

Supplementary Information

Inequities in Canadian Traffic and Congestion Exposure Under Socioeconomic Pathways

Christina Draeger^{1,*}, Martino Tran¹, Ulysses Kamps¹, and Ben Mumford¹

¹*School of Community and Regional Planning, The University of British Columbia, Vancouver, BC, Canada*

^{*}*Corresponding author: Christina Draeger, cdraeger@eoas.ubc.ca*

S1 Volume Model Architectures

Random Forest We used a Random Forest regression model (Breiman, 2001) trained with 500 trees, where the number of variables considered at each split was set to 90 % of the input features. Node splitting was based on variance reduction, and a minimum node size of 5 was enforced to limit overfitting.

Extreme Gradient Boosting (XGBoost) An XGBoost model (Chen & Guestrin, 2016) was implemented with a squared error objective. The model was trained with a learning rate of 0.15 and maximum tree depth of 6, with additional regularization via L1 ($\alpha = 1$) and L2 ($\lambda = 2$) penalties, and a minimum child weight of 5. Row and feature subsampling rates were set to 0.8 to reduce overfitting.

Support Vector Machine (SVM) Support Vector Regression (SVR; Cortes & Vapnik, 1995) with a radial basis function (RBF) kernel was implemented. The model was trained using a regularization parameter of $C = 30$, with the kernel coefficient defined as $\gamma = \frac{1}{p}$, where p is the number of input features. An ϵ -insensitive loss function with $\epsilon = 0.005$ was employed, where ϵ defines the width of the region around the regression function within which errors are not penalized. Both input features and the target variable were standardized using z-score normalization prior to training.

Multi-Layer Perceptron (MLP) We used a multi-layer perceptron (MLP; Murtagh, 1991), a fully-connected feedforward neural network, with two hidden layers (128 and 64 units), with ReLU activation functions and dropout regularization (rates of 0.05 and 0.03, respectively). The model was trained using the Adam optimizer with a batch size of 32 and mean squared error as the loss function. Training was conducted for a maximum of 100 epochs with early stopping (patience of 10) based on validation loss to mitigate overfitting. Both input features and the target variable were standardized using z-score normalization prior to training.

S2 Road Capacity Model

Ensemble Modeling Approach Our segment-level road capacity model uses an ensemble approach, averaging the results of three speed-density models (Del Castillo & Benítez, 1995; Kerner et al., 1995; Krystek, 1980) to mitigate potential biases associated with any single speed-density model. These models were selected for their relative simplicity and because they do not require additional calibration, while having demonstrated reasonable performance against real-world data (Romanowska & Jamroz, 2021). All three models describe the relationship between speed and density.

The fundamental flow equation, which is used to calculate vehicle flow Q at each Level of Service (LOS; Tab. S1), is

$$Q_{\text{base,LOS}} = N \times k_{\text{LOS}} \times v_{\text{LOS}}, \quad (1)$$

where $Q_{\text{base,LOS}}$ is the base flow of a road segment at a given LOS (veh/h), N is the number of lanes in the road segment, k_{LOS} is the vehicle density at a given LOS (veh/km/lane), and v_{LOS} is the vehicle speed at a given LOS (km/h).

The Highway Capacity Manual (HCM) provides explicit values for density at capacity for motorways and for speed at capacity for other roads, primarily in urban areas (TRB, 2022). Therefore, we use two different formulations for each model: one with speed as a function of density (for motorways and trunks) and one with density as a function of speed (for primary, secondary, and tertiary roads). Note that only the equations for motorways and trunks are shown below; the corresponding equations can be rearranged to obtain density as a function of speed. For motorways and trunks, we calculate vehicle flow at LOS using the vehicle density values from the HCM, first calculating a base flow and then adjusting for heavy vehicles using a heavy-vehicle adjustment factor. For primary, secondary, and tertiary roads, we follow the same approach, except that vehicle flow at each LOS is calculated using the average vehicle speed values from the HCM.

The three capacity models used are as follows. The speed-density equation proposed by Krystek (1980) is

$$v_{\text{LOS}} = v_f, \quad (2)$$

while the speed-density equation proposed by Kerner et al. (1995) is

$$v_{\text{LOS}} = v_f \left(\frac{1}{1 + \exp\left(\frac{k_{\text{LOS}}/k_j - 0.25}{0.06}\right)} - 3.72 \times 10^{-6} \right). \quad (3)$$

The speed-density equation proposed by Del Castillo and Benítez (1995) is

$$v_{\text{LOS}} = v_f \left(1 - \exp\left(\frac{w}{v_f} \left(1 - \frac{k_j}{k_{\text{LOS}}}\right)\right) \right). \quad (4)$$

Here, v_{LOS} is the speed at a given LOS (km/h), v_f is the free-flow speed (km/h), k_j is the jam density (veh/km/lane), k_{LOS} is the density at a given LOS (veh/km/lane), and w is the kinematic wave speed at jam density (km/h).

Heavy-Vehicle Adjustment To account for the effect of heavy vehicles on roadway performance, base flow is adjusted for all road segments using a heavy-vehicle adjustment factor. The adjusted flow, Q_{adj} , is given by

$$Q_{\text{adj,LOS}} = Q_{\text{base,LOS}} \times f_{\text{HV}}, \quad (5)$$

where $Q_{\text{base,LOS}}$ is the base flow calculated above, assuming a homogeneous traffic stream of passenger vehicles, and f_{HV} is the heavy-vehicle adjustment factor, defined as

$$f_{\text{HV}} = \frac{1}{1 + P_t(E_t - 1)}, \quad (6)$$

where P_t is the proportion of trucks and buses in the traffic stream (expressed as a decimal), and E_t is the passenger-car equivalent (PCE) factor for heavy vehicles.

Model Parameters For motorways and trunks, density thresholds for each LOS are taken directly from the HCM (Tab. S3). For primary, secondary, and tertiary roads, speed thresholds are defined as a proportion of free-flow speed (f_0) and are also taken directly from the HCM (Tab. S4). Intersection effects are implicitly incorporated into the HCM thresholds and are therefore not modeled separately.

OSM (2026) provides the posted speed limit (PSL) for each road segment. Previous research has shown that PSL is strongly correlated with observed free-flow speed (FFS), with FFS typically 5–15 % higher than PSL across road types in Canada and the United States (Alomari et al., 2022; Kirley et al., 2023; Parker et al., 2003). Therefore, free-flow speed is estimated for each road segment by applying a 10 % increase to the PSL of each road segment.

58 Jam density (k_j) represents the maximum vehicle density under fully congested conditions. We use the default
59 value specified by the HCM of 118 veh/km/lane for all road types (TRB, 2022).

60 The kinematic wave speed at jam density (w), used in the model by Del Castillo and Benítez (1995), represents the
61 speed at which the effects of congestion propagate through a traffic stream under fully congested conditions. Based on
62 a review of real-world data and existing speed-density models, w has an average value of 20 km/h and is only weakly
63 related to road characteristics (Del Castillo & Benítez, 1995). Therefore, we use a value of 20 km/h across all road
64 segments.

65 Drawing on observed heavy-vehicle volumes from Quebec’s Average Annual Daily Traffic (AADT) counts (Tab. S2;
66 TMDQ, 2026), each road segment is assigned a P_t based on its road classification and metropolitan status. The PCE
67 factor represents the relative impact of a heavy vehicle on traffic flow compared with a passenger car. An E_t greater
68 than 1 indicates that heavy vehicles occupy more space, accelerate more slowly, and consequently reduce overall traffic
69 efficiency. Following FHWA (2002), we assume $E_t = 1.5$, a value commonly used for motorways in level terrain, and
70 apply it across all road types (FHWA, 2002).

S3 Supplementary Tables

Table S1: Definitions of Levels of Service (LOS) in traffic modeling, based on AASHTO (2012) and FHWA (2002).

Level of Service	Description
A	Free flow with low volumes and high speeds
B	Reasonably free flow but speeds beginning to be restricted by traffic conditions
C	Stable flow, but most drivers are restricted in the freedom to select their own speeds
D	Approaching unstable flow; drivers have little freedom to select their own speeds
E	Unstable flow with stoppages; volume equals capacity; any minor incident or lane change creates immediate, extensive traffic queues
F	Forced or breakdown flow, congestion, stop-and-go

Table S2: Average 30th-hour K-factors and heavy-vehicle fractions by road type and metropolitan status, based on observations from Quebec (TMDQ, 2026).

Region	Road type	K-factor (30th hour) (%)	Heavy vehicles (%)
Metropolitan	Motorway	10.9	11.7
	Trunk	11.5	11.3
	Primary	11.7	8.3
	Secondary	11.9	7.9
	Tertiary	12.1	9.3
Non-metropolitan	Motorway	12.6	18.6
	Trunk	16.6	21.6
	Primary	12.8	13.2
	Secondary	12.4	12.4
	Tertiary	12.6	11.5

Table S3: Input parameters and assumptions used in the capacity calculations for highways (motorways and trunk roads).

Variable	Type	Value for primary, secondary and tertiary roads	Source
Density at each LOS (k_{LOS})	Independent variable	LOS A: 6.875 veh/km LOS B: 11.25 veh/km LOS C: 16.25 veh/km LOS D: 21.875 veh/km LOS E: 28.125 veh/km	TRB (2022)
Free-flow speed (v_f)	Segment-level variable	Posted speed limit (PSL) increased by 10 %	OSM (2026)
Jam density (k_j)	Constant	118 pc/km/ln	TRB (2022)
Kinematic wave speed at jam density (w)	Constant	20 km/h	Del Castillo and Benítez (1995)
Passenger-vehicle equivalent (E_t)	Constant	1.5	FHWA (2002)
Proportion of heavy vehicles in the traffic stream (P_t)	Constant	Variable depending on road classification and metropolitan status (Tab. S2)	TMDQ (2026)
Average speed at each LOS (v_{LOS})	Dependent variable (output)	N/A	Output variable

Table S4: Input parameters and assumptions used in the capacity calculations for primary, secondary and tertiary roads.

Input	Type	Value for primary, secondary and tertiary roads	Source
Average speed at each LOS (v_{LOS})	Independent variable	LOS A: > 80% of free-flow speed LOS B: > 66% of free-flow speed LOS C: > 50% of free-flow speed LOS D: > 40% of free-flow speed LOS E: > 30% of free-flow speed LOS F: < 30% of free-flow speed	TRB (2022)
Free-flow speed (v_f)	Segment-level variable	Posted speed limit (PSL) increased by 10 %	OpenStreetMap (2026)
Jam density (k_j)	Constant	118 pc/km/ln	TRB (2022)
Kinematic wave speed at jam density (w)	Constant	20 km/h	Del Castillo and Benítez (1995)
Passenger-vehicle equivalent (E_t)	Constant	1.5	FHWA (2002)
Proportion of heavy vehicles in the traffic stream (P_t)	Constant	Variable depending on road classification and metropolitan status (Tab. S2)	TMDQ (2026)
Density at each LOS (k_{LOS})	Dependent variable (output)	N/A	Output variable

Table S5: Test set performance of the XGBoost traffic volume model, showing AADT, measured by Root-Mean-Squared Error (RMSE), R^2 , Mean Absolute Percentage Error (MAPE), and mean bias (absolute and percentage).

Region	Road type	RMSE	NRMSE (%)	R^2	MBE	MBE (%)	MAPE (%)
Metropolitan	Motorway	3221	2.4	0.99	-31	1.5	4.9
	Trunk	989	3.2	0.98	-7	3.2	7.2
	Primary	889	6.6	0.97	35	2.0	5.9
	Secondary	1617	12.4	0.94	-34	4.6	11.9
	Tertiary	1032	11.1	0.92	-2	10.1	19.2
Non-metropolitan	Motorway	508	4.9	0.99	41	1.0	2.4
	Trunk	294	7.2	0.99	-7	0.6	3.2
	Primary	406	5.9	0.97	9	2.3	6.6
	Secondary	170	11.9	0.95	1	5.4	12.4
	Tertiary	244	19.2	0.92	5	5.2	11.1
All	All	937	8.5	0.96	1	3.6	8.5

References

- Alomari, A. H., Khedaywi, T. S., Marian, A. R. O., & Jadah, A. A. (2022). Traffic speed prediction techniques in urban environments. *Heliyon*, 8(12), e11847. <https://doi.org/10.1016/j.heliyon.2022.e11847>
- American Association of State Highway and Transportation Officials (AASHTO). (2012). *Flexibility in highway design*. CreateSpace Independent Publishing Platform.
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- Chen, T., & Guestrin, C. (2016). XGBoost: a scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794.
- Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3), 273–297. <https://doi.org/10.1007/bf00994018>
- Del Castillo, J. M., & Benítez, F. G. (1995). On the functional form of the speed-density relationship—I: general theory. *Transportation Research Part B: Methodological*, 29(5), 373–389. [https://doi.org/10.1016/0191-2615\(95\)00008-2](https://doi.org/10.1016/0191-2615(95)00008-2)
- Federal Highway Administration (FHWA). (2002). *Highway performance monitoring system field manual: Appendix N – procedures for estimating highway capacity* (tech. rep.). U.S. Department of Transportation.
- Kerner, B. S., Konhäuser, P., & Schilke, M. (1995). Deterministic spontaneous appearance of traffic jams in slightly inhomogeneous traffic flow. *Physical Review E*, 51(6), 6243.
- Kirley, B. B., Robison, K., Goodwin, A. H., Harmon, K. J., O’Brien, N. P., West, A., Harrell, S., Thomas, L. J., & Brookshire, K. (2023). *Countermeasures that work: A highway safety countermeasure guide for state highway safety offices* (tech. rep. DOT HS 813 490). National Highway Traffic Safety Administration, Office of Behavioral Safety Research. Washington, DC. <https://rosap.nhtl.bts.gov/view/dot/72947>
- Krystek, R. (1980). *Synthetic index of the traffic flow quality as control criterion* (Habilitation thesis). Technical University of Gdańsk.
- Murtagh, F. (1991). Multilayer perceptrons for classification and regression. *Neurocomputing*, 2(5-6), 183–197. [https://doi.org/10.1016/0925-2312\(91\)90023-5](https://doi.org/10.1016/0925-2312(91)90023-5)
- OpenStreetMap contributors. (2026). OpenStreetMap (OSM) [Dataset]. <https://www.openstreetmap.org/>
- Parker, J., Martin R., Sung, H.-Y., & Dereniewski, L. J. (2003). *Review and analysis of posted speed limits and speed limit setting practices in British Columbia* (tech. rep.). BC Ministry of Transportation. Victoria, BC. https://www2.gov.bc.ca/assets/gov/driving-and-transportation/reports-and-reference/reports-and-studies/planning-strategy-economy/speed-review/speed_review_report.pdf
- Romanowska, A., & Jamroz, K. (2021). Comparison of traffic flow models with real traffic data based on a quantitative assessment. *Applied Sciences*, 11(21), 9914. <https://doi.org/10.3390/app11219914>
- Transportation Research Board (TRB). (2022). *Highway capacity manual: a guide for multimodal mobility analysis* (7th). National Academies Press. <https://doi.org/10.17226/26432>
- Transports et Mobilité Durable Québec (TMDQ). (2026). Débit de Circulation [Dataset]. <https://www.donneesquebec.ca/recherche/dataset/debit-de-circulation>