A traffic simulation-informed explainable boosting machine model of emergency vehicle speeds

Farnoosh Namdarpour, Fan Zuo, Di Sha, Hannah Bonestroo, Jingqin Gao, Joseph Y. J. Chow, Kaan Ozbay
C2SMARTER Center, New York University Tandon School of Engineering
Brooklyn, NY, 11201, USA

ISTDM2025 – Montréal, QC, Canada

Introduction

Emergency vehicles (EMVs) include ambulances, fire trucks, and police cars, which respond to critical events such as medical emergencies, fire disasters, and public security crises. Emergency response time is the key indicator of a city's incidents management ability and resiliency. Reducing response time saves lives and prevents property losses. For instance, Berdowski et al. (2010) indicates that the survival rate from a sudden cardiac arrest without treatment drops 7%–10% for every minute elapsed, and there is barely any chance to survive after 8 min. EMV travel time, the time interval for an EMV to travel from a rescue station to an incident site, accounts for a major portion of the emergency response time.

However, overpopulation and urbanization have exacerbated road congestion, making it more challenging to reduce the average EMV travel time. Records (NYC, 2023) show that the average EMV response time in NYC increased across the board for all types of EMVs from NYPD, FDNY, and EMS. For example, medical emergency response times handled by FDNY increased from 10.39 minutes to 13.29 minutes, of which most of the increase is due to travel time in traffic (from 4.43 min to 6.08 min). This is a 37% increase in travel time in a matter of 8 years, or an equivalent annual growth rate of 4% per year.

Given these concerns from increasing congestion, agencies like the New York City Fire Department (FDNY) need to evaluate effective interventions to improve response times. Evaluation can be prohibitive in field conditions where wrong decisions can impact lives. EMVs behave differently from passenger vehicles in the traffic; they can gain right of way through signalized intersections, their siren technologies can inform downstream vehicles to part for them to pass, etc. As such, typical navigation tools for predicting travel times like Google Maps may not be helpful.

A prediction model of EMV travel times on different links in an arterial network can help quantify the impacts of various built environmental factors and interventions, but data availability is a problem. Prior work in this area, such as the EMVLight model from Su et al. (2023), use reinforcement learning as an optimization tool for route guidance and traffic control, not for prediction using real data. Agencies like FDNY have data only for their respective dispatches, and only for their vehicles, not including the traffic conditions downstream from those vehicles during the dispatches. As a result, many of the primary factors that impact EMV speed, such as downstream traffic densities and queues at the intersections are not known. To compensate for that, we propose an interpretable artificial intelligence (AI) model that is informed by a calibrated traffic simulation to supplement dispatch data with synthesized background traffic data. The result is a traffic simulation-informed explainable boosting machine (TSI-EBM) (see Lou et al., 2013) model for predicting emergency vehicle speeds. With such a model, several types of analyses can be conducted: (1) simulation-based analyses for the study area; (2) interpretation of the built environment features that contribute to the speed prediction; and (3) use of the predicted travel times as input to such logistics decisions as Cross Street Location (CSL) siting design – selecting optimal dispatch locations for EMVs – under different scenarios.

Proposed methodology

The EBM is a tree-based, cyclic gradient boosting generalized additive model, developed to be as accurate as state-of-the-art black box models while remaining interpretable. The structure of the EBM is highlighted in **Figure 1**(from Khattak et al., 2023), where x_j is the jth feature, y is the target variable, Θ is the link function that adapts the generalized additive model (GAM) for regression, and Γ_j is the attribute function. Training in EBM is performed repeatedly, with small trees built sequentially in each round. Each tree uses only one feature, and the residual (ε) is updated afterward. In this model, selected terms of interacting pairs of features are detected and added to standard GAMs, making the model more powerful than GAMs, but still intelligible. We propose a TSI-EBM, where certain features are obtained as an output of calibrated simulation of observed dispatch data. The output is a prediction of the EMV travel speed for a given link in the arterial network. An overview of the training and testing process for this model is shown in **Figure 2**.

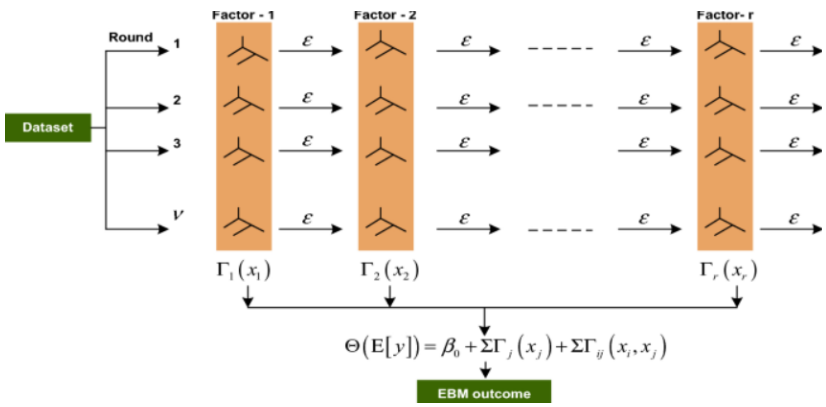

Figure 1. Illustration of EBM structure (source: Khattak et al., 2023)

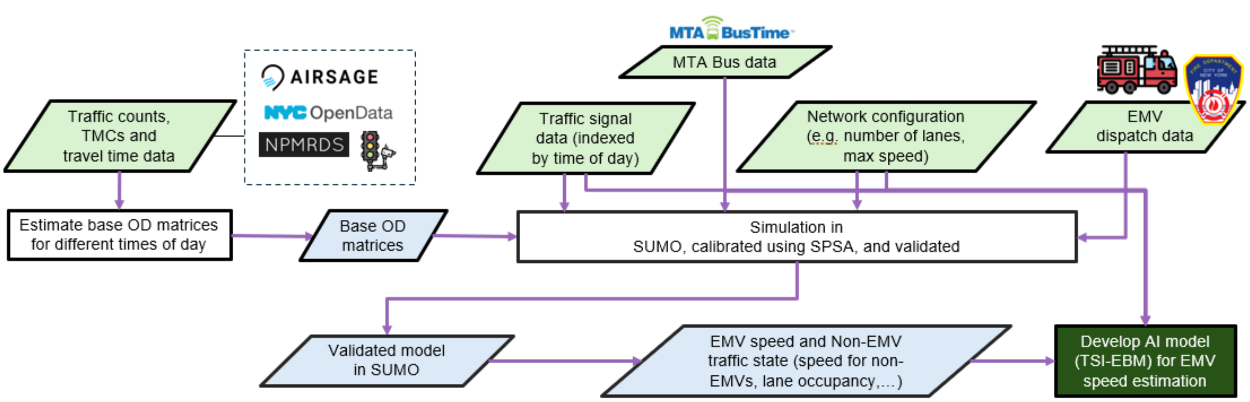

Figure 2. Flow diagram of training and testing of the traffic simulation-informed EBM.

With the trained model, three types of analyses can be done. With the underlying traffic simulation, evaluation of specific built environment changes on the distribution of response times can be ascertained. With the TSI-EBM model, we can further quantify and rank the different impacts that the features have on EMV response times. The model can also be used to predict travel times along each link to output scenario-based travel times to optimize CSLs under different settings. For the CSL optimization problem, we consider a reliability-based p-median problem (Snyder and Daskin, 2005) in which the average response time of the status quo locations can be calibrated to observed data to better reflect complexities of dispatching rules and uncertainties in availability of EMVs.

Data

The study area is the M6 District of FDNY, which covers the West Harlem neighborhood near Columbia University. As highlighted in Figure 2, a number of data sources are used to develop the traffic simulation model as well as the TSI-EBM. These data include travel time, speed, turning movement counts, signal timings, and volume and Origin-Destination data from local agencies, field observations and third-party data sources. The simulation itself is shown in **Figure 3**.

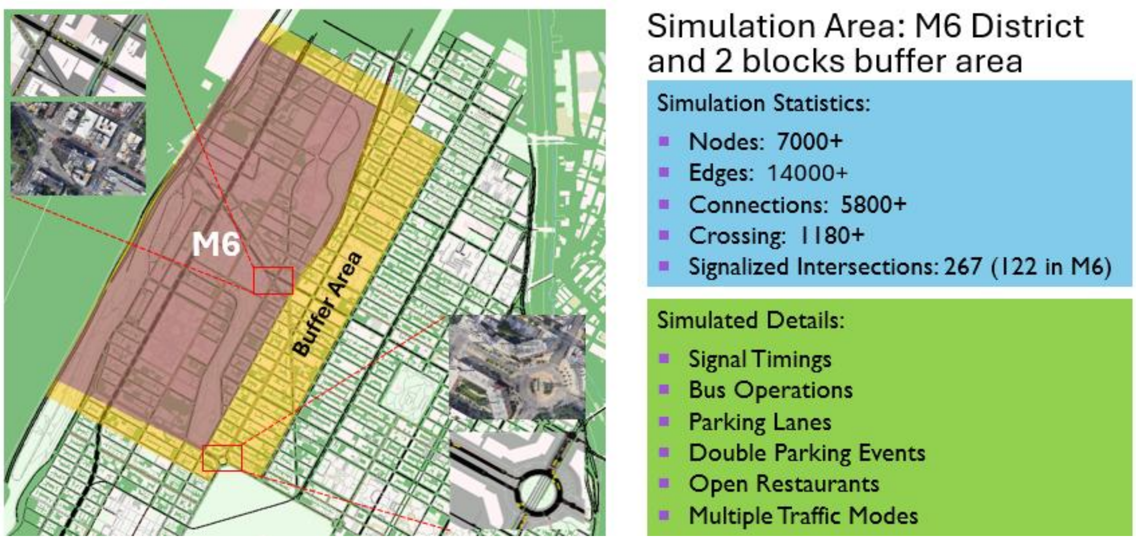

Figure 3. Traffic simulation overview in SUMO.

Link-based average EMV speeds estimated in the SUMO model were used as input data for the TSI-EBM, along with other extracted features such as background traffic speed, link occupancy, signal timings and road characteristics.

To estimate demand for the CSL optimization, EMS incidents provided by FDNY were used. The incidents were filtered to only include life-threatening cases that occurred within 2023. The optimization also used FDNY’s atom boundaries and current CSLs.

Results

The TSI-EBM was trained on the following set of features shown in **Figure 4**. Their importances are ranked. They suggest that the background passenger traffic has the most importance in predicting EMV speed. The presence of bike lanes is about 1/3 of the importance, and having non-parking protected bike lanes may serve as additional lane capacity for EMVs to increase their speeds. The model’s prediction accuracy, measured by the root mean squared error (RMSE) on the test data, is approximately 5 miles per hour.

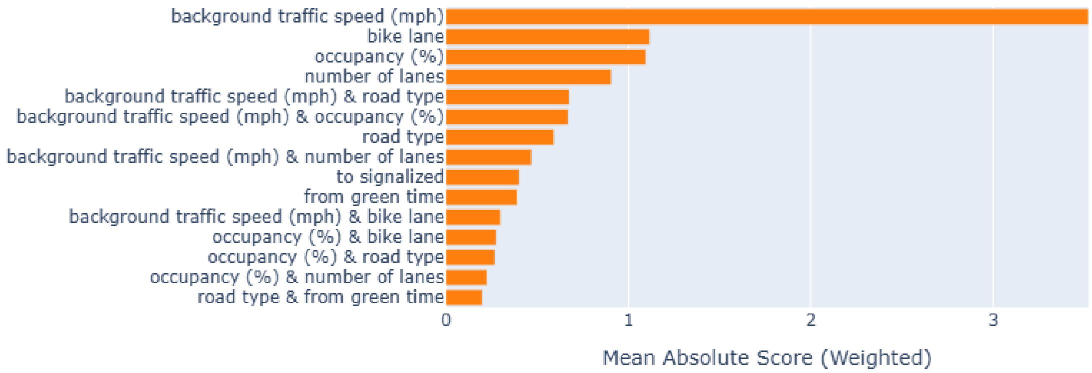

Figure 4. Interpretation of feature importance in TSI-EBM.

The presentation will include the full set of results which include the lane closure impact analysis, baseline comparison of the CSL optimization versus status quo locations using the predicted EMV travel times, comparison of that performance if locations were designed using passenger vehicle travel times instead, and an evaluation of the impact of several scenarios (bike lane removal, bike lane additions, bus stop additions/reductions, increase in traffic volumes) from all the roads in the neighborhood on the change in access time of the reoptimized CSL locations.

Acknowledgments

The research was funded by the C2SMARTER Center through USDOT #69A3552348326, with data and staff support from the Fire Department of New York, and data shared by New York City Department of Transportation and Airsage. The efforts of 15 students who helped collect field data are much appreciated. The recommendations and opinions expressed are those of the authors.

References

Berdowski, J., Berg, R. A., Tijssen, J. G., & Koster, R. W. (2010). Global incidences of out-of-hospital cardiac arrest and survival rates: systematic review of 67 prospective studies. Resuscitation, 81(11), 1479-1487.

Khattak, A., Chan, P. W., Chen, F., & Peng, H. (2023). Assessing wind field characteristics along the airport runway glide slope: An explainable boosting machine-assisted wind tunnel study. *Scientific Reports*, *13*(1), 10939.

Lou, Y., Caruana, R., Gehrke, J., & Hooker, G. (2013, August). Accurate intelligible models with pairwise interactions. In *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 623-631).

NYC, 2023. End-to-end response times. URL https://www.nyc.gov/site/911reporting/reports/end-to-end-repsonse-time.page, last accessed 8/16/2023.

Snyder, L. V., & Daskin, M. S. (2005). Reliability models for facility location: the expected failure cost case. *Transportation science*, *39*(3), 400-416.

Su, H., Zhong, Y. D., Chow, J. Y., Dey, B., & Jin, L. (2023). EMVLight: A multi-agent reinforcement learning framework for an emergency vehicle decentralized routing and traffic signal control system. *Transportation Research Part C: Emerging Technologies*, *146*, 103955.