

Machine Learning Surrogates for Optimizing Transportation Policies with Agent-Based Models

Elena Natterer^{*a}, Roman Engelhardt^a, Sebastian Hörl^b, and Klaus Bogenberger^a

^aChair of Traffic Engineering and Control, Technical University of Munich, Germany

^bInstitut de Recherche Technologique SystemX, Palaiseau, France

Extended abstract submitted for presentation at the 12th Triennial Symposium on Transportation Analysis conference (TRISTAN XII)

June 22-27, 2025, Okinawa, Japan

February 14, 2025

Keywords: Graph Neural Networks, Agent-based simulations, Multimodal Traffic Management, Surrogate Models, Optimization-Based Transportation Policies

1 INTRODUCTION

Urban growth intensifies traffic and pollution, necessitating tailored traffic management strategies such as congestion charges and repurposing car space, to reduce car usage during peak demand periods. For these measures to meet their intended goals and also achieve public acceptance, they must be designed carefully. Agent-based simulations are often employed to assess how potential policy measures impact the broader transportation system. For instance, Ben-Dor *et al.* (2024) used them to evaluate congestion charges and parking fees aimed at reducing car traffic and alleviating congestion, while Lu *et al.* (2023) examined speed limit impacts on road safety, traffic flow efficiency, and environmental factors.

A major limitation of agent-based models is their extensive runtime, which restricts the number of scenarios that can be tested and thus hinders the ability to identify optimal policy setups. Techniques such as Bayesian Optimization have been proposed to reduce the number of required simulations, enabling more efficient solution space exploration (e.g., Dandl *et al.* (2021), Huo *et al.* (2023)).

In contrast, if machine learning (ML) models can replicate essential outcomes of these agent-based models, they could serve as fast-executing surrogates, enabling large-scale simulation-based optimization. Previous research, such as Sroczynski & Czyżewski (2023), demonstrated that ML models can predict road traffic as effectively as complex microscopic simulations. Graph Neural Networks (GNNs) are particularly promising for transportation forecasting due to their ability to incorporate spatial dependencies within urban networks directly into the neural network structure. Jiang & Luo (2024) provided a comprehensive review of GNN-based traffic forecasting studies, highlighting available resources and research challenges.

This study aims to evaluate the feasibility of using Graph Neural Networks as surrogates for agent-based models in transport planning, with the long-term goal of enabling simulation-based optimization through agent-based simulations. To the authors' knowledge, only Narayanan *et al.* (2024) has provided a proof-of-concept, replicating the 4-step model with GNNs on networks of 15 to 80 nodes. Our study, by contrast, trains a GNN using MATSim simulation data for a Paris case study with 30,000 nodes. The results indicate that the trained model can successfully predict outcomes across different policy implementations.

^{*}Corresponding Author: elena.natterer@tum.de

2 METHODOLOGY

Our objective is to learn a function F that maps a set of input simulation parameters P_{sim} to the corresponding simulation output S :

$$F(P_{sim}) \rightarrow S \quad (1)$$

In our case, the simulation models the impact of an intervention on a transportation network. The output S consists of changes in car volume across different edges in the network. More specifically, for each edge $e \in E$, we aim to predict the change in car volume y_e due to the intervention. We define y_e as the simulated change and \hat{y}_e as the corresponding prediction made by our surrogate model. This defines a standard regression problem, where we learn a function $F : \mathbb{R}^{|E| \times n} \rightarrow \mathbb{R}^{|E|}$ that maps n road characteristics per edge to the predicted change due to an intervention.

To train the model, we use the Mean Squared Error (MSE) loss function, which is standard for regression tasks. It quantifies the difference between actual and predicted values:

$$\mathcal{L}(F) = \frac{1}{|E|} \sum_{e \in E} (F(x_e) - y_e)^2, \quad (2)$$

where x_e denote the road characteristics for road e . The objective is to find the optimal function F^* that minimizes this loss:

$$F^* = \arg \min_F \mathcal{L}(F). \quad (3)$$

The model utilizes a graph representation in which nodes are characterized by *static*, *positional*, and *variable* features linked to the roads of the street network. Because most GNNs are optimized for node-based predictions, we convert the simulation outputs into street network graphs and then transform them into dual graphs, where nodes correspond to road segments and edges capture the connections between segments. (1) *Static features* provide fixed attributes for each street, such as car volume and capacity in the base case, speed limit and segment length. (2) *Positional features* capture the spatial information of each street, including the coordinates of its start- and endpoint. (3) *Variable features* represent attributes modified by implemented policies, such as reductions in speed or capacity, or discounts in public transport fees.

The *base case* refers to a scenario without any policy intervention. It can be created by running the simulation tool multiple times with different random seeds and averaging the features on node (i.e., street) level.

The output predicts the change in car volume on each node (corresponding to a street in the road network) due to the policy in place. The architecture of the GNN, shown in Figure 1, begins with two layers of PointNet Convolutions (Charles *et al.* (2017)), which incorporate spatial features and process the positional attributes of each edge. This is followed by two layers of Transformer Convolutions, and one GAT Convolution layer completing the model.

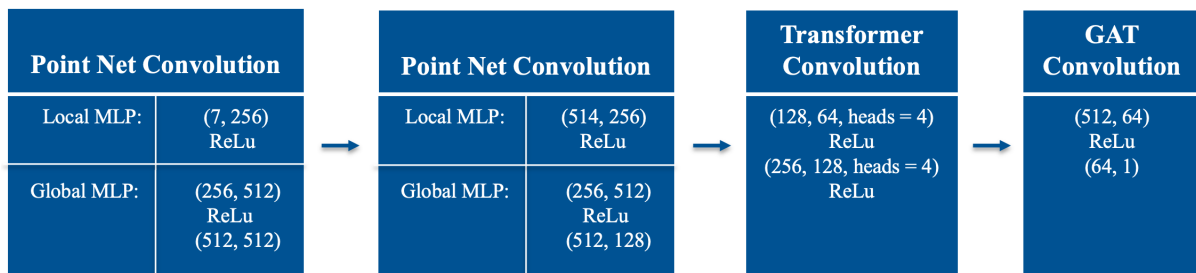


Figure 1 – Architecture of the Graph Neural Network

In addition to evaluating the model using the Mean Squared Error (MSE), as defined in Equation 2, we assess its performance with the coefficient of determination (R^2), offering a complementary measure of effectiveness. To establish a performance baseline, we define the *Naïve* MSE as the mean change in car volume across all nodes resulting from the policy.

3 RESULTS

We evaluate the proposed methodology using the MATSim model for Paris, France, based on a 1% downsampled population and 10,000 scenarios. We generate the data by first generating district combinations of its 20 districts (“Arrondissements”) by randomly sampling the districts so that the set length follows a normal distribution with a mean of 5, thereby preventing bias toward larger combinations. The tested policy involves a 50% capacity reduction on roads classified by OpenStreetMap as “Primary”, “Secondary”, or “Tertiary” within selected district combinations. Consequently, the model’s variable features are the reduction in capacity. To establish the base case, we run MATSim without policy interventions using 50 random seeds. We then simulate these scenarios to obtain the *ground truth* y . Finally, we split the resulting dataset into training, validation and test set with ratios of 80%, 15%, 5%.

The trained model is able to predict the effect of this policy interventions in random combination of districts within 0.1 seconds. It has an accuracy of 0.76 R^2 and a validation loss (MSE) of 24.95 (baseline: 103.47) on the test set. Table 1 presents the evaluation metrics for the model’s performance across various road types.

Table 1 – *Result table for different road types*

Road subset	R^2	Naïve MSE	Predicted MSE
All roads	0.76	103.47	24.95
Trunk roads	0.51	216.81	105.28
Primary roads	0.86	407.84	59.11
Secondary roads	0.65	82.67	29.10
Tertiary roads	0.58	47.84	19.87
Roads with policy in place	0.92	600.95	45.31
Roads without policy in place	0.39	36.63	22.37

The model demonstrates its highest performance in terms of R^2 , with a score of 0.92, on roads where a policy has been implemented. This suggests that the model effectively learned the relationship between capacity reduction and decreased car volume. Primary roads show the model’s second-best performance in terms of R^2 , likely due to the high traffic volume, which magnifies the effects of applied policies, allowing for more accurate predictions with reduced stochastic variation. This high traffic volume may also explain the elevated MSE baseline of 407.84: The mean car volume proves inadequate for primary roads, where usually greater traffic volume is observed. The lowest MSE appears on roads without policy intervention, with a baseline MSE of 36.63 and a predicted MSE of 22.37. This indicates that the average car volume after intervention is a good predictor for these roads, suggesting minimal impact from policies applied elsewhere in the network. This may also account for the model’s poorest R^2 performance (0.39). These findings suggest that the model effectively learns “direct” correlations, but can still improve in understanding the indirect effects of policies on non-intervention roads, such as evasive driving behavior. This observation is mirrored in Figure 2, illustrating the predicted versus actual changes in car volume at the street level for a randomly selected district combination from the test set. The model effectively captures the reduction in car volume, highlighted in varying shades of blue, indicating decreases of up to -50%; Driver evasive behavior, shown in shades of red representing increases of up to 50%, is slightly less pronounced. For this scenario, the model achieves a test loss of 3.66 (baseline is 137.67) and R^2 score of 0.81.

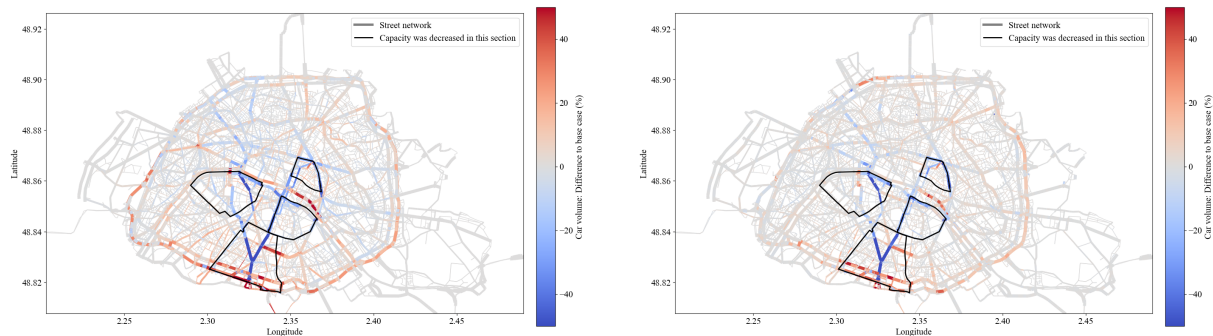


Figure 2 – Randomly chosen test data: Actual (left) and predicted (right) change in car volume

4 DISCUSSION

This paper presents a first approach of using Graph Neural Networks as surrogates for large-scale agent-based simulation models. In a case study using the MATSim model of Paris, the GNN successfully learned the impact of capacity reduction policies on the city wide traffic flow.

In the long run, this model could enable a range of applications due to the GNN’s rapid evaluation capabilities. Firstly, simulation-based optimization for determining optimal policy setups becomes feasible, allowing efficient exploration of large solution spaces. Secondly, real-time control applications become possible, as the model enables swift evaluation of scenarios when timely decision-making is crucial.

Achieving this vision requires the following future steps: 1) Expanding predictions to include multimodal impacts, such as mode shifts and public transit usage. 2) Testing a variety of policies to assess broader applicability. 3) Incorporating dynamic scenarios to enable real-time control applications. 4) Developing methods to transfer the trained model across different cities, enhancing its general applicability.

References

- Ben-Dor, Golan, Ogulenko, Aleksey, Klein, Ido, Ben-Elia, Eran, & Benenson, Itzhak. 2024. Simulation-Based Policy Evaluation of Monetary Car Driving Disincentives in Jerusalem. *Transportation Research Part A: Policy and Practice*, **183**(05), 104061.
- Charles, R. Qi, Su, Hao, Kaichun, Mo, & Guibas, Leonidas J. 2017. PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 77–85.
- Dandl, Florian, Engelhardt, Roman, Hyland, Michael, Tilg, Gabriel, Bogenberger, Klaus, & Mahmassani, Hani S. 2021. Regulating Mobility-on-Demand Services: Tri-Level Model and Bayesian Optimization Solution Approach. *Transportation Research Part C: Emerging Technologies*, **125**, 103075.
- Huo, Jinbiao, Liu, Zhiyuan, Chen, Jingxu, Cheng, Qixiu, & Meng, Qiang. 2023. Bayesian Optimization for Congestion Pricing Problems: A General Framework and Its Instability. *Transportation Research Part B: Methodological*, **169**, 1–28.
- Jiang, Weiwei, & Luo, Jiayun. 2024. Graph Neural Network for Traffic Forecasting: A Survey. *Expert Systems with Applications*, **207**(07), 117921.
- Lu, Qing-Long, Qurashi, Moeid, & Antoniou, Constantinos. 2023. Simulation-Based Policy Analysis: The Case of Urban Speed Limits. *Transportation Research Part A: Policy and Practice*, **175**(09), 103754.
- Narayanan, Santhanakrishnan, Makarov, Nikita, & Antoniou, Constantinos. 2024. Graph neural networks as strategic transport modelling alternative - A proof of concept for a surrogate. *IET Intelligent Transport Systems*, **1-19**.
- Sroczyński, Andrzej, & Czyżewski, Andrzej. 2023. Road traffic can be predicted by machine learning equally effectively as by complex microscopic model. *Scientific Reports*, **13**(1), 14523.