Enhancing Emission Estimation Through Machine Learning Emission Models: Bridging Meso- and Micro-Level Simulations

Alessandro Vissarios D'Apice*[1], Praveen Kumar Pandey*[2], Francesco Ciari[3], Luis F Miranda-Moreno[4]

[1] MSc, Civil Engineering, McGill University, Canada

[2] PhD, Civil Engineering, Polytechnique Montreal, Canada

[3] Professor, Civil Engineering, Polytechnique Montreal, Canada

[4] Professor, Civil Engineering, McGill University, Canada

1. INTRODUCTION

Climate change stands as one of the most pressing challenges facing society today. Governments around the world have made pledges to combat climate change and reduce greenhouse gas emissions, with the goal of limiting "the increase in the global average temperature to well below 2°C above pre-industrial levels" (UNCC, 2023). The transportation sector is the second biggest contributor to emissions, contributing 22% of all of Canada's carbon emissions (Environment and Climate Change Canada, 2024). There is no question that governments will need to target the transportation sector to reach their carbon goals. However, before they can reduce emissions in this sector, they must first be able to measure and understand where these emissions come from.

Vehicle emission measurement has been a key tool for environmental protection, dating back as early as 1966. Emissions factors are the main method of calculation, estimating the emissions produced by trips based on distance. This process has been continuously refined and improved upon, up to the newest modeling tool: the Environmental Protection Agency's motor vehicle emission simulator (MOVES) (USEPA, 2023). MOVES estimates a vehicle's emissions based on the second-by-second speeds and accelerations but can also estimate emissions based solely on a road network, and estimated vehicle volumes (USEPA, 2023). Naturally, the former method produces more accurate results compared to the latter (Ahn et al., 2002). The challenge for emissions modeling is getting this highly detailed data. Not only is historical data difficult to collect, but data on future trips is impossible to collect. To solve this, emissions modeling tools can be combined with transportation simulations.

This paper presents a method that integrates a car-following traffic model with a queue-based traffic model to leverage the advantages of both approaches. The queue-based approach is used to simulate the larger network, determining each agent's path and route choice, providing traffic flow validation on regional level, while the car-following model focuses on smaller subnetworks, simulating traffic flow and generating second-by-second speed and acceleration measurements for each vehicle. These measurements are then input into an emissions model powered by machine learning, which calculates vehicle emissions based on second-by-second speed data. By simulating smaller subnetworks, the microsimulation approach can cover a larger area more efficiently than modeling the entire network at once, saving computational resources while still leveraging the low computational requirements of agent-based models.

2. METHODOLOGY

A large-scale simulation of Quebec City is conducted using MATSim, leveraging the 2017 Quebec Origin-Destination (OD) survey data. Traffic flow data from counting stations is used to validate the hourly link counts generated by MATSim. There is a strong correlation between the observed traffic counts at the counting stations and the simulated link volumes from MATSim. The simulation runs for 500 iterations to achieve network equilibrium.

The integration focuses on two primary data sources: the traffic network and the travel demand. The traffic network serves as a key component in integrating MATSim with Aimsun, as both platforms represent networks in a similar manner. They utilize links to model roadways and nodes to represent intersections and turns. The process of transferring the MATSim traffic network into Aimsun involves exporting the network as a shapefile, which could then be imported into the Aimsun micro-simulator using the software's built-in import functionality.

The transfer of travel demand from MATSim to Aimsun involves processing the "events" file generated by MATSim after simulation. This file records actions taken by agents during the simulation, including events such as vehicles entering traffic, and pedestrians waiting at public transit stops.

This events file was processed to generate an origin and destination list for every user who traveld through the sub-network. Knowing the origins and destinations (OD), subnetwork (OD) matrices were generated for 24 hours. These OD matrices were the inputs used to simulate the travel demand within the Aimsun network.
Using vehicle emissions models generated by Alakus et al. (2024), emissions were predicted based on speed and acceleration at each second which was taken from the detailed second-by-second vehicle trajectories extracted from the micro-simulation.

3. Results and discussion: Quebec City case Study

To test the implementation of meso-to-micro scale integration, a case study was conducted using a subsection of the road network in Quebec City, Canada. Shown in

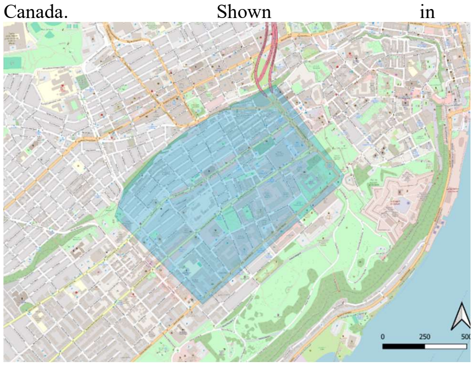

Figure 1**Error! Not a valid bookmark self-reference.**, it covers an area of 0.96 km^2, with 411links and 170 nodes. After simulating the subnetwork with the adjusted O-D matrices, the volumes reported by both simulations had a very high correlation, 0.92 for residential links and 0.89 for primary (arterial) and secondary (collector) links. Indicating that Aimsun is accurately replicating the validated traffic patterns generated by MATSim. Using trajectory-based emissions estimation, Aimsun predicted emissions demonstrated a correlation of 0.58 on residential and 0.80 for primary and secondary links. Compared to MATSim, Aimsun tended to have a lower emission estimate on these links. As shown in Figure 2, most data points lie below the equality line, indicating that MATSim consistently estimated higher emission than Aimsun.

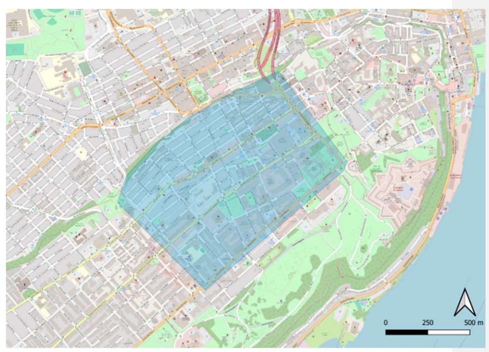

Figure 1: Sub-Network Area of Quebec City Case Study

Commented [pr1]: francesco: do you mean like local roads and arterial? Make sure you use words that are standard, this is not my field, so I cannot tell myself, but check it

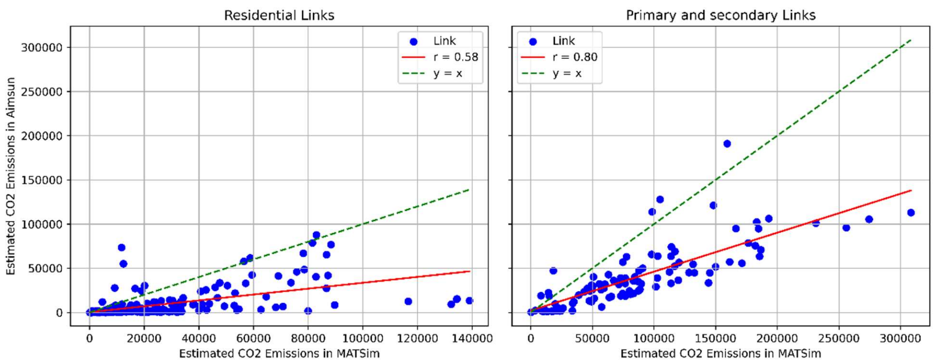

Figure 2: Comparison of Emissions Estimated Using MATSim's HBEFA Factors and Aimsun's XGBoost-Based Emission Model

The sub-network produced approximately 6.1 Tons of CO_2 over a 24-hour period. The top 10 CO_2 producing links make up 19% of those emissions, accounting for 1.17 Tons of CO_2. Figure 3 is a color-coded map of the sub-network based on the emissions generated on each link. Expectedly, there is a low amount of emissions generated by the residential streets in the sub-network, while the major links along the perimeter and crossing the center of the sub-network have much higher emissions generation.

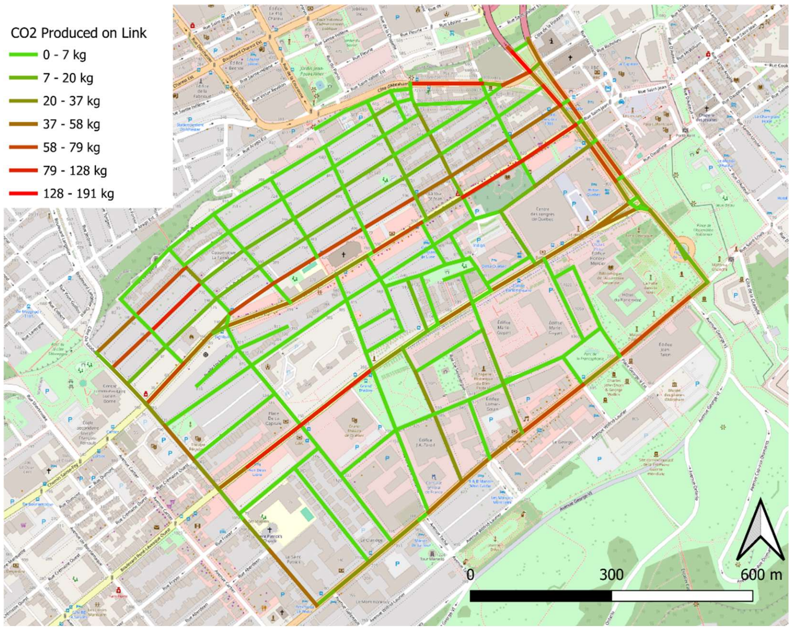

Figure 3: 24-hour CO_2 Emissions Production in Quebec City Sub-Network

The cause for the lower estimated emissions through the micro-simulation is most likely due to the smoothness of the simulated vehicle speed profiles compared to real-world speed profiles. The emissions models used in this study were trained on real world driving data and compared to simulated vehicle trajectories, real world driving data is much more erratic and inefficient. When these models get applied to simulated vehicle trajectories, they report much lower emission outputs compared to other methods of calculating emissions, such as using emissions factors, or average trip distance and fuel efficiency calculations.

4. Conclusions

This study demonstrates the integration of MATSim, an agent-based modeling tool, with Aimsun, a microscopic traffic simulation software, to estimate trajectory-based emissions and replicate traffic flow within a subnetwork of Quebec City.

The case study of the Quebec City subnetwork highlights significant differences in emission estimations produced by the two agent-based models. Emissions estimated based on average speed differ from those generated considering vehicle acceleration and speed variations. This study highlights that machine learning emissions models trained on real world data are not suitable for predicting simulated vehicle emission. This distinction underscores the importance of incorporating vehicle dynamics, such as vehicle speed, acceleration, and deceleration, when estimating vehicle emissions.

Another key area of exploration is determining the optimal subnetwork size is used to balance processing efficiency with the number of subnetworks needed, facilitating the application of this workflow to full city-scale networks. Additionally, integrating a car ownership model would enhance the realism of vehicle choice simulations, further improving model accuracy. Future studies will be focused on calculating emissions based on vehicle type in both simulation models to achieve greater realism in emissions estimation, accounting for variations in vehicle characteristics and emissions profiles.

5. References

Ahn, K., Rakha, H., Trani, A., & Van Aerde, M. (2002). Estimating vehicle fuel consumption and emissions based on instantaneous speed and acceleration levels. *Journal of transportation engineering*, *128*(2), 182-190.

Alakus, C., Miranda-Moreno, L., Labbe, A., Perez Villasenor, A., & Sun, L. (2024). Modeling Vehicle CO2 Emissions: Assessing Alternative Methods, Lag Effects, and Internal-External Factors. *Lag Effects, and Internal-External Factors (May 17, 2024)*.

Environment and Climate Change Canada. (2024). *Canadian Environmental Sustainability Indicators: Greenhouse gas emissions*. Retrieved 06 26, 2024 from www.canada.ca/en/environment-climate-change/services/environmental-indicators/greenhouse-gas-emissions.html

UNCC. (2023). *The Paris Agreement*. https://unfccc.int/process-and-meetings/the-paris-agreement

USEPA. (2023). *MOVES4 Technical Guidance: Using MOVES to Prepare Emission Inventories for State Implementation Plans and Transportation Conformity*. (EPA-420-B-23-011). U.S. Environmental Protection Agency