Beitbridge Border Traffic Volume Estimating Model

Shumel Siraj, Michael Schultz, Satya Prasad, Daniel Saslavsky

THE GEORGE WASHINGTON UNIVERSITY
WASHINGTON, DC

The World Bank

1. Introduction

The main goal of this project is to create a dependable and scalable system for predicting traffic at international border crossings, with a specific emphasis on monitoring traffic from Zimbabwe to South Africa, particularly at the Beitbridge border.

This research will focus on key border crossings in Zimbabwe and South Africa, with a particular emphasis on traffic heading towards Musina, South Africa, which originates from Beitbridge, Zimbabwe.

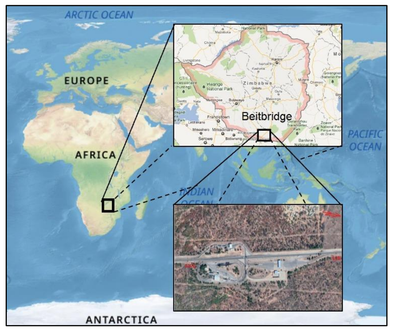


Fig 1.1 Beitbridge station 3201 And Beitbridge station 3202

Use Case:

- **Congestion Management:** Helps in identifying bottlenecks and high-traffic areas so that measures can be taken to alleviate congestion.
- **Urban Planning:** Provides data to inform the layout of roads, intersections, and public transportation.
- **Policy Decisions:** Helps in deciding fare rates, toll collections, and other transportation-related policies.
- **Resource Allocation:** Helps in effective deployment of emergency services and road repair teams.
- **Real-time Updates:** Used in GPS and navigation systems to provide real-time route recommendations.

2. Data

This model utilizes comprehensive traffic count data sourced from two toll road stations near the Beitbridge border. This data represents the most reliable approximation of actual traffic flow in the area, offering a near-ground truth for our analyses.

Key features of this Toll Road Station dataset include:

- **Data Coverage:** 01-01-2018 to 12-01-2022
- **Interval-Based Counts:** Traffic counts are recorded in three-hour intervals, providing a granular view of traffic flow dynamics throughout the day and night.
- **Station Specifics:** Beitbridge Station 3201 and Beitbridge Station 3202:
- **Vehicle Classification:** The traffic count data is segmented into various vehicle types for a more detailed analysis. These classifications include Light vehicles, as well as Short-, Medium-, and Long Heavy vehicles.

To enhance the richness of features, additional data has been incorporated.

- Stock data representing the economies of South Africa and Zimbabwe
- Crude oil price (OPEC)

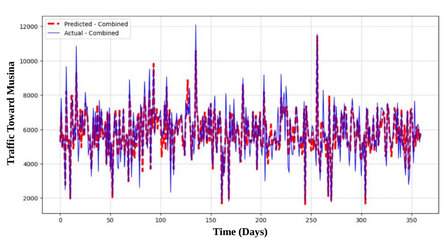


Fig 2.1 Time Series Of Actual VS Predicted

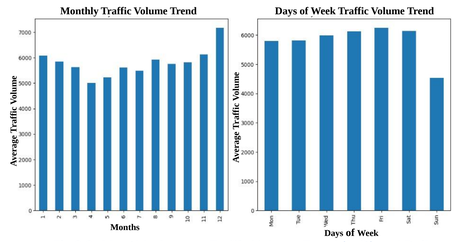


Fig 2.2: Monthly Average Traffic Volume Trend (Left) and Daily Average Traffic Volume Trend (Right)

3. Methodology

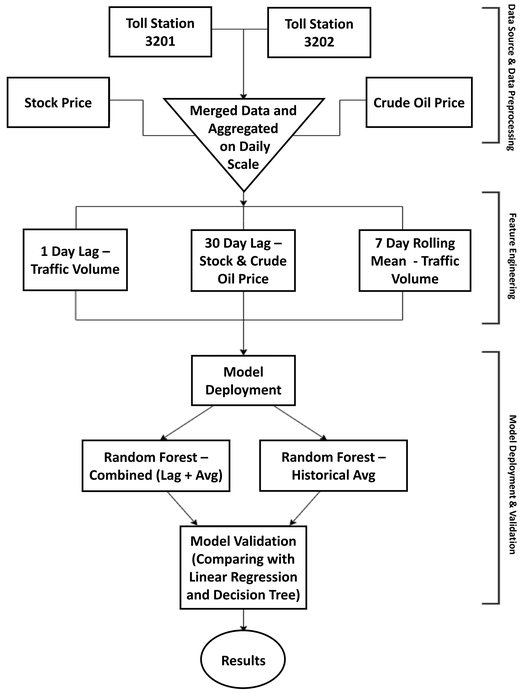


Fig 3.1 Methodological Flowchart

Data Collection: Traffic counts are gathered from Beitbridge TCC stations every 3 hours.

Data Integration: Traffic data is combined with daily stock and crude oil prices, providing a comprehensive overview that is aggregated daily.

Feature Engineering:

- **1-Day Lag for Traffic Volume:** Incorporates previous day's data to account for immediate temporal influences.
- **30-Day Lag for Economic Indicators:** Assesses the impact of stock and crude oil prices from the past month on current traffic trends.
- **7-Day Rolling Mean for Traffic Volume:** Applied to historical data to mitigate daily volatility and reveal underlying patterns.

Model Development and Validation:
Random Forest models were developed to predict traffic patterns using historical averages and time-lagged economic data. Their performance was benchmarked against Linear Regression and Decision Tree models to ensure accuracy and reliability.

4. Results

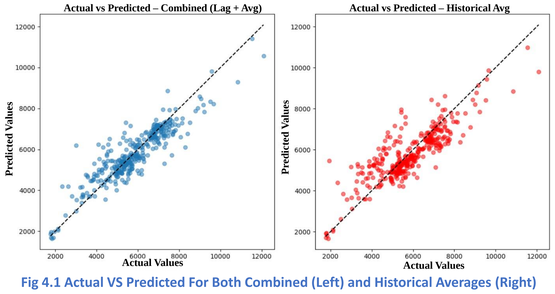


Fig 4.1 Actual VS Predicted For Both Combined (Left) and Historical Averages (Right)

Combined Model (Blue):
Demonstrates a close match between predicted and actual traffic volumes, signifying high accuracy.

Historical Averages Model (Red):
Shows a reasonable match but with more variance, particularly at higher traffic volumes, implying lower accuracy compared to the combined model.

Top 10 Features
To_Beit_Bridge_Total_3201_7_day_avg
To_Beit_Bridge_Light_3201_1day_lag
To_Beit_Bridge_Total_3201_1day_lag
To_Beit_Bridge_Total_3202_7_day_avg
To_Beit_Bridge_Light_3202_7_day_avg
To_Beit_Bridge_Short_HV_3201_1day_lag
To_Beit_Bridge_Light_3201_7_day_avg
To_Beit_Bridge_Medium_HV_3202_1day_lag
To_Beit_Bridge_Light_3202_1day_lag
To_Beit_Bridge_Total_3202_1day_lag

Fig 4.2 Top 10 Features

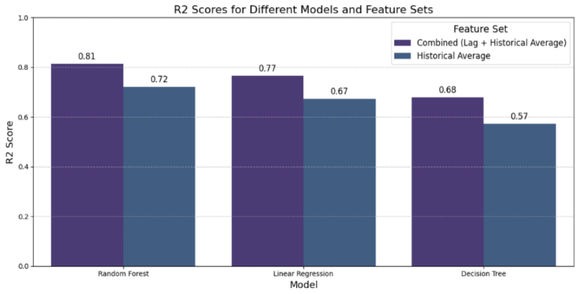


Fig 4.3 R2 Scores For Different Models For Both Combined Approach And Historical Averages

5. Conclusion

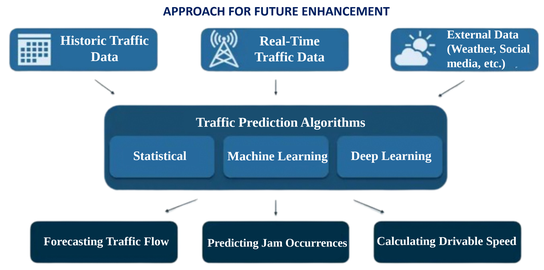
The Random Forest, combining lag features and historical averages, is the most accurate model for predicting traffic volumes at the Beitbridge border, with an R^2 score of 0.81. This model outperforms other methods, endorsing its use for efficient traffic management and planning.

6. Future Enhancement

In my future work, I aim to enhance traffic prediction accuracy and scope through:

Data Enrichment: We'll expand our data sources by:

- Leveraging historic traffic data for deeper temporal analysis.
- Integrating real-time traffic data for dynamic predictions.
- Incorporating external data like weather conditions, social media trends, and public events to account for a wider range of traffic flow influences.



Predictive Outcomes:

By combining these methods, we aim to enhance traffic flow prediction, considering historical trends and real-time data. This includes predicting traffic jams and calculating drivable speeds.

7. References

Drouyer, S., & de Franchis, C. (2019, July). Highway traffic monitoring on medium resolution satellite images. In *IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium* (pp. 1228-1231). IEEE.

Zhou, J., Gao, D., & Zhang, D. (2007). Moving vehicle detection for automatic traffic monitoring. *IEEE transactions on vehicular technology*, *56*(1), 51-59.

Zhu, Z., Xu, G., Yang, B., Shi, D., & Lin, X. (2000). VISATRAM: A real-time vision system for automatic traffic monitoring. *Image and Vision Computing*, *18*(10), 781-794.

Drouyer, S., & de Franchis, C. (2020). Parking occupancy estimation on sentinel-1 images. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *2*, 821-828.