, which only identify deterioration after it has occurred, predictive models facilitate preventive maintenance by pinpointing potential risk factors before they culminate in significant damage. This foresight allows transportation agencies to allocate resources more efficiently, mitigate safety hazards, and extend the service life of infrastructure [4].

3. Summary of Literature Review

The literature reveals that traditional highway deterioration models often depend on isolated datasets, such as the HPMS, which limits their ability to account for the complex interactions between traffic, freight, and environmental factors. Pioneering research has suggested that including data from sources like LTPP provides additional context by incorporating variables related to pavement structures and environmental conditions [3]. Additionally, the emergence of predictive models has marked a significant shift from reactive detection approaches toward proactive maintenance strategies. Predictive methods are inherently preventive, they anticipate deterioration and help inform timely interventions before severe degradation occurs [2]. Our proposed research aims to extend this body of work by fusing HPMS with FAF data, thus capturing critical variables such as freight dynamics that have been largely overlooked. This integrated approach is expected to refine deterioration forecasts, enhance the precision of maintenance planning, and ultimately bridge the gap between traditional, detection-based models and advanced predictive analytics.

4. Research Questions

In this study, we will develop a robust predictive model by integrating HPMS data from the Federal Highway Administration with FAF data from the Bureau of Transportation Statistics, covering the period from 2013 to 2022, to improve highway deterioration forecasts based on the International Roughness Index (IRI). Our methodology involves rigorous data cleaning, normalization, and feature engineering to extract key variables, such as traffic volume, freight movement, and pavement condition indices. By leveraging advanced machine learning techniques, we aim to quantitatively assess how these factors influence pavement performance and optimize maintenance strategies. The following research questions guide our investigation:

1. How does the fusion of HPMS and FAF datasets (2013–2022) enhance the predictive performance of highway deterioration models in estimating IRI, compared to traditional statistical and machine learning approaches using single-source data?
2. What are the most influential predictive features—such as traffic volume, freight load, and pavement condition indices—derived from the integrated datasets, and how do their contributions vary across different machine learning models?
3. How effectively can the proposed predictive model, leveraging data fusion and advanced machine learning techniques, minimize forecasting errors and improve the optimization of maintenance scheduling to reduce unplanned highway repairs?

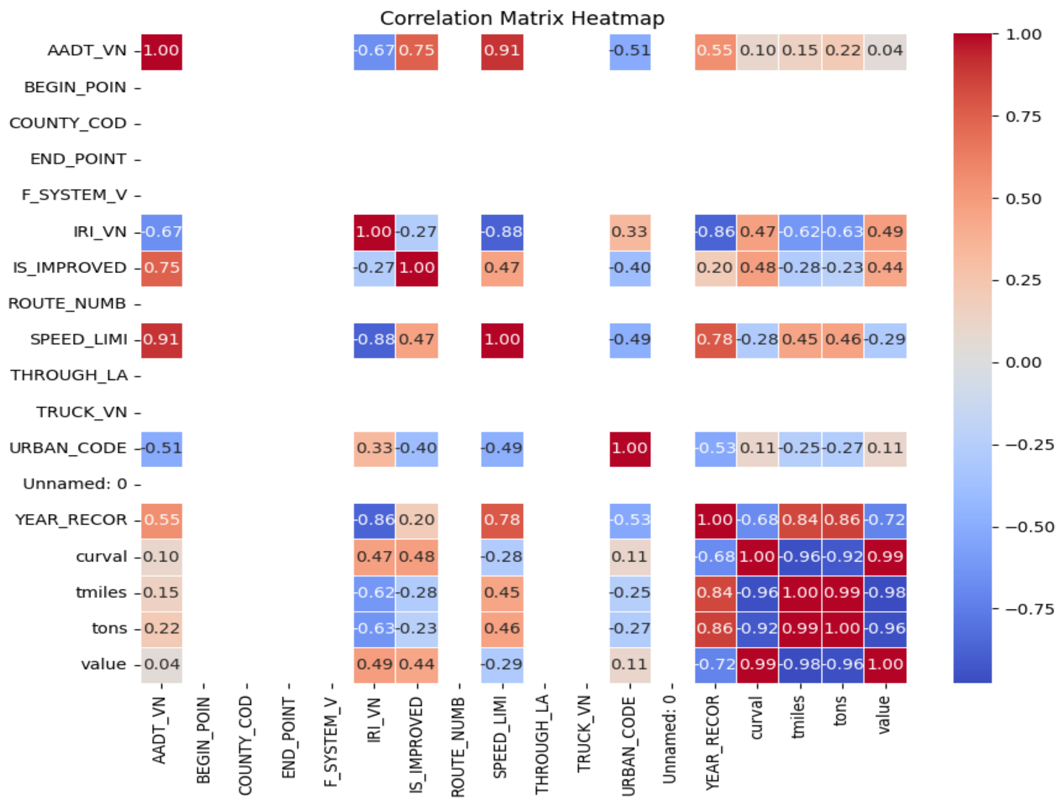
4. What is the optimal approach to forecasting IRI at different levels of granularity—both for entire highway routes (RouteID level) and for specific highway sections (0.1-mile segments)—to support more precise maintenance planning?
5. What might be the most effective method for visualizing and presenting findings to highway maintenance teams, like using geospatial mapping to find the roughest sections along a highway and their projected deterioration over time?

By systematically addressing the above listed research questions, we aim to develop an advanced, practical, and predictive model that allows transportation agencies to make data-driven decisions for proactive highway maintenance.

5. Preliminary Implications

Preliminary implications of this research are highly promising, as the integration of HPMS and FAF datasets is expected to significantly enhance our understanding of highway deterioration. By combining these diverse data sources, our study aims to extend current knowledge by providing a more holistic view of the factors affecting pavement performance.

Our initial analysis revealed meaningful correlations that underscore the potential impact of our approach. Specifically, the International Roughness Index (IRI) shows a strong correlation with speed limits and year, telling that roadway design parameters and the age of infrastructure play a critical role in pavement performance. Additionally, moderate correlations with average annual daily traffic(AADT) and freight tons highlights the influence of traffic volume and freight movement on deterioration. These findings not only validate the relevance of our data fusion methodology but also provide a clear direction for future research. Subsequent studies can build on these correlations to refine predictive models further, explore causal relationships, and adapt strategies to different regions, ultimately advancing the field of predictive maintenance in highway infrastructure.



6. Conclusion

In conclusion, this research proposal presents a novel approach to highway deterioration prediction by integrating HPMS and FAF datasets from 2013 to 2022, thereby addressing critical gaps in current maintenance planning methodologies. By employing advanced machine learning techniques to analyze key variables ranging from traffic volume and freight movement to roadway design parameters the proposed model moves beyond traditional detection methods to enable proactive, data-driven decision-making. The preliminary analysis, demonstrating significant correlations between IRI and factors such as speed limits and infrastructure age, reinforces the potential of this data fusion strategy to improve predictive accuracy and optimize resource allocation. Ultimately, the findings from this study are expected to not only enhance the precision of pavement condition forecasts but also pave the way for future research in predictive maintenance, contributing to more cost-effective and sustainable highway infrastructure management.

7. Reference

[1] (Yuanjiao Hu et al., 2022). Evaluation of pavement surface roughness performance under multi-features conditions based on optimized random forest. https://ieeexplore.ieee.org/document/9816255

[2] (Maher Mahmood et al., 2020). Multi-Types of Flexible Pavement Deterioration Prediction Models. https://ieeexplore.ieee.org/document/9122932/

[3] (Moein Latifi et al., 2021). A deep reinforcement learning model for predictive maintenance planning of road assets: Integrating LCA and LCCA. https://arxiv.org/abs/2112.12589

[4] (Yingjie Du et al., 2025). Enhancing Road Maintenance Through Cyber-Physical Integration: The LEE-YOLO Model for Drone-Assisted Pavement Crack Detection. https://ieeexplore.ieee.org/document/10902021