The Journal of Transport and Land Use jtlu.org
Vol. 19 No. 1 [2026] pp. 333-360
https://doi.org/10.5198/jtlu.2026.2867

JTLU

A data-driven recalibration framework for spatially transferring the vehicle ownership module of an agent-based integrated urban model

Ifratul Hoque, The University of British Columbia Okanagan, ifratulhoque1@gmail.com
Mahmudur Fatmi (corresponding author), The University of British Columbia Okanagan, mahmudur.fatmi@ubc.ca
Muntahith Mehadil Orvin, City of Edmonton, muntahith.orvin@edmonton.ca
Mohamad Ali Khalil, The University of British Columbia Okanagan, mohamad.khalil@ubc.ca

Abstract: Traditional transferability approaches can update only a single model at a time and are not designed for large systems. This constrains the spatial transferability of Integrated Urban Models, which comprise multiple micro-models. Recalibrating each micro-model independently with existing techniques would require detailed disaggregate data, undermining the practical objective of enhancing transferability. This study, therefore, proposes a recalibration framework that leverages aggregate data to update all micro-models within a complex modeling system. The framework combines a Steady-State Genetic Algorithm with a Random Forest surrogate model to identify optimal parameter sets. It was applied to the vehicle ownership module of the STELARS model, which represents vehicle transaction and type choice through six micro-models. The module, originally estimated for the Okanagan region, was transferred to the Greater Vancouver Area using only aggregate data sources. The recalibrated model reproduced observed household vehicle ownership distributions within about four percentage points, and 71.6% of the recalibrated parameters remained within 75% of their original estimates, indicating substantial preservation of the underlying behavioral structure. Parameter deviations reflected consistent behavioral differences between the regions rather than model misspecification. The proposed framework offers a scalable and resource-efficient approach to enhancing the spatial transferability of complex urban models.

Keywords: integrated urban models, vehicle ownership modeling, spatial transferability, data-driven modeling, recalibration

1 Introduction

Integrated Urban Models (IUMs) provide a comprehensive framework for simulating dynamic interactions among land use, transportation systems, demographic change, and environmental processes within urban regions. By representing how households and individuals make interconnected decisions such as residential location, travel behavior, and vehicle ownership, IUMs enable long-term policy analysis that captures feedback effects across multiple urban subsystems (Waddell, 2002). Prominent examples include the Integrated Land Use, Transport, and Environment (ILUTE) model, which integrates modules for demographic transitions, residential mobility, vehicle ownership, and travel behavior (Salvini & Miller, 2005). Each module typically comprises several behavioral micro-models estimated using detailed disaggregate data, such as specialized household travel and vehicle ownership surveys (Mohammadian & Miller, 2003; Roorda et al.,

---

ISSN: 1938-7849
Received February 23, 2026; Revised June 13, 2026; Accepted September 7, 2026; Available online September 22, 2026

The *Journal of Transport and Land Use* is the official journal of the World Society for Transport and Land Use (WSTLUR) and is published and sponsored by the University of Minnesota Center for Transportation Studies.

2000). Although these models have demonstrated substantial analytical capability, their reliance on extensive and highly context-specific detailed datasets renders them geographically constrained, thereby limiting their spatial transferability. This limitation has motivated growing interest in enhancing model transferability. Previous research in this area within the travel demand modeling domain has largely focused on transferring individual behavioral models, such as mode-choice or trip-generation models. However, transferring an entire IUM or its modules presents greater complexities due to the interdependence of multiple micro-models, the large number of parameters involved, and the need for disaggregate datasets.

This study aims to address this gap by developing a recalibration framework for transferring complex IUM modules using only aggregate data. The proposed approach was applied to the Vehicle Ownership (VO) module of the Simulator for Transportation, Energy, and Land use for Regional Systems (STELARS) model, transferring it from the Okanagan region to the Greater Vancouver Area (GVA). By demonstrating that reliable performance can be achieved using aggregate indicators, the study advances a scalable and resource-efficient approach for enhancing the spatial transferability of integrated urban models.

2 Literature review

2.1 Spatial transferability of integrated urban models and its components

Viewed over time, research on spatial transferability has followed an uneven trajectory. Interest was concentrated in the 1980s and 1990s, when disaggregate choice modeling prompted extensive testing of whether estimated models could be transferred across regions, primarily for mode-choice and trip-generation models. Research activity slowed through the 2000s and early 2010s, coinciding with a broader period of reduced momentum in integrated urban modeling and with the heavy disaggregate data requirements that made transfer studies costly. A renewed and growing interest has emerged since the mid-2010s, driven largely by machine-learning and data-driven methods that reframe transferability as a recalibration or transfer-learning problem rather than full re-estimation. This more recent phase emphasizes adapting existing models with limited or aggregate local data.

Early empirical investigations primarily focused on individual behavioral components. Agyemang-Duah and Hall (1997) demonstrated that modest adjustments improved predictive accuracy for shopping trip models within the Greater Toronto Area. Arentze et al. (2002) transferred a rule-based activity–travel model across Dutch municipalities and found that most behavioral components performed adequately, although mode-choice predictions were particularly sensitive to contextual differences. Koppelman and Wilmot (1982) showed that retaining structural form while recalibrating parameters improved work-trip mode-choice transferability. Subsequent research emphasized the importance of contextual similarity and selective recalibration. Sikder and Pinjari (2013) evaluated the transferability of daily activity generation and time-use models between regions within Florida and between Florida and California. Updating alternative-specific constants (ASCs) improved aggregate outcomes but failed to capture policy responsiveness, revealing that spatial proximity and urban similarity are important determinants of successful transfers. In a subsequent study, Sikder et al. (2014) assessed time-of-day choice models across San Francisco Bay Area counties and found that models estimated with pooled regional data achieved better generalizability than those estimated separately for each county. Fatmi and Habib (2015) observed that behavioral drivers shifted when transferring residential mobility models between Canadian cities, while Yasmin et al. (2015) and Bowman and Bradley (2017) highlighted that re-estimation or targeted updating was often necessary for acceptable performance.

Recent studies have explored a broader range of data-driven transfer strategies, spanning econometric updating, transfer estimation, and machine-learning methods. Tang et al. (2018) showed that neural-network models deteriorate when directly transferred across dissimilar regions but improve when limited local data are introduced through transfer learning. Linh et al. (2019) transferred the FEATHERS activity-based model from Belgium to Ho Chi Minh City and found that although the overall framework could be implemented, nearly all behavioral components required recalibration, with mode and location choice models showing the lowest transferability due to differences in infrastructure and cultural norms. Nohekhan et al. (2022) evaluated alternative transfer strategies for activity duration and activity type models and demonstrated that combined transfer estimation could achieve predictive performance comparable to full local re-estimation when even a small fraction of local data was incorporated. Kavta et al. (2024) investigated the transfer of mixed logit mode-choice models from Amsterdam to Manchester and showed that updating ASCs and selectively adjusting parameters substantially improved predictive reliability relative to direct transfer. Orvin et al. (2025) assessed the transferability of housing price models from Vancouver to Kelowna and reported that while aggregate predictions were acceptable, parameter recalibration was necessary to account for demographic and market-structure differences.

Collectively, this body of literature demonstrates that direct spatial transfer without adjustment consistently reduces predictive reliability. Most of this research, however, has focused on isolated micro-models rather than the interconnected, multi-component modules that characterize IUMs.

2.2 Model updating and recalibration techniques

To enhance transferability without full model re-estimation, several recalibration techniques have been proposed. One of the earliest and most widely applied approaches is updating ASCs. Koppelman and Wilmot (1982) examined the transfer of work-trip mode-choice models across regions in the Washington, DC, area and showed that recalibrating ASCs significantly improved predictive performance while retaining the original model structure. In subsequent work, Koppelman and Wilmot (1986), Koppelman and Rose (1985), and Koppelman et al. (1985) further demonstrated that adjusting constants, and in some cases scale parameters, could restore a substantial proportion of predictive accuracy in transferred disaggregate models. Santoso and Tsunokawa (2005) applied ASC adjustments when transferring mode-choice models between Ho Chi Minh City and Phnom Penh and reported notable improvements in aggregate market share predictions. However, Sikder and Pinjari (2013) observed that while ASC updating improved overall fit, it did not necessarily preserve behavioral responsiveness to policy variables.

Transfer scaling was introduced to address systematic differences in error variance between datasets. Ben-Akiva and Morikawa (1990) formalized this method, showing that scaling parameters could correct for differences in utility variance across contexts. Gunn et al. (1985) tested selective scaling of variable groups, while Badoe and Miller (1995) applied temporal scaling adjustments and found improved aggregate predictions. Walker et al. (1997) demonstrated that even small supplementary datasets could be used for scaling adjustments in regional travel demand models, and Karasmaa and Pursula (1997) compared multiple updating methods and found transfer scaling effective for simpler model structures.

Bayesian updating provides a more statistically rigorous alternative by treating transferred parameters as prior distributions and updating them with local observations. Atherton and Ben-Akiva (1976) showed that Bayesian updating could achieve reliable adaptation with limited data. Later applications by Karasmaa and Pursula (1997), Zhang and Mohammadian (2008), Rashidi et al. (2013), and Xu et al. (2014) demonstrated its effectiveness in transferring travel behavior and crash prediction models, particularly under small-sample conditions.

Combined transfer estimation, proposed by Ben-Akiva and Bolduc (1987), pools base and target data to estimate parameters that minimize mean square error. Santoso and Tsunokawa (2005) found that this approach produced the largest predictive improvements among tested methods. Joint context estimation extends pooled estimation by explicitly modeling cross-regional variance differences; Bradley and Daly (1991), Badoe and Miller (1998), Bowman et al. (2014), and Karasmaa (2007) showed that joint estimation improved generalizability and parameter stability across regions.

Model re-specification represents a more structural approach to improving transferability by modifying explanatory variables or functional forms to better capture locally relevant behavioral determinants. Koppelman and Wilmot (1986) showed that enriching model specifications could enhance absolute transferability, while Sikder et al. (2013) emphasized the importance of incorporating context-specific variables beyond conventional level-of-service measures. Similarly, Santoso and Tsunokawa (2005, 2010) demonstrated that improved specification, particularly when combined with other updating techniques, led to better predictive performance in transferred models. However, such modifications typically require extensive local behavioral data and may approach the effort of full re-estimation. More broadly, this limitation extends to all previously discussed updating strategies, whether constant adjustments, scaling, Bayesian methods, pooled estimation, or re-specification, as they depend on model-specific disaggregate data for parameter adjustment or validation. Consequently, these methods are not readily applicable for simultaneously updating multiple interdependent micro-models within an entire IUM module.

2.3 Present study

The reviewed literature points to a consistent conclusion: direct spatial transfer without adjustment degrades predictive reliability, yet the established remedies are difficult to apply at the scale of a complete IUM module. Updating techniques such as constant adjustments, transfer scaling, Bayesian updating, pooled estimation, and re-specification all depend on model-specific disaggregate data and were developed to adapt individual behavioral models one at a time. None is designed to recalibrate the multiple interdependent micro-models that constitute an IUM module simultaneously. Although recent machine-learning and data-driven studies have begun to reframe transferability as a recalibration rather than a re-estimation problem, they too have remained largely confined to isolated models. This leaves a clear methodological gap: a means of transferring an entire, multi-component module to a new region without requiring disaggregate target-region data. The present study develops a framework to address this gap.

To do so, this study makes three specific contributions. First, it recalibrates the VO module of the STELARS model using only aggregate data sources, eliminating the need for disaggregate revealed- or stated-preference data in the target region. Second, it tests transferability between two structurally very different contexts, transferring the module from the low-density, auto-oriented Okanagan region to the high-density, transit-rich Greater Vancouver Area, which provides a rigorous test under substantial regional contrast. Third, it introduces a data-driven, machine-learning optimization approach that couples a Steady-State Genetic Algorithm (SSGA) with a Random Forest (RF) surrogate model to jointly recalibrate the module's interdependent micro-models efficiently while preserving their original behavioral structure.

3 Conceptual framework of STELARS

3.1 Overview of STELARS

STELARS is an agent-based model designed to simulate dynamic interactions among land use, transportation, and energy systems. Although not originally developed for a specific region, it was first estimated and validated for the Okanagan region. The framework represents individuals and households as decision-making agents, while vehicles and dwellings are treated as non-agent entities characterized by attributes but without independent behaviors. STELARS adopts an event-based hybrid of discrete- and continuous-time simulation. Agents remain inactive until triggered by life-course or contextual events such as birth, death, relocation, or employment transitions. This structure enables dynamic representation of demographic evolution, residential mobility, and behavioral change over time. The framework is implemented in Python using an object-oriented structure, which facilitates modular development.

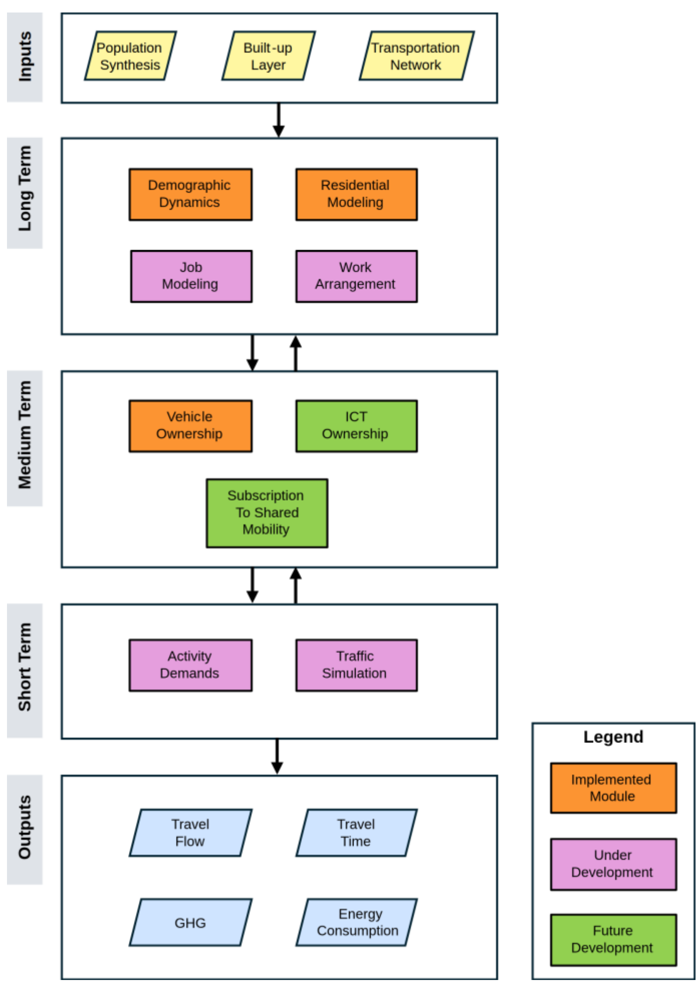

Figure 1. Conceptual flowchart of STELARS

As illustrated in Figure 1, which is adapted from Khalil (2025), STELARS is organized into three interconnected decision phases: long-term, medium-term, and short-term. The model relies on three primary inputs: (1) a synthetic population, (2) a built environment representation, and (3) a transportation network. The synthetic population developed by Rahman and Fatmi (2023) incorporates demographic, socio-economic, and residential attributes generated using a Bayesian network and generalized raking techniques to ensure consistency with regional control totals.

The long-term phase includes Demographic Dynamics, Residential Modeling, Job Modeling, and Work Arrangement modules. The Demographic Dynamics module, developed by Khalil et al. (2025), simulates life-course events such as aging, marriage, childbirth, and employment transitions. The Residential Modeling module represents relocation as a multi-stage process including mobility decision, search, bidding, and location choice (Orvin et al., 2024; Orvin & Fatmi, 2024). Job Modeling captures employment transitions, while the Work Arrangement module distinguishes between commuting, remote, and hybrid work patterns.

The medium-term phase models household-level asset and subscription decisions, including VO, Information and Communication Technology (ICT) ownership, and subscription to shared mobility. The VO module simulates vehicle acquisition, trade, disposal, and attribute decisions. ICT and subscription to shared mobility modules are planned to complement vehicle ownership by capturing digital connectivity and multimodal access.

The short-term phase represents daily activity-travel behavior and network performance. The Activity Demands module simulates activity participation, duration, timing, and mode choice (Khaddar et al., 2024; Khalil & Fatmi, 2023), distinguishing between mandatory and discretionary activities. The Traffic Simulation module, implemented in MATSim, translates activity-travel patterns into network-level outputs such as flows, travel times, greenhouse gas emissions, and energy consumption. The modular and feedback-driven architecture enables STELARS to simulate the co-evolution of urban form, mobility, and energy systems.

3.2 Vehicle ownership module

The VO module is a core component of the medium-term phase and constitutes the focus of this study. It was originally estimated for the Okanagan region using the Travel Technology and Mobility Survey (TTMS) (Hossain, Fatmi, & Enam, 2026; UBC integrated Transportation Research Laboratory [UiTR], 2022), a retrospective household survey containing residential history, employment transitions, vehicle ownership records, and life-cycle events. The module represents household-level vehicle ownership dynamics within an event-based microsimulation framework (Hossain, Fatmi, & Khalil, 2026).

As shown in Figure 2, the VO module receives updated demographic and residential information from the long-term phase, ensuring consistency between household composition, dwelling characteristics, and vehicle decisions. The VO module is organized into two stages, vehicle transaction and vehicle type choice. The vehicle transaction stage contains two models, First Vehicle Purchase, which models the transition from zero to first vehicle ownership, and Vehicle Fleet Adjustment, which captures subsequent additions, trades, or disposals. Each transaction triggers a vehicle type choice model, jointly determining body type, fuel type, vintage, and presence of advanced technology. Body type alternatives include subcompact, compact, midsize/large, SUV, and truck/van. Fuel type includes gasoline, diesel, and alternative fuel vehicles (AFVs). Vintage is categorized as new (≤1 year), used (2–5 years), or old

(>5 years). Technology choice distinguishes between vehicles with and without advanced features. These six models account for household socio-economic attributes, life-cycle events, built environment characteristics, and prior ownership history. The module outputs detailed household-level vehicle ownership outcomes at the end of the simulation period, including total fleet size and vehicle attributes. Although Figure 2 indicates a conceptual feedback loop between vehicle ownership and short-term activity decisions, short-term modules were not simulated in this study.

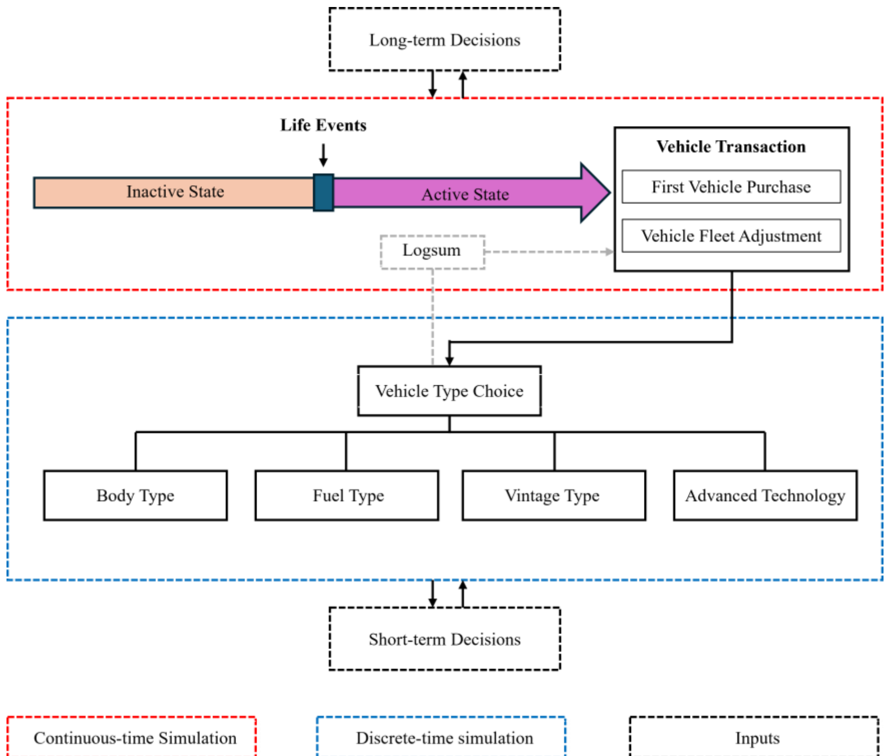

Figure 2. Conceptual flowchart of the vehicle ownership module

4 Study area and data sources

4.1 Study area

As mentioned earlier, this study aims to enhance the transferability of the VO module of STELARS by updating the model estimated for the Okanagan region to perform well in the GVA.

The Okanagan region represents a low-density, mixed rural and urban environment characterized by dispersed settlements and smaller urban centers, including Kelowna, West Kelowna, Vernon, Peachland, and Lake Country. Kelowna serves as the regional hub, with roughly 145,000 residents and a density of about 680 persons per square kilometer, while surrounding municipalities are smaller and less dense. The region is largely auto-oriented, with relatively limited public transit services. In contrast, the GVA is a high-density metropolitan region of more than 3 million residents, accounting for over half of British Columbia's population, and comprising 21 municipalities. It features a well-developed and integrated transportation network, including the SkyTrain rapid transit system, SeaBus, RapidBus corridors, and an extensive bus network. The two study areas are shown in Figure 3 and Figure 4.

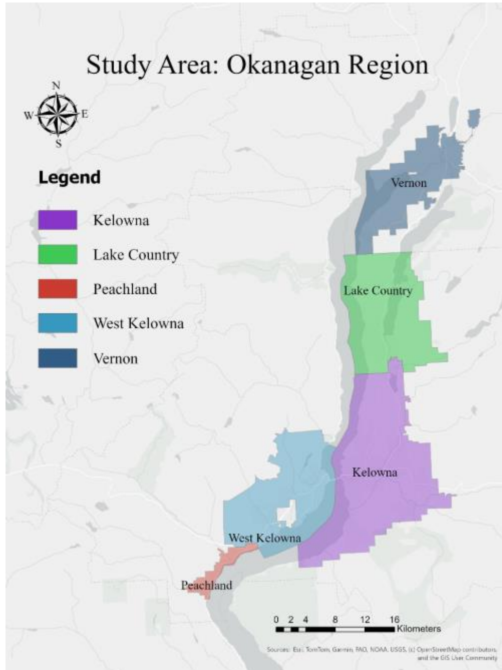

Figure 3. Okanagan region

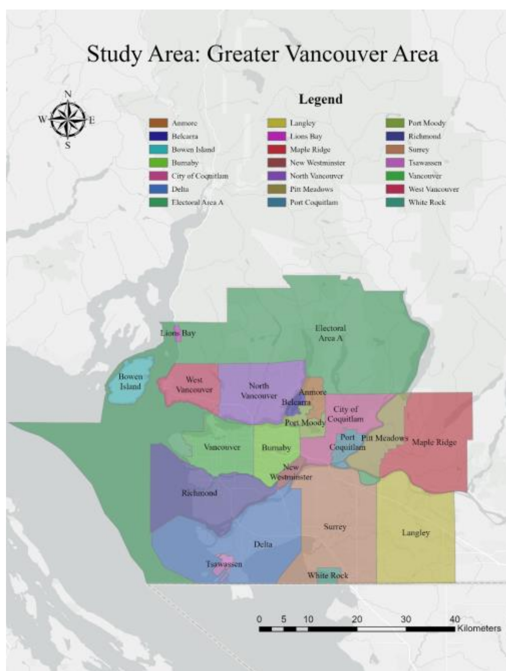

Figure 4. Greater Vancouver Area

4.2 Data sources

The recalibration framework in this study relies on two aggregate data sources: (1) the Insurance Corporation of British Columbia (ICBC) Vehicle Population Dataset (ICBC, 2025), and (2) the British Columbia Activity Time Use Survey (BC ATUS) (UBC integrated Transportation Research Laboratory, 2024).

4.2.1 ICBC vehicle population dataset

The ICBC dataset provides municipality-level counts of registered personal vehicles for 2024 and serves as the primary source for vehicle attribute distributions. The dataset includes filtering options for municipality, body style, fuel type, make, model, model year, and anti-theft device presence.

Vehicle attributes were reformatted to align with the STELARS VO module structure. Body types were classified into subcompact, compact, midsize/large, SUV, and truck/van categories based on make and model. Vehicle vintage was derived from model year relative to 2024 and categorized as new (≤1 year), used (2–5 years), or old (>5 years). Fuel types were grouped into gasoline, diesel, and AFVs, including hybrid and electric vehicles. The presence of advanced technology was proxied using the anti-theft device indicator, as it was the only consistently available technology-related variable. The ICBC vehicle population dataset is publicly available (Insurance Corporation of British Columbia, 2025).

4.2.2 British Columbia Activity Time Use Survey

The BC ATUS is a 2024 activity time-use survey that captures travel behavior and socio-demographic characteristics across British Columbia (Fatmi & Ahmed, 2025). Importantly, it includes household-level VO information. After data cleaning, the final sample comprised 2,391 individuals across 1,594 households. The distribution of total household vehicles (0, 1, 2, and 3+ vehicles) was extracted and used as the target ownership distribution for recalibration. Comparison with the 2021 Census indicates that the survey closely reflects the demographic composition of the GVA population, supporting its use as a reliable benchmark.

5 Simulation setup for the Greater Vancouver Area

To implement the spatial transferability assessment, the STELARS framework was applied to the GVA over the 2011 to 2024 simulation horizon. Because the simulation began in 2011, it required a complete 2011 setup for the GVA. This setup comprises four primary input files: (1) a 2011 synthetic population of the GVA, (2) a dynamic dwelling supply file, (3) a baseline household-level vehicle ownership database, and (4) a temporal vehicle price file. The 2011 synthetic population for the GVA was generated using the same procedure as the original STELARS implementation (Rahman & Fatmi, 2023), in which demographic, socio-economic, and residential attributes were synthesized using a Bayesian network and generalized raking to match 2011 Census control totals for the GVA. The dynamic dwelling supply file, constructed from BC Assessment records (BC Assessment, 2025), provides the annual stock of available dwelling units that the Residential Modeling module draws on to allocate households over the simulation horizon. The baseline vehicle ownership database supplies the initial household-level vehicle holdings for 2011, and the temporal vehicle price file provides vehicle prices across the simulation period. The aggregate ICBC and BC ATUS datasets introduced in Section 4.2 serve as the recalibration targets for vehicle attributes and household ownership, respectively.

Although the central focus of this study is the recalibration of the VO module, the STELARS architecture requires the Demographic Dynamics and Residential Modeling modules to run before the VO module, so that vehicle decisions reflect accurate, up-to-date household and residential conditions. For the simulation to represent the actual GVA context, these two upstream modules must themselves be calibrated for the region. They had already been recalibrated for the GVA, with their performance documented in Hoque

(2025), and were adopted here without modification. The operational engine, comprising Demographic Dynamics, Residential Modeling, and VO modules, was then run annually from 2012 to 2024 (with 2011 serving as the base year). The demographic and residential outputs were validated against 2016 and 2021 Census data and reproduced realistic population growth, household formation, and spatial distribution patterns. Because these upstream modules performed satisfactorily, the recalibration effort could focus on the VO module alone, without revisiting the modules that preceded it.

In addition, the relationship between residential location and vehicle ownership in STELARS operates through a directional information flow rather than joint estimation. The Residential Modeling module, described in Section 3, draws on the dynamic dwelling supply file as a supply-side input. The dwelling supply file informs the location-choice process but does not itself perform location choice. The resulting residential and dwelling outcomes are then passed to the VO module as inputs, so that residential conditions influence vehicle decisions while household composition, dwelling characteristics, and vehicle outcomes remain consistent. Because this linkage is directional, recalibrating the VO module adjusts only how vehicle outcomes respond to these residential inputs; it does not modify the Residential Modeling module or its location-choice outputs, which were retained as calibrated for the GVA.

6 Methodology

6.1 Recalibration framework

Transfer strategies span a spectrum. At one end is naïve transfer, the direct application of source-region parameters without adjustment. At the other end is full econometric re-estimation of the VO module for the GVA, which would yield statistically validated and behaviorally interpretable parameters but requires disaggregate revealed- or stated-preference data for the target region. The proposed recalibration occupies the practical middle ground: it adapts the existing Okanagan model structure to match observed aggregate ICBC and BC ATUS distributions, recovering aggregate predictive validity without requiring disaggregate target-region data. The proposed recalibration framework, comprising the stages of surrogate-assisted optimization, is illustrated in Figure 5.

The process begins with the direct deployment of the Okanagan-estimated parameters in the GVA setting to establish a baseline reference and identify recalibration needs. The Mean Absolute Error (MAE), which measures the distance between simulated and observed outcome distributions, was selected as the error function to be minimized during optimization. This objective function was formulated independently for each recalibration category to allow modular adjustment of distinct behavioral components. The search space was then defined by bounding each parameter within ±100% of its original estimate while preserving its sign to maintain behavioral consistency. Latin Hypercube Sampling (LHS) was used to generate an initial set of parameter vectors that uniformly cover the multidimensional search space. Each sampled vector was evaluated through the full STELARS simulation to obtain true MAE values, which serve as training data for an RF surrogate model.

The core optimization stage employs an SSGA, which iteratively explores the parameter space using surrogate-predicted fitness values rather than repeatedly invoking the full simulation. Tournament selection, Simulated Binary Crossover, and bounded Polynomial Mutation are applied to generate new candidate solutions, while poor-performing solutions are gradually replaced to preserve diversity and avoid premature convergence. Upon convergence, the best-performing parameter vector identified through the surrogate-assisted search was validated using the full STELARS simulation. This final evaluation ensures that the recalibrated parameters achieve genuine improvements in aggregate vehicle ownership and attribute distributions while maintaining structural consistency with the original VO specification.

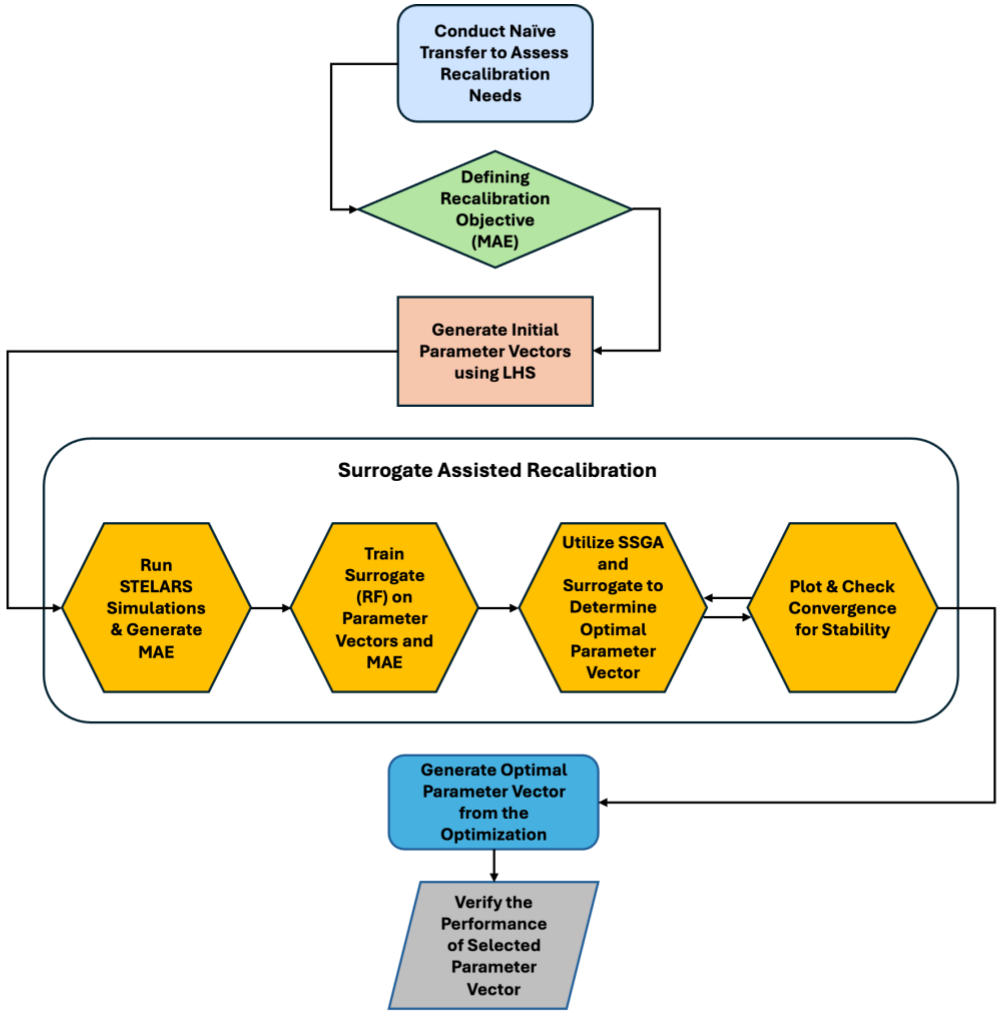

Figure 5. Recalibration flowchart

6.2 Naïve transfer assessment

Prior to recalibration, the VO module estimated for the Okanagan region was directly applied to the GVA without any parameter adjustment. The simulation was initialized using 2011 GVA inputs and executed through 2024. Figure 6 compares the simulated and observed (BC ATUS) household-level vehicle ownership distributions for 2024. The naïve transfer overestimated one- and two-vehicle households while underestimating zero-vehicle households. This pattern reflects behavioral differences between the two regions: the auto-oriented and lower-density Okanagan context embeds higher ownership propensities that do not translate directly to the denser, transit-rich GVA environment.

Figures 7 to 10 present comparisons of simulated and observed (ICBC) vehicle attribute distributions across body type, fuel type, vintage, and technology dimensions in 2024. The model overpredicted SUVs and gasoline-powered vehicles while underpredicting compact and alternative fuel vehicles, indicating a failure to capture the stronger prevalence of compact forms and cleaner technologies in GVA. Similarly, newer and technologically advanced vehicles are underpredicted, suggesting that the original Okanagan calibration does not fully reflect GVA's adoption dynamics.

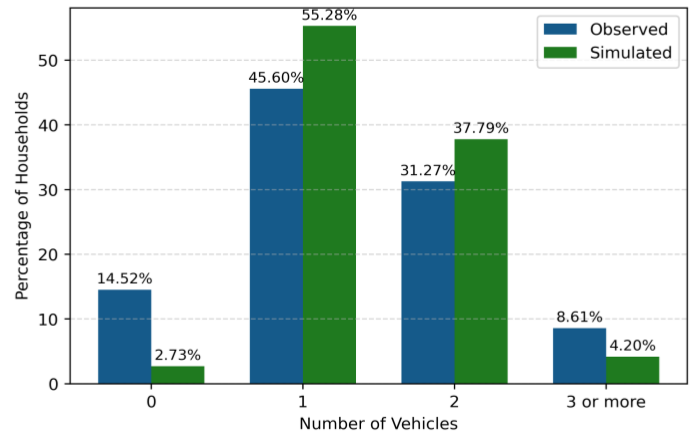

Figure 6. Household-level vehicle ownership distribution comparison for naïve transfer

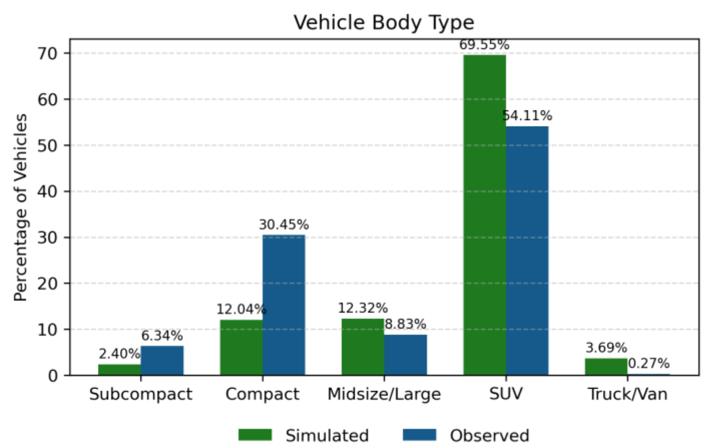

Figure 7. Vehicle body type distribution comparison for naïve transfer

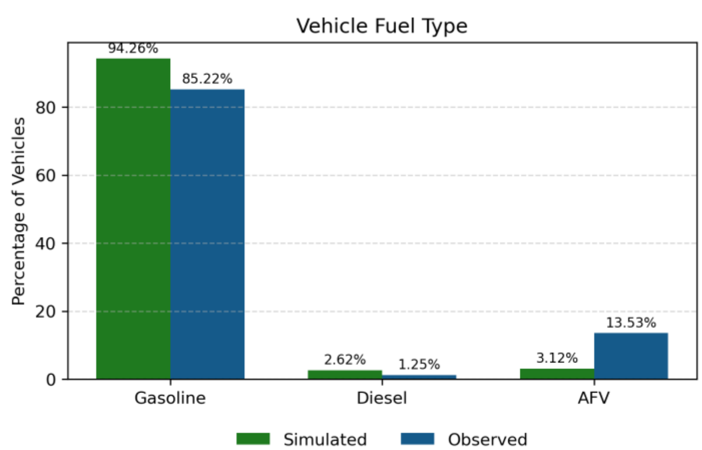

Figure 8. Vehicle fuel type distribution comparison for naïve transfer

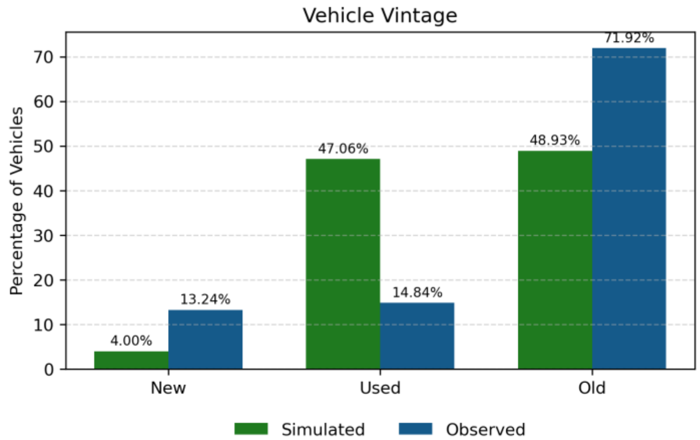

Figure 9. Vehicle vintage type distribution comparison for naïve transfer

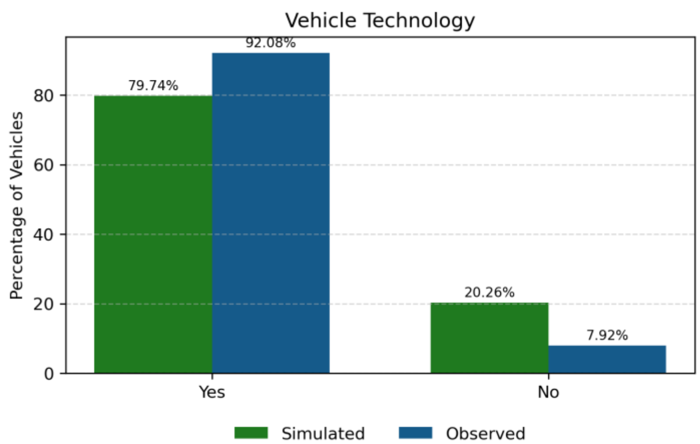

Figure 10. Vehicle technology type distribution comparison for naïve transfer

6.3 Recalibration objective function

The recalibration objective, MAE, is formulated as a distributional matching problem, where simulated outputs from the VO module are aligned with observed aggregate targets for the GVA. MAE was selected due to its scale-independence, interpretability, and reduced sensitivity to extreme deviations compared to squared-error metrics (Willmott & Matsuura, 2005). Because the recalibration targets are aggregate proportions rather than disaggregate likelihood contributions, MAE provides a direct and robust measure of distributional fit. By operating on shares rather than raw counts, the loss function remains invariant to population size and focuses exclusively on replicating behavioral patterns. It is defined as:

$$MAE = \frac{1}{N} \sum_{i=1}^{N} | s_i^{sim} - s_i^{obs} | \quad (1)$$

Here, s_i^{sim} denotes the simulated proportion of recalibration category i, s_i^{obs} denotes the corresponding observed proportion, and N is the total number of outcome categories within the same model. The MAE objective is applied independently to each of the five recalibration categories, namely household vehicle ownership, body type, fuel type, vintage, and technology. The VO module is built from six micro-models, but these map onto five recalibration categories because two of the models share a single aggregate target. The First Vehicle Purchase model and the Vehicle Fleet Adjustment model (the latter comprising the addition, trade, and disposal equations) are recalibrated together as one category, since jointly they determine the household-level vehicle ownership distribution, which forms a single aggregate target (the 2024 BC ATUS household vehicle counts). MAE for this category is computed against that distribution. The remaining four categories correspond to the vehicle type choice attributes. Each attribute is modeled for both first-purchased and transacted vehicles, so its category draws on the equations from both the first purchase and the transaction side, and MAE is computed against the corresponding 2024 ICBC distribution. The recalibration is therefore organized by aggregate target distribution rather than by individual model, which is why six micro-models yield five recalibration categories.

6.4 Parameter space definition and initialization

To preserve behavioral consistency while allowing contextual adaptation, each parameter was permitted to vary within ±100% of its base value, subject to sign preservation. This constraint ensures that the directional effects of explanatory variables remain consistent while allowing sufficient flexibility to accommodate regional differences. The width of this bound was established empirically rather than assumed. An initial pilot using a narrower ±50% range was conducted to test whether the parameters could adapt within it. A substantial proportion of parameters across categories reached the ±50% boundary, indicating that this range was too restrictive to accommodate the degree of adaptation required between the Okanagan and GVA contexts. The bound was therefore extended to ±100%, which preserves parameter signs while allowing each coefficient to vary from near zero up to twice its base value. This range is an operational search bound rather than a statistical interval, and its adequacy is confirmed in Section 7.3, where no recalibrated parameter reached the ±100% limit, indicating that the extended bound was sufficiently wide without being excessive. Logsum parameters associated with nested structures were held fixed at their original values to maintain structural coherence and avoid destabilizing substitution patterns across alternatives. As a result, the recalibration focuses exclusively on utility coefficients representing behavioral sensitivities.

To generate an initial set of candidate parameter vectors, LHS was employed. LHS ensures stratified and space-filling coverage of high-dimensional parameter spaces, improving representativeness relative to purely random sampling (McKay, 1992). For each recalibration category containing k parameters, the number of initial samples N was defined as:

$$N = 15 \times k \tag{2}$$

This sampling density follows established recommendations for surrogate-assisted optimization workflows, where approximately 10–20 samples per parameter provide sufficient coverage for reliable surrogate training (Afzal et al., 2017). The resulting parameter vectors formed the initial evaluation set and were used to generate true simulation-based loss values.

6.5 Surrogate-assisted optimization

After generating the initial parameter vectors θ by LHS, each vector was evaluated using the full VO simulation to compute the MAE. To ensure comparability across candidate solutions within each recalibration category, the MAE was normalized using min-max normalization:

$$MAE_{norm}(\theta) = \frac{MAE\ (\theta) - \ min(\ MAE\ (\theta'))}{max(\ MAE\ (\theta')) - \ min\ (\ MAE\ (\theta'))} \tag{3}$$

This normalization maps all performance values to the interval [0,1], where lower values indicate better fit. The resulting dataset of parameter-loss pairs was then used to train an RF surrogate model f, which approximates the true loss surface:

$$\hat{f}(\theta_{new}) \approx \ MAE_{norm}\ (\theta_{new}) \tag{4}$$

The RF surrogate (Breiman, 2001) is well suited for high-dimensional, nonlinear, and simulation-based optimization problems due to its robustness and resistance to overfitting (Wang & Jin, 2020; Zheng et al., 2019).

The SSGA operates on this surrogate-estimated fitness landscape. At each iteration t, two parent parameter vectors $\theta_{p1}, \theta_{p2} \in P_t$ are selected from the current population P_t using tournament selection. Offspring are generated through Simulated Binary Crossover (Deb & Agrawal, 1995) followed by Polynomial Mutation:

$$\theta_{new} = Mut\left(Cross(\theta_{p1}, \theta_{p2})\right) \tag{5}$$

The surrogate model predicts the performance of the new candidate. Then replacement follows a steady-state strategy: if the predicted loss of the new candidate is lower than the worst-performing individual in the current population, the population is updated as:

$$if\ \hat{f}(\theta_{new}) < \ \max_{\theta \in P_t} \hat{f}(\theta) \ \Rightarrow P_{t+1} = \ P_t \setminus \{\theta_{worst}\} \cup \{\theta_{new}\} \tag{6}$$

This incremental replacement preserves high-performing solutions while maintaining diversity (Whitley, 1994) and is particularly effective for expensive optimization problems (Jin, 2011).

Full STELARS evaluations were limited to the one-time initial LHS training set of N = 15 × k parameter vectors per category. The SSGA search operated entirely on surrogate predictions and added no further STELARS runs during optimization, which is the central reason the surrogate-assisted design is computationally efficient relative to invoking the full simulation at every candidate evaluation. Upon convergence, the best parameter vector identified by the SSGA for each category was run once through the full STELARS simulation to obtain its true, non-surrogate MAE. The final recalibrated vector θ* was then the fully evaluated candidate with the lowest true MAE, as expressed in Equation (7):

$$\theta^* = arg \min_{\theta \in E} MAE(\theta) \tag{7}$$

where E denotes the set of all parameter vectors evaluated using the full VO simulation, namely the initial LHS set together with the SSGA-selected vector for each category.

The search was terminated using an explicit convergence criterion: optimization stopped once the best objective value remained unchanged for five consecutive generations, indicating that the surrogate-assisted search had stabilized. This criterion, rather than visual inspection of the loss curve, governed the number of generations reported for each category.

7 Results and discussion

7.1 Recalibration framework performance

The performance of the recalibration framework was evaluated using statistical fit indicators, error reduction metrics, and convergence diagnostics. The objective was to assess how efficiently the optimization framework achieved stable parameter solutions. Table 1 summarizes the performance of the recalibration framework across the five categories. All simulations and optimization procedures were executed on a workstation equipped with an Intel Xeon® W-1270 CPU @ 3.40 GHz and 32 GB RAM. Each full STELARS run (2011 to 2024) required an average of 4 minutes and 26 seconds. To improve scalability, the VO module was vectorized to eliminate iterative for-loops, substantially reducing runtime during repeated evaluations within the optimization process.

Table 1. Summary of performance across recalibration categories

Recalibration Category Number	Category Type	Number of Parameters	Parameter Sets for Surrogate Training	Final Surrogate R^2	Average Generations for Convergence	Naïve Transfer MAE	Calibrated Transfer MAE
C1	Household Vehicle Ownership	51	765	0.8587	144	0.081	0.019
C2	Vehicle Body Type	89	1335	0.8800	288	0.089	0.018
C3	Vehicle Fuel Type	18	270	0.8228	56	0.069	0.018
C4	Vehicle Vintage Type	35	525	0.8529	99	0.215	0.036
C5	Vehicle Technology Type	22	330	0.8540	83	0.123	0.013

The final surrogate R^2values ranged from 0.8228 to 0.8800, indicating that more than 82% of the variation in simulation-based MAE is explained by the surrogate models across all categories. The body-type model achieved the highest R^2(0.8800), while the fuel-type model showed the lowest (0.8228), reflecting its smaller training sample (18 parameters, 270 LHS parameter sets). Each LHS parameter set represented one training observation for the RF surrogate, where the input is the parameter vector and the output is the MAE computed from a full STELARS simulation. A 70/30 train-test split was applied at the level of parameter sets, with the 30% test partition held out entirely from training. The Random Forest surrogate used a combination of standard and tuned hyperparameters: most settings were kept at their standard values, while key hyperparameters (such as the number of trees and maximum tree depth) were selected using 5-fold cross-validation performed on the training partition only. No tuning involved the test partition, so the reported R^2 values reflect genuine out-of-sample performance:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(\mathrm{MAE}_{\mathrm{norm},i}^{\mathrm{true}} - \mathrm{MAE}_{\mathrm{norm},i}^{\mathrm{pred}}\right)^2}{\sum_{i=1}^{n}\left(\mathrm{MAE}_{\mathrm{norm},i}^{\mathrm{true}} - \overline{\mathrm{MAE}_{\mathrm{norm}}^{\mathrm{true}}}\right)^2} \tag{8}$$

where $MAE_{norm,i}^{true}$ denotes the true normalized MAE for test vector i, $MAE_{norm,i}^{pred}$ the surrogate-predicted normalized MAE, and the barred term the mean true normalized MAE in the test set.

Substantial reductions in calibration error were achieved across all categories. The household vehicle ownership model reduced MAE from 0.081 to 0.019 (77% reduction). The body-type model reduced MAE from 0.089 to 0.018 (80% reduction). Fuel-type error decreased from 0.069 to 0.018 (74% reduction), vintage from 0.215 to 0.036 (83% reduction), and technology from 0.123 to 0.013 (89% reduction).

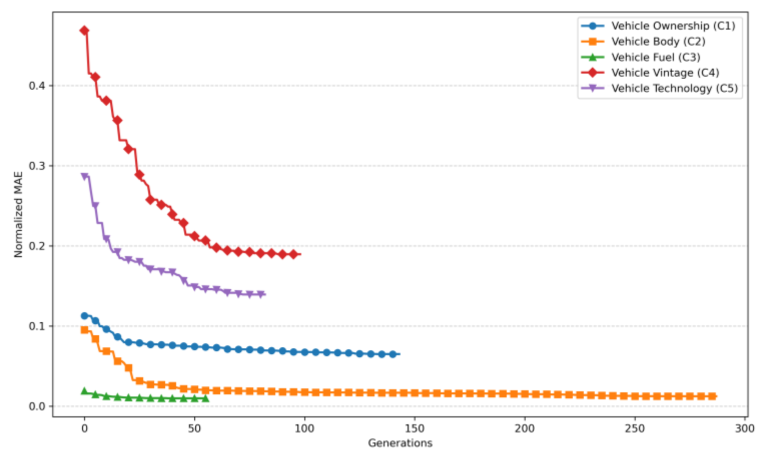

Figure 11. Convergence of normalized MAE across generations

A consistent decline in the evolution of normalized MAE across generations for all categories is observed in Figure 11, indicating stable improvement in fitness as the SSGA explores the parameter space. Simpler categories such as fuel (Category 3) and technology (Category 5) converged rapidly (within 56 and 83 generations, respectively). In contrast, higher-dimensional categories such as body type (89 parameters) required 288 generations, reflecting increased search complexity. The vintage model exhibited the highest initial MAE and a more gradual decline, suggesting a more irregular fitness landscape. Notably, the RF maintained sufficient predictive reliability throughout the optimization process, and no oscillatory or unstable convergence patterns were observed. The number of generations reported for each category corresponds to the five-generation stopping criterion. Apparent early flattening of the normalized MAE curves does not indicate that the criterion had been met, since small improvements in the best objective value continued beyond the visual plateau until it remained unchanged for five consecutive generations, particularly in the higher-dimensional body-type category. Categories that converged earlier, such as household vehicle ownership and fuel type, were terminated at their respective convergence generations.

7.2 Verification of the recalibrated model

Figure 12 compares the simulated and observed household-level vehicle ownership distributions in the GVA for 2024. The recalibrated model demonstrates strong agreement across all ownership categories, with deviations generally within 2–3 percentage points. Households with no vehicles account for 12.16% in the simulation compared to 14.52% observed, while the dominant single-vehicle category is closely reproduced at 48.65% simulated versus 45.60% observed. Two-vehicle households are similarly well aligned, representing 32.01% in the simulation compared to 31.27%

observed. The three-or-more vehicle category shows a minor underestimation, with 7.18% simulated relative to 8.61% observed.

Figures 13 to 16 further verify the performance of the recalibrated vehicle attribute models across body type, fuel type, vintage, and technology dimensions. For body type, compact (28.37% simulated versus 30.45% observed) and SUV vehicles (51.74% versus 54.11%) dominate both datasets, jointly accounting for over 80% of the fleet. Minor overestimations are observed for midsize/large vehicles (11.48% versus 8.83%) and subcompact vehicles (8.19% versus 6.34%), while the truck/van category remains negligible. Fuel type distributions show strong consistency, with gasoline vehicles comprising 87.98% simulated compared to 85.22% observed, diesel vehicles nearly identical at approximately 1.2%, and alternative fuel vehicles slightly underestimated at 10.84% versus 13.53%. The vintage distribution similarly aligns well, with older vehicles accounting for 66.47% simulated compared to 71.92% observed, new vehicles 15.94% versus 13.24%, and used vehicles 17.59% versus 14.84%. Finally, technology adoption is closely replicated, with advanced-technology-equipped vehicles representing 93.35% in the simulation and 92.08% observed, and non-equipped vehicles 6.65% versus 7.92%. The overall close correspondence across categories indicates that the recalibrated VO module effectively captures the underlying household vehicle ownership structure in the GVA.

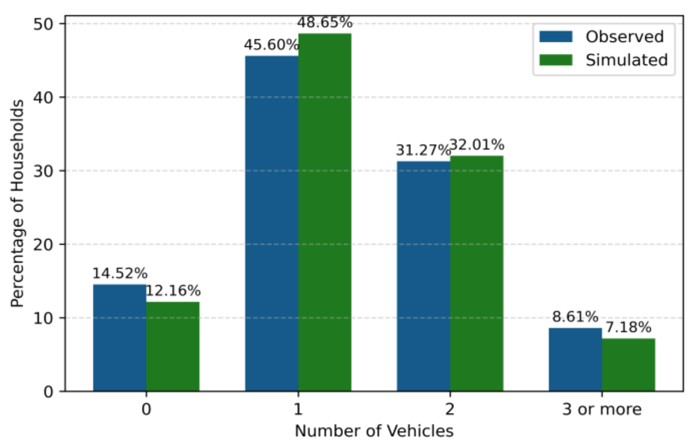

Figure 12. Household-level vehicle ownership distribution comparison after recalibration

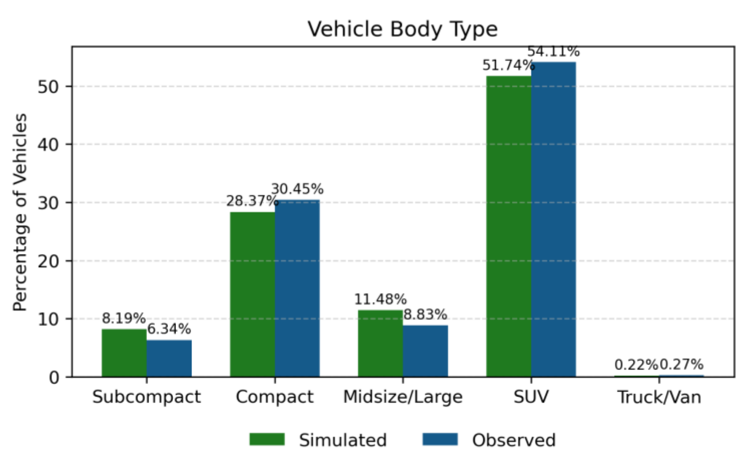

Figure 13. Vehicle body type distribution comparison after recalibration

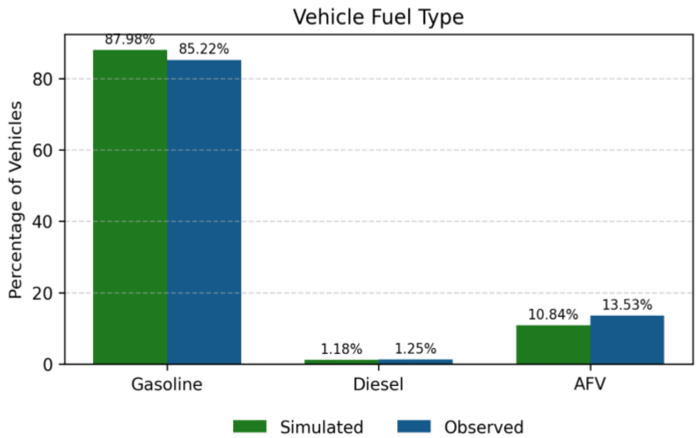

Figure 14. Vehicle fuel type distribution comparison after recalibration

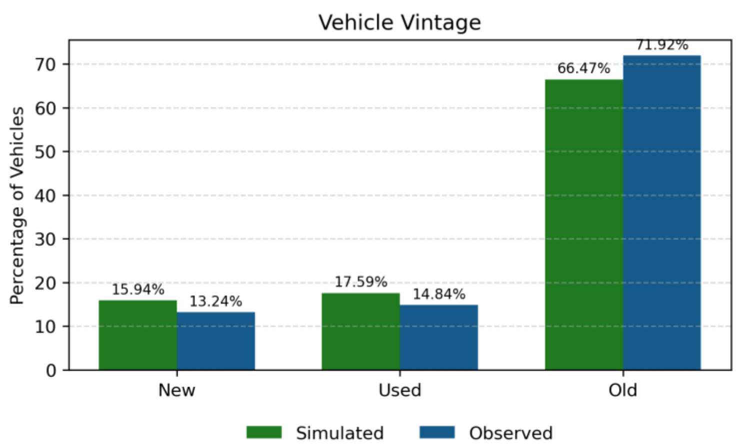

Figure 15. Vehicle vintage type distribution comparison after recalibration

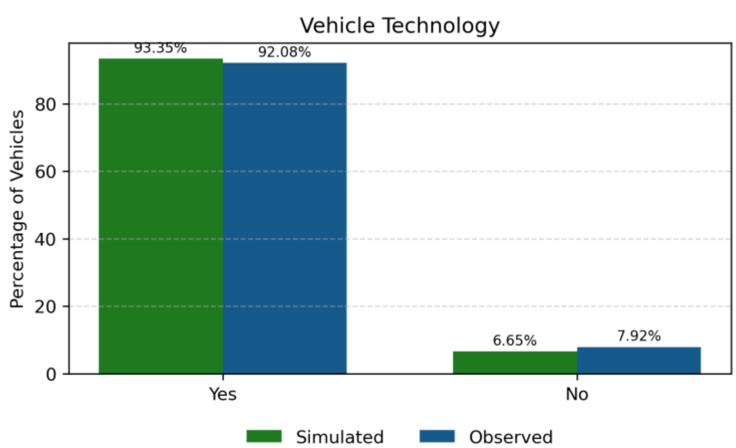

Figure 16. Vehicle technology type distribution comparison after recalibration

7.3 Recalibrated parameter comparison

To evaluate how behavioral sensitivities changed following recalibration, Figures 17 to 21 compare the original Okanagan parameter estimates with the recalibrated values for the GVA across the five VO categories. In these figures, blue bars represent the

Okanagan parameters, black dots indicate the recalibrated values, dashed lines denote a ±75% diagnostic band around the original estimates, and red crosses identify parameters exceeding this threshold. The ±75% interval is not a confidence interval and carries no statistical meaning. It is a fixed diagnostic reference positioned at the midpoint between the discarded ±50% pilot range and the ±100% search bound, used only to flag parameters that moved substantially toward the boundary and to visualize how far the recalibrated values shifted from their original estimates. The operational optimization bounds were ±100% with sign preservation.

Parameters are denoted using the notation CXPY, where X identifies the recalibration category (1 to 5, as defined in Table 1 and labeled C1 to C5 in Figure 11) and Y is the index of an individual parameter within that category. Each parameter corresponds to a specific coefficient in the underlying VO choice models. These coefficients fall into a few broad groups: alternative-specific constants, life-cycle event variables (such as birth of a child, residential move, or job change), household socio-economic variables (such as income, number of adults, and number of children), built-environment and accessibility variables (such as land-use index, transit access, and distance to urban core), and dwelling and prior-ownership variables. For example, C2P40 denotes the 40th parameter in Category 2 (Vehicle Body Type). It is the alternative-specific constant for the SUV alternative in the first-vehicle body type choice model, applicable to the Year 1 to 2 simulation timing segment, relative to the starting year of the simulation. Several alternative-specific constants in the original specification are defined separately for ranges of simulation years (Year 1 to 2, Year 3, Year 4, and Year 5 and onwards) to allow baseline propensities to vary over the simulation horizon, and these segmented positions are retained and recalibrated here. The complete set of recalibrated parameters, with descriptions and both original and recalibrated values, is provided in the Appendix.

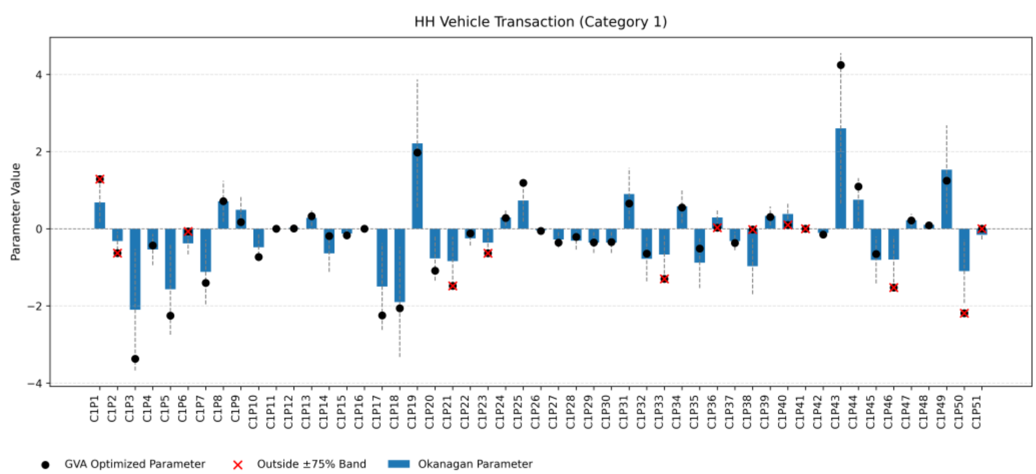

Figure 17. Household vehicle ownership recalibrated parameter comparison

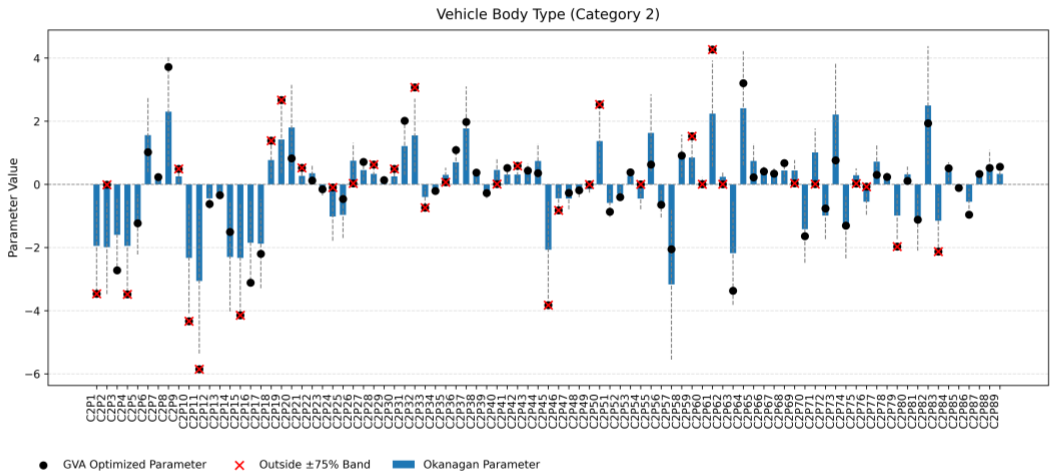

Figure 18. Vehicle body type recalibrated parameter comparison

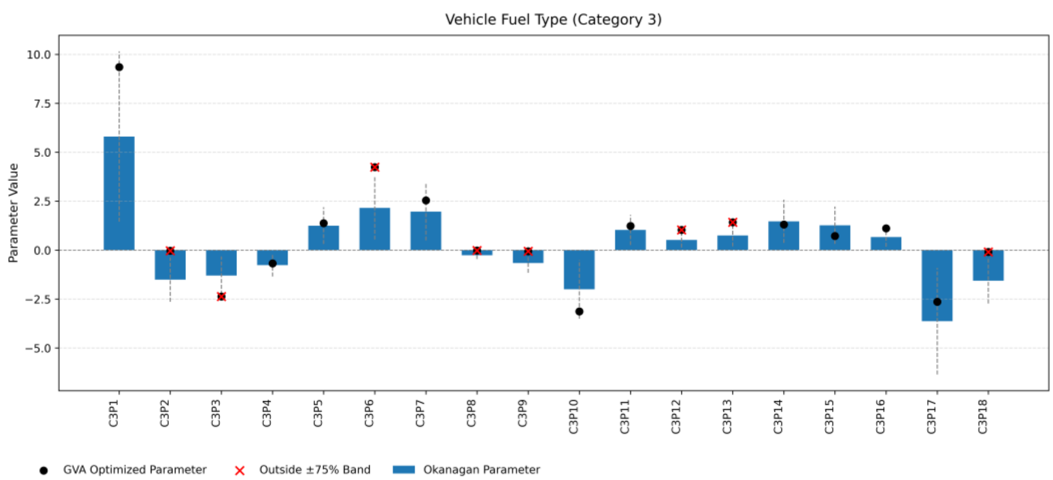

Figure 19. Vehicle fuel type recalibrated parameter comparison

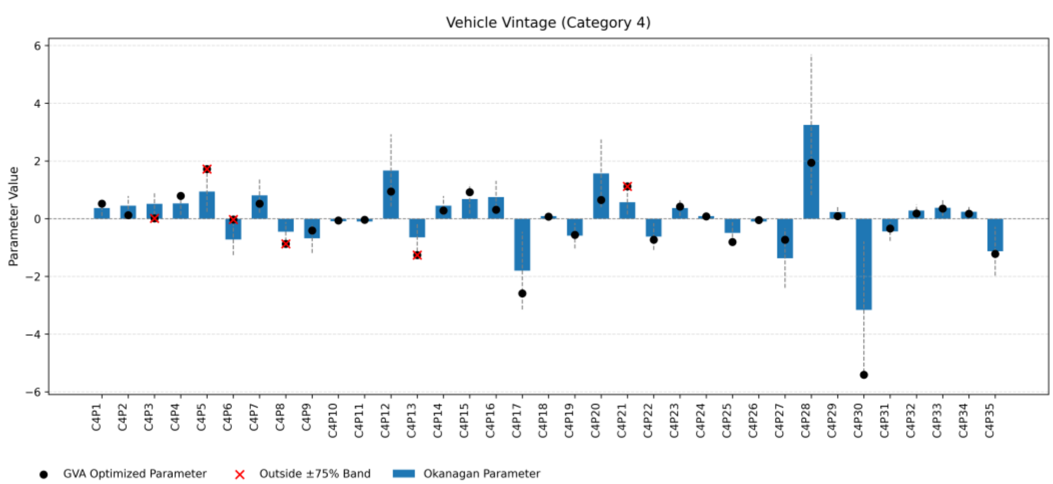

Figure 20. Vehicle vintage type recalibrated parameter comparison

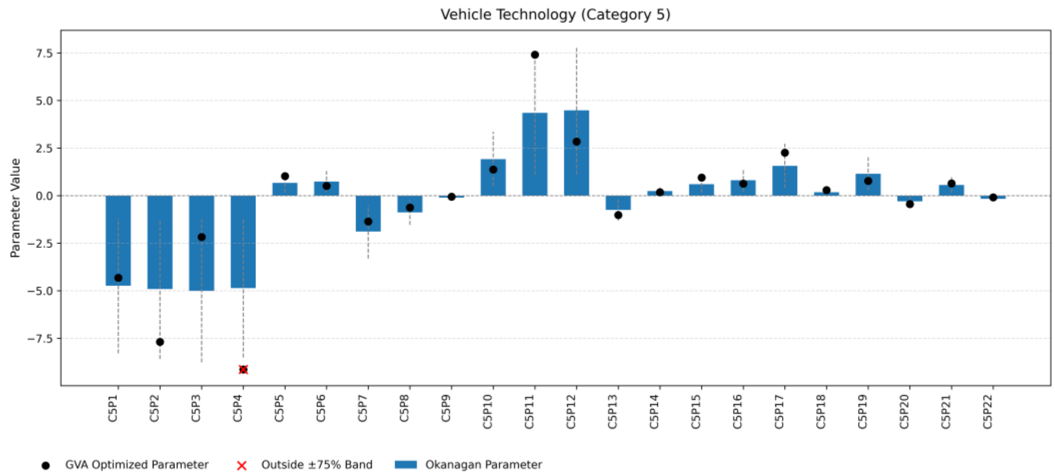

Figure 21. Vehicle technology type recalibrated parameter comparison

Across categories, most parameters remain within the ±75% interval, indicating substantial structural stability in behavioral relationships. The Household Vehicle Ownership (Figure 17) and Vehicle Vintage (Figure 20) categories exhibit relatively moderate deviations, suggesting that core ownership dynamics and fleet aging patterns are broadly transferable. In contrast, Vehicle Body Type (Figure 18) and Vehicle Fuel Type (Figure 19) show a higher concentration of parameters approaching or exceeding the ±75% band, reflecting stronger contextual differences in vehicle market composition and fuel preferences between the two regions. The Vehicle Technology category (Figure 21) displays targeted adjustments in parameters associated with advanced technology adoption, consistent with region-specific diffusion patterns.

Importantly, no parameter reached the ±100% search boundary, and all sign constraints were preserved, maintaining theoretical interpretability. The recalibration framework further mitigates overfitting risks by minimizing the per-category MAE objective and validating the surrogate model using a 70/30 train-test split with strong out-of-sample performance.

7.4 Recalibrated parameter interpretation

Across all categories, the recalibrated parameters remain mostly behaviorally meaningful and directionally consistent with the original model specification. As mentioned before, no coefficient reached the ±100% search boundary, and all sign constraints were preserved, which confirms that the underlying structural relationships governing vehicle ownership decisions remain valid across regions, while the strength of these relationships adapts to reflect local contextual conditions.

The most pronounced adjustments occur in the vehicle body type and fuel type categories. In the GVA, parameters associated with compact and midsize vehicle preferences exhibit stronger effects relative to the source region, reflecting higher urban density, constrained parking supply, and greater transit availability. Conversely, preferences associated with larger vehicles are comparatively moderated. These shifts are consistent with urban transportation literature demonstrating that dense built environments and multimodal accessibility reduce the attractiveness of larger vehicle classes. Similarly, coefficients associated with AFV adoption increase in magnitude in the metropolitan region, reflecting higher market maturity, greater exposure to charging infrastructure, and stronger policy incentives supporting zero-emission vehicles.

Adjustments are also evident in variables related to household composition and life-cycle dynamics. Parameters linked to children, employment transitions, and residential mobility display altered magnitudes, indicating that vehicle acquisition and disposal decisions are more sensitive to life-course changes in the metropolitan context. The recalibrated coefficients on accessibility and travel-related variables, including commute distance, proximity to transit, and access to urban amenities, generally increased in magnitude in the metropolitan context, while coefficients on dwelling type and baseline household characteristics were comparatively moderated. These magnitude changes align with the denser, multimodal character of the GVA, but they represent optimization-derived adjustments to existing coefficients rather than estimated behavioral effects. Importantly, many parameters exhibit moderate adjustments rather than extreme deviations, suggesting partial transferability of behavioral relationships across regions. Core determinants such as income categories and built-environment indicators remain influential in both contexts, though their magnitudes adjust to reflect differences in affordability, urban structure, and market conditions. Larger deviations are concentrated in vehicle type and technology-related parameters, indicating that these dimensions are most context-sensitive and closely tied to regional market maturity and infrastructure.

8 Conclusions

This study proposes a data-driven recalibration framework for enhancing the spatial transferability of the VO module within the agent-based IUM system, STELARS. Rather than treating model transfer as a binary choice between naïve application and full re-estimation, surrogate-assisted recalibration is positioned as a structured intermediate strategy for adapting complex behavioral systems across urban contexts using only aggregate data.

Applied to the Okanagan-to-GVA transfer, the framework reproduced the observed household vehicle ownership distribution within about four percentage points and reduced calibration error by 74% to 89% across the five recalibration categories. No parameter reached the ±100% search bound and all behavioral signs were preserved, so the existing model structure remained adequate and contextual differences were mostly captured through changes in parameter magnitude rather than structural reformulation. The built-environment, land-use, and accessibility inputs were the actual GVA values, supplied through the GVA simulation input files, so the framework recalibrated the Okanagan coefficients against genuine target-region conditions rather than carrying over source-region environments. Notably, these are optimization-derived adjustments made to match aggregate targets, not separately estimated behavioral sensitivities, and they should not be interpreted as causal effects.

The recalibration also surfaced contextual drivers absent from the original specification, such as parking constraints and higher insurance and operating costs, which the model partially absorbed through parameter magnitude. This points to the limits of recalibration when key explanatory variables are missing. Recalibration is therefore appropriate when behavioral mechanisms are broadly similar across regions and parameters adjust within reasonable bounds, whereas structural contextual differences may require re-specification through new variables or revised functional forms.

Several limitations remain. The framework was demonstrated on a single transfer case, so generalizability across urban systems is untested, and the recalibrated coefficients, while directionally consistent and empirically aligned, are not validated behavioral parameters. In addition, because the recalibration targets were regional aggregate distributions, model performance is assessed at the regional level, and evaluation at a finer spatial scale, such as the municipality level, was outside the scope of this study. Future research should integrate partial disaggregate data to strengthen interpretability, incorporate omitted contextual variables where needed, and conduct cross-regional validation across diverse spatial and policy settings, as well as assess model performance at finer spatial scales such as the municipality level.

Acknowledgments

The authors thank the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant program and the Environment and Climate Change Canada Climate Action and Awareness Fund for their financial support.

Author contribution

The authors confirm their contribution to the paper as follows: study conception and design: I. Hoque, M. Fatmi, M. M. Orvin, M. A. Khalil; data collection: M. Fatmi; analysis and interpretation of results: I. Hoque, M. Fatmi, M. M. Orvin, M. A. Khalil; draft manuscript preparation: I. Hoque, M. Fatmi, M. M. Orvin, M. A. Khalil. All authors reviewed the results and approved the final version of the manuscript.

Data and code availability

This study was conducted as part of a broader research project. The BC ATUS data are subject to participant confidentiality and ethics requirements. This dataset, the STELARS model, the recalibration framework, and all associated code developed in this study will be released publicly as part of the project's deliverables, in accordance with applicable confidentiality and ethics requirements.

Appendix

Appendix available as a supplemental file at https://doi.org/10.5198/jtlu.2026.2867.

References

Afzal, A., Kim, K. Y., & Seo, J. W. (2017). Effects of Latin hypercube sampling on surrogate modeling and optimization. *International Journal of Fluid Machinery and Systems*, *10*(3), 240–253. https://doi.org/10.5293/IJFMS.2017.10.3.240

Agyemang-Duah, K., & Hall, F. L. (1997). Spatial transferability of an ordered response model of trip generation. *Transportation Research Part A: Policy and Practice*, *31*(5), 389–402. https://doi.org/10.1016/S0965-8564(96)00035-3

Arentze, T., Hofman, F., Van Mourik, H., & Timmermans, H. (2002). Spatial transferability of the Albatross model system: Empirical evidence from two case studies. *Transportation Research Record*, *1805*, 1–7. https://doi.org/10.3141/1805-01

Atherton, T. J., & Ben-Akiva, M. (1976). Transferability and updating of disaggregate travel demand models. *Transportation Research Record*, *610*, 12–18. https://trid.trb.org/View/53666

Badoe, D. A., & Miller, E. J. (1995). Analysis of temporal transferability of disaggregate work trip mode choice models. *Transportation Research Record, 1493*, 1–11. https://trid.trb.org/View/452588

Badoe, D. A., & Miller, E. J. (1998). Modelling mode choice with data from two independent cross-sectional surveys: An investigation. *Transportation Planning and Technology*, *21*(4), 235–261. https://doi.org/10.1080/03081069808717611

BC Assessment. (2025). *BC assessment data advice and inventory extracts, 2016–2022* (Version 6) [Data set]. Abacus Data Network.

Ben-Akiva, M., & Bolduc, D. (1987). Approaches to model transferability and updating: The combined transfer estimator. *Transportation Research Record, 1139,* 1–7.

Ben-Akiva, M., & Morikawa, T. (1990). Estimation of switching models from revealed preferences and stated intentions. *Transportation Research Part A: General*, *24*(6), 485–495. https://doi.org/10.1016/0191-2607(90)90037-7

Bowman, J. L., & Bradley, M. (2017). Testing spatial transferability of activity-based travel forecasting models. *Transportation Research Record*, *2669*, 62–71. https://doi.org/10.3141/2669-07

Bowman, J. L., Bradley, M., Castiglione, J., & Yoder, S. L. (2014). Making advanced travel forecasting models affordable through model transferability. Paper presented at the 93rd Annual Meeting of the Transportation Research Board, Washington, DC.

Bradley, M. A., & Daly, A. J. (1991, May). Estimation of logit choice models using mixed stated preference and revealed preference information. Paper presented at the 6th International Conference on Travel Behaviour, Quebec, Canada.

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Deb, K., & Agrawal, R. B. (1995). Simulated binary crossover for continuous search space. *Complex Systems*, *9*(2), 115–148.

Fatmi, M. R., & Ahmed, M. (2025). Investigating activity, travel and time use patterns by work arrangements: Evidence from the BC ATUS – Wave 1. Paper presented at the 104th Annual Meeting of the Transportation Research Board, Washington, DC.

Fatmi, M. R., & Habib, M. A. (2015). Spatial transferability of a microresidential mobility model in the Integrated Land Use, Transportation, and Environment modeling system. *Transportation Research Record*, *2494*, 29–36. https://doi.org/10.3141/2494-04

Gunn, H. F., Ben-Akiva, M., & Bradley, M. A. (1985). Tests of the scaling approach to transferring disaggregate travel demand models. *Transportation Research Record*, *1037,* 21–30. https://trid.trb.org/view/272223

Hoque, I. (2025). *Data-driven recalibration methodology for spatially transferring the vehicle ownership module of an agent-based integrated urban model* [Master's thesis, University of British Columbia]. UBC Theses and Dissertations. https://doi.org/10.14288/1.0451067

Hossain, M. S., Fatmi, M. R., & Enam, A. (2026). What type of vehicles do households own? A joint model for vehicle body, vintage, fuel, and advanced driving assistive systems. *Transportation Letters,* 1–23. https://doi.org/10.1080/19427867.2026.2669759

Hossain, M. S., Fatmi, M. R., & Khalil, M. A. (2026). Microsimulation of vehicle ownership decisions within an agent-based integrated urban model: An event-based hybrid of continuous and discrete simulation approach. *Journal of Transport and Land Use*, *19*(1), 79–106. https://doi.org/10.5198/jtlu.2026.2665

Insurance Corporation of British Columbia. (2025). *Vehicle population dataset* [Data set]. https://public.tableau.com/app/profile/icbc/viz/VehiclePopulationIntroPage/VehiclePopulationData

Jin, Y. (2011). Surrogate-assisted evolutionary computation: Recent advances and future challenges. *Swarm and Evolutionary Computation*, *1*(2), 61–70. https://doi.org/10.1016/j.swevo.2011.05.001

Karasmaa, N. (2007). Evaluation of transfer methods for spatial travel demand models. *Transportation Research Part A: Policy and Practice*, *41*(5), 411–427. https://doi.org/10.1016/j.tra.2006.09.009

Karasmaa, N., & Pursula, M. (1997). Empirical studies of transferability of Helsinki metropolitan area travel forecasting models. *Transportation Research Record*, *1607*, 38–44. https://doi.org/10.3141/1607-06

Kavta, K., Bösehans, G., Bell, M. C., Liao, F., de Almeida Correia, G. H., & Dissanayake, D. (2024). Assessing the spatial transferability of mode choice models: A case of shared electric mobility hubs (eHUBS) in Amsterdam and Manchester. *Transport Policy*, *156*, 101–111. https://doi.org/10.1016/j.tranpol.2024.07.017

Khaddar, S., Varghese, V., Fatmi, M. R., & Chikaraishi, M. (2024). Jointly modelling activity start time, travel mode, companionship, and destination location choices. *Transportmetrica A: Transport Science*, *22*(1). https://doi.org/10.1080/23249935.2024.2372025

Khalil, M. A. (2025). *Towards a next generation of integrated urban models: Advanced machine learning and microsimulation modelling for transportation and urban systems* [Doctoral dissertation, University of British Columbia]. UBC Theses and Dissertations. https://doi.org/10.14288/1.0449896

Khalil, M. A., & Fatmi, M. R. (2023). Modeling daily in-home activities using machine learning techniques. *Travel Behaviour and Society*, *31*, 374–385. https://doi.org/10.1016/j.tbs.2023.02.001

Khalil, M. A., Fatmi, M. R., & Orvin, M. M. (2025). Developing and microsimulating demographic dynamics for an integrated urban model: A comparison between logistic regression and machine learning techniques. *Transportation*, *52*(4), 1621–1655. https://doi.org/10.1007/s11116-024-10468-7

Koppelman, F. S., Kuah, G.-K., & Wilmot, C. G. (1985). Transfer model updating with disaggregate data. *Transportation Research Record, 1037*, 102–107.

Koppelman, F. S., & Rose, G. (1985). Geographic transfer of travel choice models: Evaluation and procedures. In B. G. Hutchinson, P. Nijkamp, & M. Batty (Eds.), *Optimization and discrete choice in urban systems* (pp. 272–309). Springer. https://doi.org/10.1007/978-3-642-51020-5_16

Koppelman, F. S., & Wilmot, C. G. (1982). Transferability analysis of disaggregate choice models. *Transportation Research Record*, *895,* 18–24. https://trid.trb.org/View/189355

Koppelman, F. S., & Wilmot, C. G. (1986). The effect of omission of variables on choice model transferability. *Transportation Research Part B: Methodological*, *20*(3), 205–213. https://doi.org/10.1016/0191-2615(86)90017-2

Linh, H. T., Adnan, M., Ectors, W., Kochan, B., Bellemans, T., & Tuan, V. A. (2019). Exploring the spatial transferability of FEATHERS — an activity based travel demand model — for Ho Chi Minh City, Vietnam. *Procedia Computer Science*, *151*, 226–233. https://doi.org/10.1016/j.procs.2019.04.033

McKay, M. (1992). Latin hypercube sampling as a tool in uncertainty analysis of computer models. *Proceedings of the 24th Conference on Winter Simulation* (pp. 557–564). Association for Computing Machinery. https://www.osti.gov/biblio/7196246

Mohammadian, A., & Miller, E. J. (2003). Dynamic modeling of household automobile transactions. *Transportation Research Record*, *1831*, 98–105. https://doi.org/10.3141/1831-11

Nohekhan, A., Samimi, A., & Zahedian, S. (2022). Spatial transferability of a daily activity type and duration MDCEV model. *Civil Engineering Infrastructures Journal*, *55*(1), 161–181. https://doi.org/10.22059/CEIJ.2021.313889.1721

Orvin, M. M., Enam, A., & Fatmi, M. R. (2024). A joint model of residential mobility and location choice decisions. *International Journal of Urban Sciences*. https://doi.org/10.1080/12265934.2024.2358415

Orvin, M. M., & Fatmi, M. R. (2024). A residential location search model based on the reasons for moving out. *Transportation Letters*, *16*(6), 566–580. https://doi.org/10.1080/19427867.2023.2222990

Orvin, M. M., Fatmi, M. R., & Hoque, I. (2025). Housing price dynamics within an integrated urban model: Testing the spatial and temporal transferability. *International Journal of Urban Sciences*. https://doi.org/10.1080/12265934.2025.2517097

Rahman, M. N., & Fatmi, M. R. (2023). Population synthesis accommodating heterogeneity: A Bayesian network and generalized raking technique. *Transportation Research Record*, *2677*(6), 41–57. https://doi.org/10.1177/03611981221144289

Rashidi, T. H., Auld, J., & Mohammadian, A. (2013). Effectiveness of Bayesian updating attributes in data transferability applications. *Transportation Research Record, 2344*, 1–9. https://doi.org/10.3141/2344-01

Roorda, M. J., Mohammadian, A., & Miller, E. J. (2000). Toronto area car ownership study: A retrospective interview and its applications. *Transportation Research Record, 1719,* 69–76. https://doi.org/10.3141/1719-09

Salvini, P., & Miller, E. J. (2005). ILUTE: An operational prototype of a comprehensive microsimulation model of urban systems. *Networks and Spatial Economics*, *5*(2), 217–234. https://doi.org/10.1007/s11067-005-2630-5

Santoso, D. S., & Tsunokawa, K. (2005). Transferability and updating analysis of mode choice model between two developing countries: Predictive performance point of view. *Journal of the Eastern Asia Society for Transportation Studies*, 6, 173–185. https://doi.org/10.11175/easts.6.173

Santoso, D. S., & Tsunokawa, K. (2010). Comparison of updating techniques in transferability analysis of work trip mode choice models in developing countries. *Journal of Advanced Transportation*, *44*(2), 89–102. https://doi.org/10.1002/atr.112

Sikder, S., Augustin, B., Pinjari, A. R., & Eluru, N. (2014). Spatial transferability of tour-based time-of-day choice models. *Transportation Research Record*, *2429*, 99–109. https://doi.org/10.3141/2429-11

Sikder, S., & Pinjari, A. R. (2013). Spatial transferability of person-level daily activity generation and time use models. *Transportation Research Record*, *2343*, 95–104. https://doi.org/10.3141/2343-12

Sikder, S., Pinjari, A. R., Srinivasan, S., & Nowrouzian, R. (2013). Spatial transferability of travel forecasting models: A review and synthesis. *International Journal of Advances in Engineering Sciences and Applied Mathematics, 5*(2), 104–128. https://doi.org/10.1007/s12572-013-0090-6

Tang, L., Xiong, C., & Zhang, L. (2018). Spatial transferability of neural network models in travel demand modeling. *Journal of Computing in Civil Engineering*, *32*(3), 04018010. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000752

UBC integrated Transportation Research Laboratory. (2022). *Travel technology and mobility survey (TTMS)*. https://uitr.ok.ubc.ca/2022/04/28/travel-technology-and-mobility-survey-ttms/

UBC integrated Transportation Research Laboratory. (2024). *British Columbia activity time use survey* [Data set]. https://survey.uitrlab.ok.ubc.ca/v2.0/

Waddell, P. (2002). UrbanSim: Modeling urban development for land use, transportation, and environmental planning. *Journal of the American Planning Association*, *68*(3), 297–314. https://doi.org/10.1080/01944360208976274

Walker, W. T., Brady, S. H., & Taylor, C. (1997). Updating existing travel simulation models with small-sample survey data using parameter scaling methods. *Transportation Research Record, 1607,* 55–61. https://doi.org/10.3141/1607-08

Wang, H., & Jin, Y. (2020). A random forest-assisted evolutionary algorithm for data-driven constrained multiobjective combinatorial optimization of trauma systems. *IEEE Transactions on Cybernetics*, *50*(2), 536–549. https://doi.org/10.1109/TCYB.2018.2869674

Whitley, D. (1994). A genetic algorithm tutorial. *Statistics and Computing*, *4*(2), 65–85. https://doi.org/10.1007/BF00175354

Willmott, C. J., & Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research*, *30*(1), 79–82. https://doi.org/10.3354/cr030079

Xu, C., Wang, W., Liu, P., Guo, R., & Li, Z. (2014). Using the Bayesian updating approach to improve the spatial and temporal transferability of real-time crash risk prediction models. *Transportation Research Part C: Emerging Technologies*, *38*, 167–176. https://doi.org/10.1016/j.trc.2013.11.020

Yasmin, F., Morency, C., & Roorda, M. J. (2015). Assessment of spatial transferability of an activity-based model, TASHA. *Transportation Research Part A: Policy and Practice*, *78*, 200–213. https://doi.org/10.1016/j.tra.2015.05.008

Zhang, Y., & Mohammadian, A. (2008). Bayesian updating of transferred household travel data. *Transportation Research Record*, *2049,* 111–118. https://doi.org/10.3141/2049-13

Zheng, Y., Fu, X., & Xuan, Y. (2019). Data-driven optimization based on random forest surrogate. *2019 6th International Conference on Systems and Informatics,* (pp. 487–491). ICSAI. https://doi.org/10.1109/ICSAI48974.2019.9010547