

Article

Not peer-reviewed version

Advancing Rural Mobility: Identifying Operational Determinants for Effective Autonomous Road-Based Transit

[Shenura Jayatilleke](#) , [Ashish Bhaskar](#) , [Jonathan Bunker](#) *

Posted Date: 26 August 2025

doi: 10.20944/preprints202508.1835.v1

Keywords: Sustainable Mobility; Public Transport; Shared Autonomous Vehicles; First Mile and Last Mile; General Additive Model; Extreme Gradient Boost; South-East Queensland



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Advancing Rural Mobility: Identifying Operational Determinants for Effective Autonomous Road-Based Transit

Shenura Jayatilleke, Ashish Bhaskar and Jonathan Bunker *

School of Civil and Environmental Engineering, Queensland University of Technology, 2 George Street, Brisbane, QLD 4000, Australia

* Correspondence: j.bunker@qut.edu.au

Abstract

Rural communities face persistent transport disadvantages due to low population density, limited-service availability, and high operational costs, restricting access to essential services and exacerbating social inequality. Autonomous public transport systems offer a transformative solution by enabling flexible, cost-effective, and inclusive mobility options. This study investigates the operational determinants for autonomous road-based transit systems in rural and peri-urban South-East Queensland, employing a structured survey of 273 residents and analytical approaches, including General Additive Model (GAM) and Extreme Gradient Boosting (XGBoost). Findings indicate that small shuttles suit flexible, non-routine trips, minibus shuttles enhance first-mile and last-mile connectivity, and standard-sized buses are optimal for high-capacity school and emergency transport. Hybrid models integrating autonomous and conventional buses are preferred, with autonomous taxis raising equity concerns. Multipurpose services and Mobility-as-a-Service (MaaS) integration will help improve efficiency of the system. Policy recommendations include deploying small shuttles in university and tourist areas, minibus shuttles in town centres with accessibility standards, and standard-sized buses for school transport with subsidies to ensure equitable access. These insights guide policymakers in designing autonomous transit systems to improve rural connectivity and quality of life.

Keywords: sustainable mobility; public transport; shared autonomous vehicles; first mile and last mile; general additive model; extreme gradient boost; south-east queensland

1. Introduction

The provision of efficient and accessible public transport in rural areas remains a persistent challenge, driven by low population densities, dispersed settlements, and constrained financial resources. These factors contribute to significant transport disadvantages, limiting access to essential services such as education, healthcare, employment, and social activities, which in turn exacerbates social isolation and inequality. In rural and remote regions, public transport options are often scarce, infrequent, or entirely absent, compelling residents to rely heavily on private vehicles, which may not be viable for all, particularly low-income individuals, seniors, or those with disabilities. Autonomous public transport systems, enabled by advancements in vehicle automation and artificial intelligence, offer a transformative opportunity to address these challenges. Unlike traditional fixed-route systems, autonomous vehicles can provide flexible, on-demand services tailored to the irregular and dispersed demand patterns characteristic of rural areas. By eliminating driver-related costs, which constitute a significant portion of public transport expenses, autonomous systems could reduce operational costs, enabling extended service hours, including 24/7 operations, and broader geographic coverage. However, the successful deployment of such systems requires a nuanced

understanding of user preferences and the operational determinants that influence their adoption in rural contexts, where transport needs differ markedly from urban environments.

Researchers such as Scheltes and de Almeida Correia (1), Lau and Susilawati (2), Roy, Dadashev (3), Gurumurthy, Kockelman (4), have primarily focused on Autonomous Demand Responsive Transit (ADRT) as the preferred service type, highlighting its prominence in their studies. The high flexibility of autonomous shuttle services has significantly contributed to their increased adoption. Successful implementation of these services hinges on effective service design and user acceptance. Although numerous studies have explored service design through simulations, there remains a critical need to incorporate public perception into service design to identify key predictors of operational success. This study addresses this critical knowledge gap by empirically investigating the operational determinants of autonomous public transport systems in rural and peri-urban communities of South-East Queensland.

Through a structured survey targeting residents in low-density areas, the research explores preferences for various autonomous vehicle types and service offerings. The analysis exhibits two complementary analytical approaches: General Additive Models (GAM) and Extreme Gradient Boosting (XGBoost). GAM is particularly suited for this study due to its ability to model non-linear relationships between predictors such as trip purpose, demographic groups, and land use and user preferences, which may exhibit complex, non-linear patterns. XGBoost, a powerful machine learning algorithm, excels in handling high-dimensional data and capturing intricate interactions among variables, enabling the identification of the most influential factors driving user preferences.

The significance of this research lies in its contribution to the emerging field of autonomous transport, particularly in rural settings where traditional public transport solutions have proven inadequate. By providing empirical evidence on user preferences across diverse demographic groups, trip purposes, and land-use contexts, this study offers actionable insights for policymakers, transport planners, and industry stakeholders. These insights can inform the design and implementation of autonomous public transport systems that are tailored to the unique needs of rural communities, thereby reducing transport disadvantage, enhancing access to essential services, and improving overall quality of life.

The remainder of the paper is divided into the following sections. Study methodology including the survey design, data collection and analysis in Section 2. Model estimation and validation results of GAM and XGBoost in Section 3. A comprehensive discussion, conclusion of the study and the policy implications are provided in Sections 4, 5 and 6. Finally, the research limitations and the future research directions are presented in Section 7.

2. Materials and Methods

This section details the methodology undertaken, focusing on survey design and data collection, data characteristics and processing, GAM, and XGBoost model fitting and validation.

2.1. Survey Design and Data Collection

To empirically identify the operational determinants for rural autonomous road-based transits, a structured questionnaire was developed. The survey design process was structured in three phases to ensure robustness and validity. First, a comprehensive literature review was conducted to identify and conceptualize key constructs, including trip purpose, land Use, demographic groups, vehicle types, and service offering. These references provided a theoretical foundation for understanding travel behavior, land use dynamics, vehicle suitability, and service models in rural and peri-urban settings. In the second phase, semi-structured interviews with domain experts from the Australian transport sector were conducted to validate the identified variables and ensure their relevance to rural environments (5). The survey instrument, encompassing measurement items and introductory content, was iteratively refined through consultations with an expert supervisory review panel to enhance content validity and contextual appropriateness (5, 6). Final list of study attributes is listed on Table 1.

Table 1. Study attributes.

Variable	Item	Description
Demand Drivers		
Trip Purpose (TP)	TP1	Work
	TP2	School
	TP3	University
	TP4	Shopping
	TP5	Medical
	TP6	Leisure
	TP7	Emergency
	TP8	Special events or gatherings
Demographic Group (DG)	DG1	School children
	DG2	University students
	DG3	Working professionals
	DG4	Senior citizens
	DG5	Tourists
	DG6	Leisure travellers
	DG7	People with physical disabilities
	DG8	People with sensory disabilities
	DG9	People with cognitive disabilities
	DG10	Low-income individuals
	DG11	Middle-income individuals
	DG12	High-income individuals
Built-Environment Factors		
Land Use (LU)	LU1	Residential neighbourhoods
	LU2	Industrial/ business parks
	LU3	University precincts
	LU4	Agricultural land areas
	LU5	Tourist Destinations
	LU6	Town Centres
Supply Features		
Vehicle Type (VT)	VT1	Small shuttle
	VT2	Minibus shuttle
	VT3	Standard-sized conventional bus
Service Offering (SO)	SO1	Completely replace conventional buses
	SO2	Operate as a connector to existing fixed-route bus services
	SO3	Connector to longer distance services
	SO4	Operate as private taxi services
	SO5	Accommodate as a multipurpose service
	SO6	Integrated with other transport offerings
	SO7	Operate 24/7

In the third phase, a pilot study was administered to Higher Degree Research (HDR) students, selected for their expertise in survey methodologies and analytical rigor. This pilot aimed to identify ambiguities, structural inconsistencies, and potential misinterpretations within the questionnaire. Feedback informed further refinements, improving the instrument’s clarity, interpretability, and face validity. The finalized questionnaire was approved by the University Human Research Ethics Committee (UHREC: 9083-HE09), ensuring compliance with ethical standards. The target population consisted of residents in rural and peri-urban communities within the South-East Queensland (SEQ) region, specifically in the Lockyer Valley, Moreton Bay, Sunshine Coast, Scenic Rim, and Redland local government areas, as identified by Mortoja and Yigitcanlar (7). To align with global rural

classification standards, only localities with a population density below 300 inhabitants per square kilometer were included (8).

Sample size determination was guided by population-based sampling adequacy, targeting a sample of 300 responses for a population exceeding 1,000,000, at a 95% confidence level with a 6% margin of error (9), sufficient for the exploratory analyses employed in this study. Data collection was facilitated by Qualtrics using a convenience random sampling approach to reach rural residents within the study area. A total of 357 responses were initially collected. The final sample size was 273, after a rigorous data cleaning process addressed four key issues: unengaged respondents, outliers, missing data, and multicollinearity (10, 11). The socio-demographic profile: gender, age, education level, employment status, and annual household income of the respondents is presented in Figure 1(a-e).

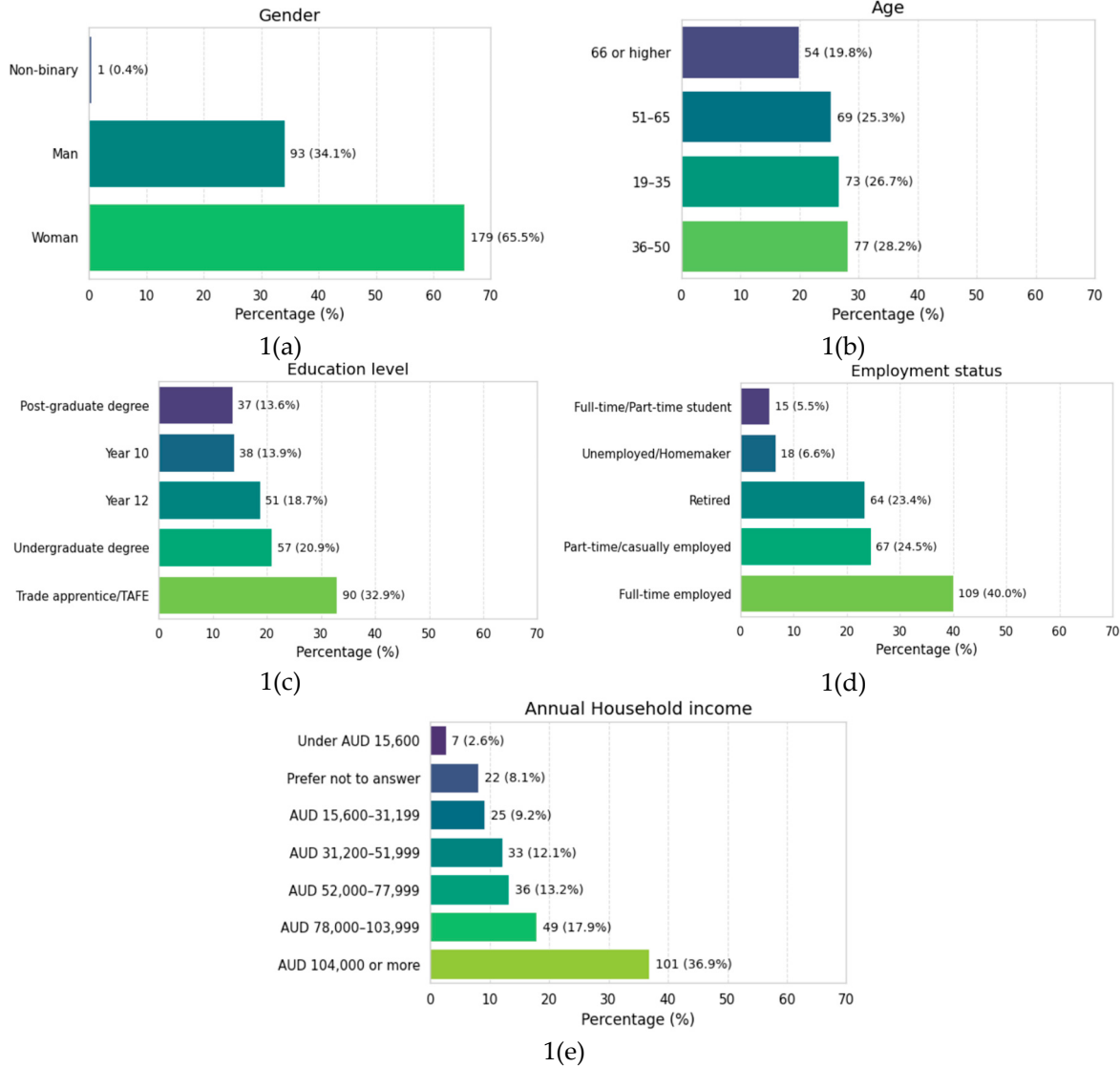


Figure 1. Socio-demographic profile (a) Gender. (b) Age. (c) Education level. (d) Employment status. (e) Annual household income.

The socio-demographic profile of the survey participants (N = 273) revealed a predominantly female sample, with women comprising the majority (n = 179, 65.5%), followed by men (n = 93, 34.1%) and a negligible non-binary representation (n = 1, 0.4%). Age distribution was relatively even across younger and middle-aged groups, peaking at 36–50 years (n = 77, 28.2%) and lowest among those aged 66 years or older (n = 54, 19.8%). Education levels were highest for trade/apprenticeship or TAFE qualifications (n = 90, 32.9%), while Year 10 (n = 38, 13.9%) and postgraduate degrees (n = 37, 13.6%) were the least common. Employment status showed full-time employment as dominant (n = 109,

40.0%), with full-time or part-time students representing the smallest group (n = 15, 5.5%). Household income was skewed towards higher brackets, with AUD 104,000 or more the most prevalent (n = 101, 36.9%) and under AUD 15,600 the least (n = 7, 2.6%). Internal consistency was evaluated using Cronbach’s Alpha, with values ranging from 0.705 to 0.984, indicating acceptable to excellent reliability across constructs (12, 13). Following the data cleaning, the dataset underwent further processing to verify the assumptions underlying the model fit, as denoted in Figure 2.

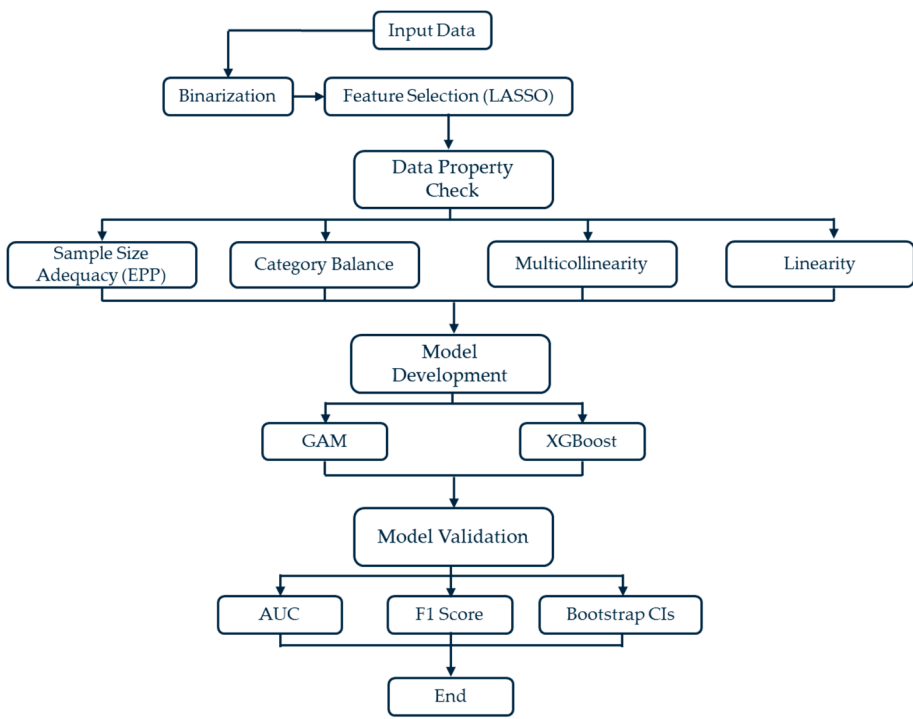


Figure 2. Data analysis process.

2.2. Data Characteristics and Processing

The study employed a sophisticated statistical pipeline for data processing, including binarization, feature selection, sample size adequacy, category balance, multicollinearity and linearity checks.

2.2.1. Binarization

The dependent variables are binarized to transform Likert-scale responses into binary outcome. Equation 1 represents the logistic regression model used to model binarized dependent variables. The logistic function was initially implemented and later transformed to incorporate non-linear effects of GAM.

$$P\left(Y=1|X\right)=\frac{1}{1+e^{-\left(\beta_0+\beta_1X_1+...+\beta_PX_P\right)}}$$

(1)

Where, β are the coefficients, and X are the predictors of the logistic function. Dependent variables are binarized at a threshold of ≥ 4 (coding 4–5 as 1, else 0), based on Norman (14), who argues that Likert-scale data can be treated as binary when high scores indicate agreement. Sensitivity analysis is conducted at thresholds of 3 and 5 to assess impacts on class balance, ensuring robustness. The default threshold of 4 is selected, considering the lower Akaike Information Criterion (AIC) value (15).

2.2.2. Feature Selection

Feature selection is performed using the Least Absolute Shrinkage and Selection Operator (LASSO) to identify relevant predictors, addressing the high dimensionality of the dataset (26 predictors). LASSO imposed an L1 penalty on logistic regression, shrinking irrelevant coefficients to zero (16).

Minimize $\left\{-\frac{1}{n}\sum_{i=1}^n[y_i\log(p_i)+(1-y_i)\log(1-p_i)+\lambda\sum_{j=1}^p|\beta_j|\right\}$

(2)

Where, $P_i = \frac{1}{1+e^{-(\beta_0+\beta_1X_1+\dots+\beta_pX_p)}}$, The negative log-likelihood (log-loss) for logistic regression is averaged over n observations. The L1 penalty shrinks the coefficients with λ controlling the penalty strength. Ten-fold cross-validation was conducted to select the optimal λ , maximising the Area Under the Curve (AUC). AUC is the probability that a randomly chosen positive instance ranks higher than a negative one, computed as the integral of the Receiver Operating Characteristic (ROC) curve (1, 2). LASSO's ability to handle multicollinearity and select sparse models makes it ideal for high-dimensional datasets. Predictors with non-zero coefficients were retained for each dependent variable as highlighted in Table 2 (2, 3). Supply side features were modelled as dependent variables, while demand drivers and built-environment factors were considered as independent variables.

Table 2. LASSO feature selection results.

Dependent Variable	Selected Variables
VT1	TP1, TP2, TP5, TP8, DG5, DG6, DG7, DG12, LU1, LU3
VT2	TP4, TP5, TP6, TP8, DG2, DG3, DG4, DG5, DG6, DG8, DG10, LU1, LU2, LU3, LU6
VT3	TP1, TP2, TP6, TP7, TP8, DG1, DG2, DG10, DG12, LU2, LU3, LU4, LU5, LU6
SO1	TP2, TP7, TP8, DG1, DG7, DG8, DG9, LU4, LU5, LU6
SO2	DG3, LU3
SO3	TP1, TP3, TP4, TP8, DG1, DG2, DG6, DG7, DG10, LU1, LU2, LU6
SO4	TP1, TP2, TP3, TP4, TP5, TP7, DG1, DG2, DG4, DG5, DG7, DG8, DG10, DG11, DG12, LU1, LU2, LU3, LU6
SO5	TP1, TP6, DG3, DG6, DG11, LU2, LU3, LU4
SO6	TP1, TP8, DG2, DG4, DG6, DG10, LU2, LU3, LU5
SO7	TP1, TP6, DG3, DG10, LU3, LU4

2.2.3. Sample Size and Category Balance Check

The sample size is evaluated to ensure sufficient Events Per Parameter (EPP) for stable model estimates, given the relatively smaller sample size (20). An $EPP \geq 10$ is recommended to avoid overfitting and ensure reliable estimates (20). This threshold was evaluated post-LASSO feature selection to account for the reduced number of predictors, aligning with recommendations for robust model estimation in smaller samples (21). Imbalanced classes can bias machine learning models toward the majority class, reducing predictive accuracy for minority classes (22). Category balance is assessed by calculating the minority class count, and a minimum of 20 minority cases is required to ensure balanced classes, as suggested by King and Zeng (22).

2.2.4. Multicollinearity Check

Multicollinearity among predictors can inflate standard errors and destabilize model estimates. Variance Inflation Factors (VIF) are calculated to quantify multicollinearity as follows:

$$VIF_j = \frac{1}{1 - R_j^2}$$

(3)

Where, R_j^2 is the coefficient determinations from regressing predictor j on all other predictors, with $VIF > 5$ indicating redundancy (4).

2.2.5. Linearity Check

The linearity assumption of logistic regression is tested by comparing the AIC of logistic regression (linear) versus GAM (non-linear). Lower GAM AIC or significant non-linear residual patterns (via loess smoothing) indicated non-linearity, justifying GAM's use (5, 6). Loess smoothing fits a local regression to residuals, revealing non-linear patterns that violate logistic regression assumptions (7).

2.3. General Additive Model (GAM)

GAMs offer a robust framework for modelling complex, non-linear relationships between predictors and outcomes, extending Generalized Linear Models (GLMs) by incorporating smooth functions to capture non-linear effects (25). For binary outcomes, as analysed in this study, GAMs employ a logistic link function to model the log-odds of the response variable as a sum of smooth functions of predictors. The logistic GAM is shown below:

$$g(\mu_i) = \log \frac{\mu_i}{1-\mu_i} = \beta_0 + \sum_{j=1}^p f_j(X_{ij}) \quad (4)$$

Where, $\mu_i = P(Y_i = 1)$ is the expected probability of the binary outcome $Y_i \in \{0,1\}$, β_0 is the intercept, and $f_j(X_{ij})$ are smooth functions for each predictor X_{ij} . These smooth functions enable GAMs to model non-linear relationships without assuming a specific functional form (8). In this study, the smooth functions are implemented using cubic regression splines, which are piecewise cubic polynomials joined at knots, ensuring continuity up to the second derivative (9). The basis dimension is set to $k=4$, providing a balance between flexibility for capturing non-linear effects and robustness against overfitting, as recommended for moderate-sized datasets (10). Each smooth function is expressed as:

$$f_j(X_j) = \sum_{k=1}^{k_j} \beta_{jk} b_{jk}(X_j) \quad (5)$$

Where, $b_{jk}(X_j)$ are cubic spline basis functions, β_{jk} are coefficients estimated during model fitting, and $k_j \approx 4$ is the number of basis functions, determined by the number of knots and boundary constraints. To prevent overfitting, GAM employed penalized regression, optimizing a penalized log-likelihood:

$$Objective = \ell(\beta, y) - \sum_{j=1}^p \left(\lambda_j \int [f_j''(x)]^2 dx + \lambda_j^* \int [f_j(x)]^2 dx \right) \quad (6)$$

Where, $\ell(\beta, y)$ is the binomial log-likelihood:

$$\ell = \sum_{i=1}^n y_i \log(\mu_i) + (1 - y_i) \log(1 - \mu_i) \quad (7)$$

The first penalty term, $\lambda_j \int [f_j''(x)]^2 dx$, controls the noise of each smooth function by penalizing the squared second derivative, and the second penalty term, $\lambda_j^* \int [f_j(x)]^2 dx$, shrinks non-significant smooth terms to zero, facilitating automatic feature selection (11). The smoothing parameters λ_j and λ_j^* are estimated using Restricted Maximum Likelihood (REML), balancing model fit and complexity (8). The Effective Degrees of Freedom (EDF) quantifies the complexity of each smooth function after penalization. Defined as the trace of the influence matrix for each smooth term, $EDF_j = trace(F_j)$, the EDF reflects the effective number of parameters used by f_j . An EDF near 1 indicates a nearly linear relationship, values greater than 1 indicate non-linearity, and values near 0 suggest the predictor has been penalized out of the model (8). This study uses EDF to assess the non-linear contributions of predictors, with ensuring the non-significant terms are excluded to enhance model parsimony.

2.4. Extreme Gradient Boost (XGBoost)

XGBoost is an advanced, scalable implementation of Gradient Boosting Machines (GBMs), originally introduced by (12) as a method for constructing predictive models through an ensemble of weak learners, typically decision trees. XGBoost, developed by (13), enhances this framework by

incorporating regularization terms to prevent overfitting, efficient tree construction algorithms for handling large datasets, and support for parallel computing. It is particularly suited for binary classification tasks with non-linear relationships. The model's objective is to minimize a regularized loss function iteratively, where interpretability and predictive accuracy are crucial. XGBoost builds an additive ensemble of regression trees, where each tree corrects the residuals of the previous ones. For a binary classification task, the predicted probability $\hat{y}_i$ for the positive class for observation i is given by:

$$\hat{y}_i = \sigma(\sum_{m=1}^M f_m(x_i)) \quad (8)$$

Where, $\sigma(z) = \frac{1}{1+e^{-z}}$ is the sigmoid (logistic) function that maps the additive log-odds to a probability between 0 and 1, M is the total number of trees (boosting iterations), x_i is the vector of selected predictors (post-LASSO features) and f_m is the output of the m -th tree, a piecewise constant function defined over a partition of the feature space (12, 13). Each tree f_m assigns a score (leaf weight) to each terminal node, and the final prediction aggregates these scores. This formulation extends generalized linear models by allowing flexible, non-parametric capturing of interactions and non-linearities.

Training minimizes a regularized objective function that balances predictive accuracy (empirical risk) with model complexity to avoid overfitting:

$$\mathcal{L} = \sum_{i=1}^n \ell(y_i, \hat{y}_i) + \sum_{m=1}^M \Omega(f_m) \quad (9)$$

Where, $\ell(y_i, \hat{y}_i)$ is the loss function measuring prediction error, $\Omega(f_m)$ and penalizes complex trees. For binary logistic regression, the loss is the binomial deviance (negative log-likelihood):

$$\ell(y_i, \hat{y}_i) = -[y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)] \quad (10)$$

which penalizes confident wrong predictions more heavily. The regularization term $\Omega(f_m)$ is:

$$\Omega(f_m) = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T w_j^2 + \frac{1}{2} \alpha \sum_{j=1}^T |w_j| \quad (11)$$

where T is the number of leaves in the tree, w is the weight (score) of the j -th leaf, γ controls the minimum loss reduction required for a split (pruning), λ applies L2 (ridge) regularization to shrink leaf weights, and α applies L1 (lasso) regularization to promote sparsity by setting some weights to zero (13). This hybrid regularization prevents overfitting in high-dimensional settings, as justified by Tibshirani (14) and Hoerl and Kennard (15) for ridge penalties.

With the objective defined, XGBoost optimizes it via second-order gradient boosting, approximating the loss with a Taylor expansion for efficiency. At iteration m , a new tree f_m is added to minimize:

$$\mathcal{L}^{(m)} \approx \sum_{i=1}^n [g_i f_m(x_i) + \frac{1}{2} h_i f_m(x_i)^2] + \Omega(f_m) \quad (12)$$

Where,

$g_i = \frac{\partial \ell(y_i, \hat{y}_i^{(m-1)})}{\partial \hat{y}_i^{(m-1)}}$ is the first-order gradient (residual direction), and $h_i = \frac{\partial^2 \ell(y_i, \hat{y}_i^{(m-1)})}{\partial \hat{y}_i^{(m-1)2}}$ is the second-order Hessian (curvature, weighting uncertain residuals less) (12). For logistic loss, $g_i = \hat{y}_i^{(m-1)} - y_i$ and $h_i = \hat{y}_i^{(m-1)}(1 - \hat{y}_i^{(m-1)})$. Tree construction uses an exact greedy algorithm to find splits maximizing gain:

$$Gain = \frac{1}{2} \left[\frac{G_L^2}{H_L + \lambda} + \frac{G_R^2}{H_R + \lambda} - \frac{(G_L + G_R)^2}{H_L + H_R + \lambda} \right] - \gamma \quad (13)$$

where G_L, G_R are summed gradients in left/right child nodes, and H_L, H_R are summed Hessians (13). Splits occur only if $Gain > 0$, ensuring parsimony. The learning rate η shrinks tree contributions: $f_m \leftarrow \eta f_m$, controlling step size to prevent overshooting.

Once the core optimization is established, hyperparameters are tuned to adapt the model to the specific dataset. This is performed via grid search over a predefined set of values: tree depth $\in \{3, 6\}$, which limits the maximum depth of each tree to control complexity and prevent overfitting by

restricting the number of splits (lower depths yield simpler trees, reducing variance at the potential cost of bias; Hastie et al., 2009); learning rate $\eta \in \{0.1, 0.3\}$, which scales the contribution of each tree to slow down learning and improve generalization (Friedman, 2002); and number of boosting iterations $\in \{50, 100\}$, determining the ensemble size where too few may underfit and too many may overfit (13).

Grid search operates by exhaustively evaluating all combinations in the Cartesian product of the parameter spaces, resulting in $2 \times 2 \times 2 = 8$ candidates in this case. For each combination $\theta =$ (tree depth, learning rate, number of boosting iterations), the model is trained and evaluated to find the θ^* that maximizes a performance metric. Mathematically, grid search solves:

$$\theta^* = \operatorname{argmax}_{\theta \in \Theta} \mathcal{M}(\theta) \quad (14)$$

Where, Θ is the grid of hyperparameters, and $\mathcal{M}(\theta)$ is the cross-validated performance metric (16). This exhaustive approach is computationally feasible for small grids like this one but can be inefficient for larger spaces; however, it is empirically effective for GBMs in moderate datasets, as it avoids local optima common in random search (16). To evaluate θ , 5-fold cross-validation is employed, which partitions the training data into $k = 5$ equally sized folds to estimate out-of-sample performance and minimize overfitting (17). During cross-validation, the mean test AUC across folds is evaluated. Early stopping is integrated, halting training if the validation AUC does not improve for 10 consecutive rounds to prevent overfitting and reduce computation (18). The best parameters θ^* are those maximizing Cross-validation AUC. The final model is then trained on the full training set (70% of data) while hyperparameters are tuned robustly. This nested approach balances bias-variance trade-off in the data sample (19). Feature importance in XGBoost quantifies each predictor's contribution to the model. It yields three metrics: Gain (primary), Cover, and Frequency (Weight), aggregated across all trees. Gain measures the total improvement in loss from splits on a feature, reflecting its predictive power. Cover quantifies the relative number of observations affected by splits on the feature, averaged across trees, and the frequency (weight) is the number of times the feature is used in splits across the ensemble. Gain was prioritized as it directly ties to model improvement, while cover and frequency provided complementary views.

Shapley Additive Explanation (SHAP) values decompose predictions into additive feature contributions, addressing the black-box nature of ensemble models (20). Rooted in cooperative game theory (21) SHAP assigns each feature a value representing its marginal contribution to the prediction, averaged over all possible coalitions of features. For a prediction $\hat{y}_i = f(x_i)$, SHAP decomposes it as:

$$\hat{y}_i = \phi_0 + \sum_{j=1}^P \phi_{j,i} \quad (15)$$

Where, $\phi_0 = E[f(x)]$ is the base value (expected model output over the dataset, often mean log-odds for logistic models), and $\phi_{j,i}$ is the SHAP value for feature j in instance i :

$$\phi_{j,i} = \sum_{S \subseteq N \setminus \{j\}} \frac{|S|!(p - |S| - 1)!}{p!} [f_{S \cup \{j\}}(x_{S \cup \{j\}}) - f_S(x_S)] \quad (16)$$

Where, $N = \{1, \dots, p\}$ is the set of features, S is a subset (coalition), and f_S is the model trained on S (or approximated). This satisfies efficiency ($\sum \phi_{j,i} = \hat{y}_i - \phi_0$), symmetry, dummy (zero for non-contributing features), and additivity (20).

2.5. Model Validation

Validation is paramount in the methodology to ensure models reliably predict user preferences. Validation in this study employs a held-out test set approach, computing a comprehensive suite of multi-metrics derived from the confusion matrix, and estimating uncertainty through bootstrap resampling. Predictions on the test set are binarized using a 0.5 threshold ($\hat{y}_i = 1$ if $\hat{P}_i \geq 0.5$), where $\hat{P}_i$ is the predicted probability from trained models. The resulting predictions yield a 2x2 confusion matrix C , tabulating true positives (TP: correct positive predictions), true negatives (TN: correct negative predictions), false positives (FP: incorrect positive predictions), and false negatives (FN:

incorrect negative predictions) (22). The performance metrics computed from the confusion matrix includes AUC and F1 score. AUC is the probability that a randomly chosen positive instance ranks higher than a negative one, computed as the integral of the ROC curve, or equivalently via the Wilcoxon-Mann-Whitney statistic (1). Precision is the proportion of positive predictions that are correct:

$$\text{Precision} = \frac{TP}{TP+FP} \quad (17)$$

Precision is critical in contexts where false positives carry significant costs, such as misclassifying non-agreement (23). Sensitivity is the proportion of actual positives correctly identified:

$$\text{Sensitivity} = \frac{TP}{TP+FN} \quad (18)$$

Specificity complements sensitivity, focusing on correct identification of non-agreement cases. Specificity is the proportion of actual negatives correctly identified:

$$\text{Specificity} = \frac{TN}{TP+FP} \quad (19)$$

The F1 score balances precision and sensitivity, providing a single measure of performance that accounts for both false positives and false negatives. F1 score is the harmonic mean of precision and sensitivity:

$$F1 = 2 \cdot \frac{\text{Precision} \cdot \text{Sensitivity}}{\text{Precision} + \text{Sensitivity}} \quad (20)$$

To quantify uncertainty in model performance, bootstrap resampling with $R = 100$ iterations 95% Confidence Intervals (CIs) for accuracy and F1 scores. For a test set $\mathcal{D}_{test} = \{(x_i, y_i)\}_{i=1}^{n_{test}}$, $B = 100$ resamples $\mathcal{D}_b^*$ are drawn with replacement, each of size n_{test} . For each resample B , the GAM and XGBoost model is refitted (or predictions are recomputed for XGBoost using the fitted model). This method assumes approximate normality of the bootstrap distribution for large B , providing reliable intervals without parametric assumptions (24).

3. Results

Model estimation results of GAM significant predictor EDF values and XGBoost features importance, are presented in Figures 3 and 4 respectively. Small shuttles (VT1) in this study are defined as vehicles with a capacity of up to 6 passengers, designed for autonomous road-based transit. The GAM for VT1 identified several significant predictors: Work trips (EDF=0.796, $p=0.018$), School trips (EDF=0.999, $p=0.014$), Tourists (EDF=1.384, $p=0.040$), Leisure travellers (EDF=0.964, $p=0.048$), People with physical disabilities (EDF=1.514, $p=0.032$), and University precincts (EDF=0.909, $p=0.001$). The XGBoost model highlighted DG6 (Gain=0.473) and LU3 (Gain=0.253) as the most influential features, followed by TP5 (Gain=0.090), DG7 (Gain=0.065), TP8 (Gain=0.039), DG5 (Gain=0.027), DG12 (Gain=0.015), TP2 (Gain=0.014), TP1 (Gain=0.010), and LU1 (Gain=0.008).

Minibus shuttles (VT2) are defined as autonomous shuttles with a capacity of up to 20 passengers. GAM for VT2 (minibus shuttle) identified significant predictors: Shopping trips (EDF=1.819, $P=0.039$), Tourists (EDF=1.107, $P=0.024$), People with sensory disabilities (EDF=2.317, $P=0.008$), and University precincts (EDF=1.275, $P=0.001$). The XGBoost model highlighted Leisure travellers (Gain=0.275), Tourists (Gain=0.199), and Industrial/business parks (Gain=0.128) as the most influential features, followed by University students (Gain=0.098), Town centres (Gain=0.065), University precincts (Gain=0.058), Shopping trips (Gain=0.039), Senior citizens (Gain=0.030), Low-income individuals (Gain=0.027), People with sensory disabilities (Gain=0.025), Residential neighbourhoods (Gain=0.014), Special events trips (Gain=0.011), Working professionals (Gain=0.009), Leisure trips (Gain=0.008), and Medical trips (Gain=0.007).

Standard-sized conventional autonomous bus (VT3) refers to vehicles that can carry up to 60 passengers for autonomous road-based transit. GAM for VT3 (standard-sized conventional bus) identified significant predictors: School trips (EDF=0.834, $P=0.018$), Emergency trips (EDF=1.140,

P=0.021), School children (EDF=1.963, P=0.000), Agricultural land areas (EDF=0.711, P=0.038), Tourist destinations (EDF=1.960, P=0.026), and Town centres (EDF=1.837, P=0.024). The XGBoost model prioritized School trips (Gain=0.427) and School trips (Gain=0.148), followed by Industrial/business parks (Gain=0.075), Work trips (Gain=0.066), Emergency trips (Gain=0.064), Low-income individuals (Gain=0.049), University students (Gain=0.040), Agricultural land areas (Gain=0.032), Tourist destinations (Gain=0.029), Town centres (Gain=0.023), High-income individuals (Gain=0.015), Special events trips (Gain=0.014), Leisure trips (Gain=0.008), and University precincts (Gain=0.004).

The operational scenario: complete replacement of conventional buses with autonomous shuttle buses is denoted as SO1. GAM identified significant predictors for this scenario: School trips (EDF=0.816, P=0.005), People with sensory disabilities (EDF=0.677, P=0.016), Agricultural land areas (EDF=0.936, P=0.001), and Town centres (EDF=0.897, P=0.001). The XGBoost model highlighted People with physical disabilities (Gain=0.208) and Town centres (Gain=0.142) as the most influential features, followed by Agricultural land areas (Gain=0.112), School trips (Gain=0.106), Emergency trips (Gain=0.102), School children (Gain=0.102), Tourist destinations (Gain=0.080), People with sensory disabilities (Gain=0.057), Special events trips (Gain=0.053), and People with cognitive disabilities (Gain=0.033).

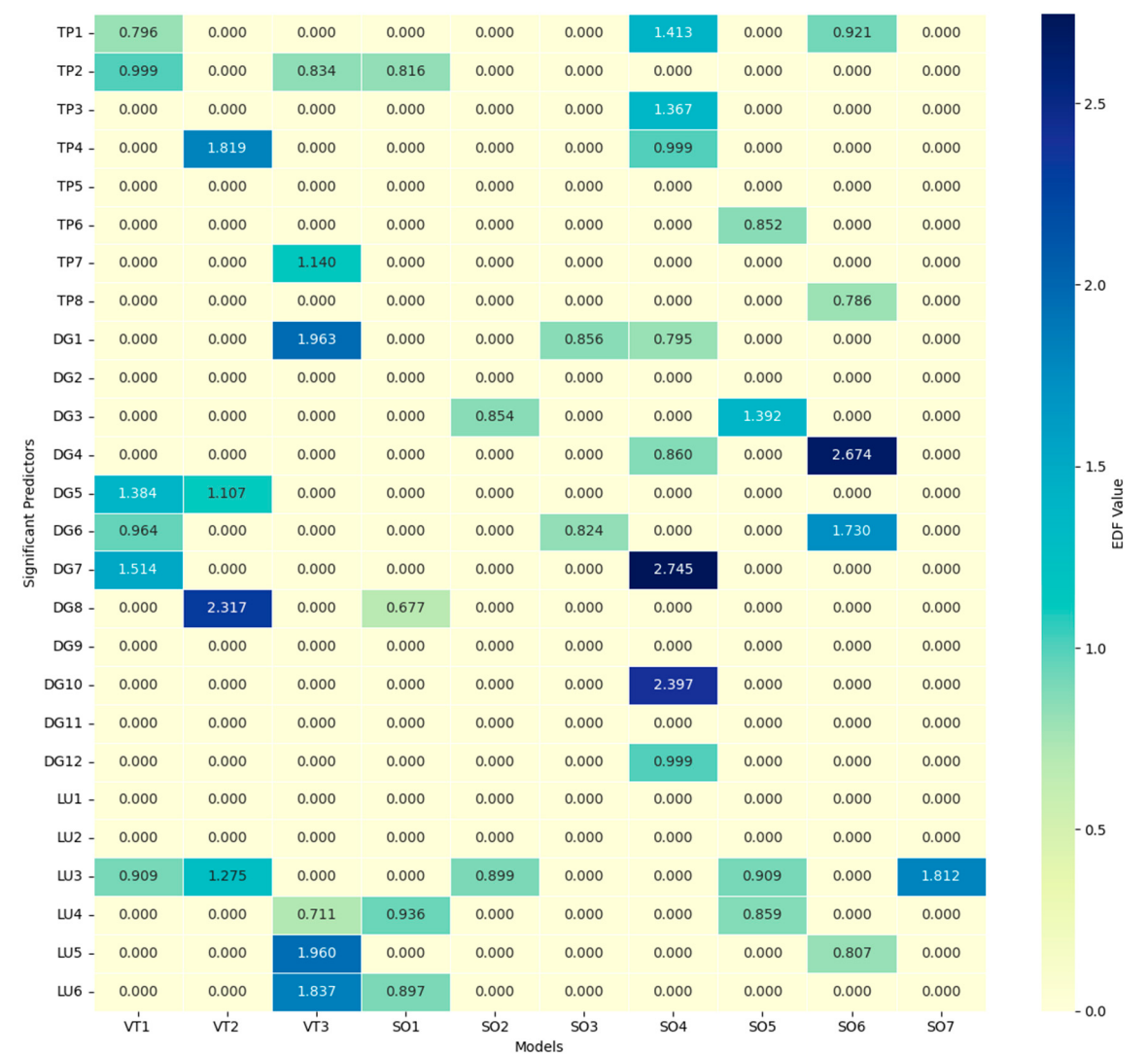


Figure 3. GAM significant predictor EDF values.

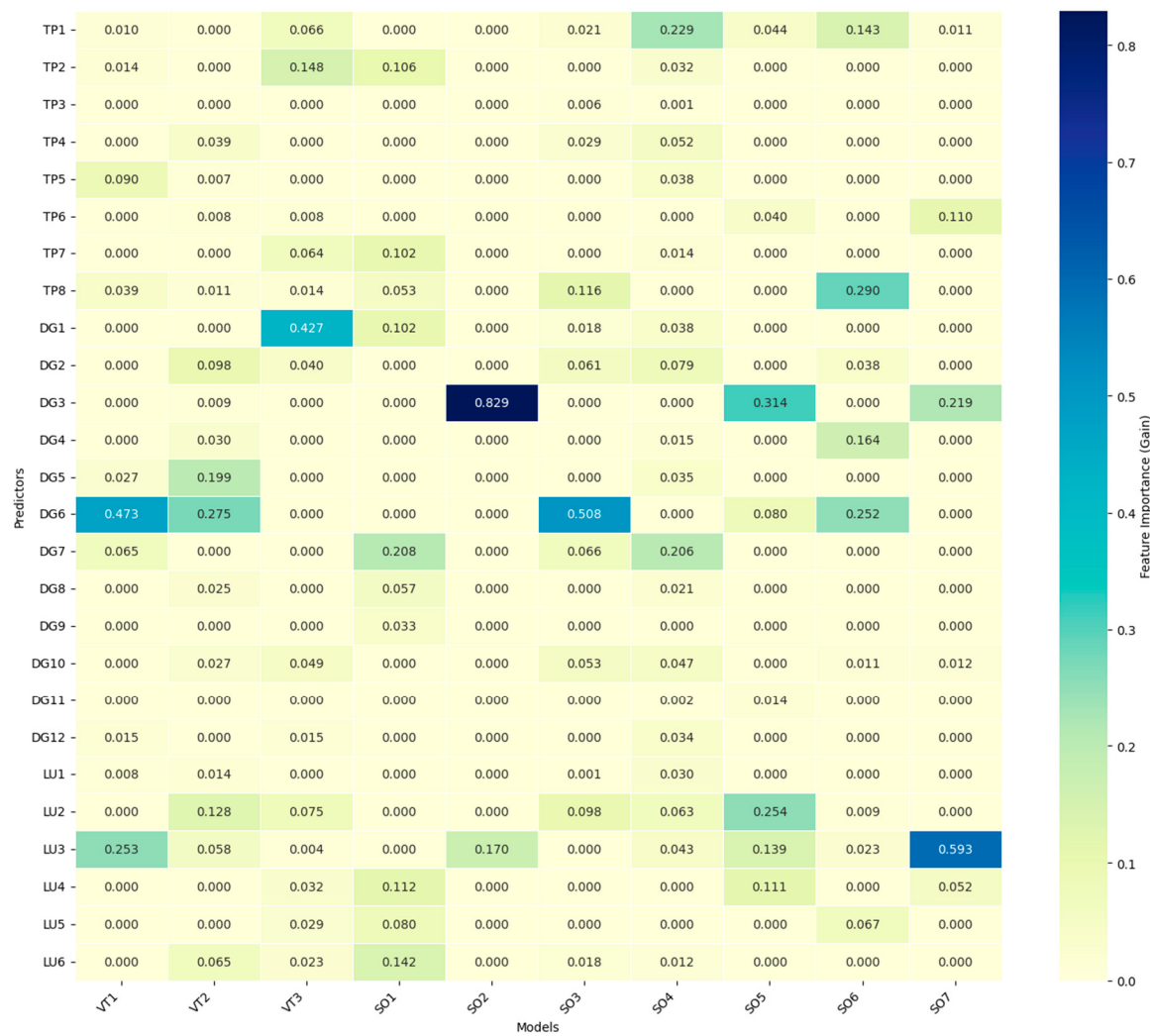


Figure 4. XGBoost predictors feature importance (Gain).

The operational scenario SO2 refers to having autonomous shuttles as a connector to existing fixed-route bus services, where the autonomous road-based transit complements the traditional bus services. GAM identified significant predictors: Working professionals (EDF=0.854, P=0.001) and University precincts (EDF=0.899, P=0.001). XGBoost mirrored these findings, highlighting Working professionals (Gain=0.829) and University precincts (Gain=0.170) as the most influential features. It is imperative to note the lower number of predictors for the model fit, due to the penalized splines and regularization. However, the models were cross validated to minimise the risk of underfitting (14).

The operational scenario SO3 underscores the potential of autonomous shuttles to enhance the connectivity of long-distance road/rail-based services. School children (EDF=0.856, P=0.004) and Leisure travellers, EDF=0.824, P=0.003). The XGBoost model prioritized Leisure travellers (Gain=0.508) and Special events trips (Gain=0.116) as the most influential features, followed by Industrial/business parks (Gain=0.098), People with physical disabilities (Gain=0.066), University students (Gain=0.061), Low-income individuals (Gain=0.053), Shopping trips (Gain=0.029), Work trips (Gain=0.021), School children (Gain=0.018), Town centres (Gain=0.018), University trips (Gain=0.006), and Residential neighbourhoods (Gain=0.001).

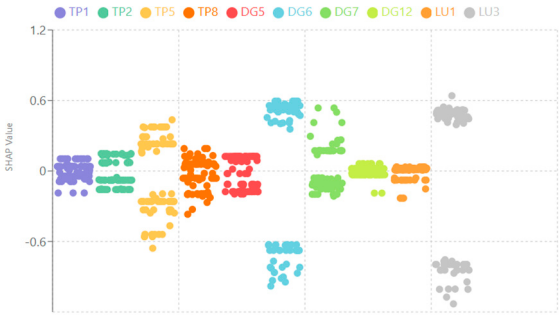
Autonomous shuttles operating as a private taxi service functions similar to traditional taxi services (SO4), providing on-demand, point-to-point transport. GAM identified several significant predictors: Work trips (EDF=1.413, P=0.002), University trips (EDF=1.367, P=0.042), Shopping trips (EDF=0.999, P=0.004), School children (EDF=0.795, P=0.049), Senior citizens (EDF=0.860, P=0.036), People with physical disabilities (EDF=2.745, P=0.001), Low-income individuals (EDF=2.397, P=0.016),

and High-income individuals (EDF=0.999, P=0.011). The XGBoost model ranked worked trips (Gain=0.229) and People with physical disabilities (Gain=0.206) as the most influential features, followed by University students (Gain=0.079), Industrial/business parks (Gain=0.063), Shopping trips (Gain=0.052), Low-income individuals (Gain=0.047), University precincts (Gain=0.043), School children (Gain=0.038), Medical trips (Gain=0.038), Tourists (Gain=0.035), High-income individuals (Gain=0.034), School trips (Gain=0.032), Residential neighbourhoods (Gain=0.030), People with sensory disabilities (Gain=0.021), Senior citizens (Gain=0.015), Emergency trips (Gain=0.014), Town centres (Gain=0.012), Middle-income individuals (Gain=0.002), and University trips (Gain=0.001).

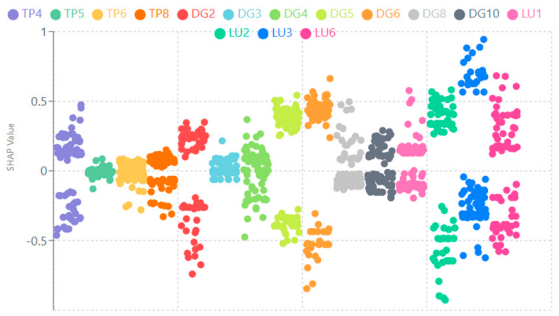
Accommodating autonomous shuttles as a multipurpose service (SO5) combines passenger transport with light freight delivery. Leisure trips (EDF=0.852, P=0.002), Working professionals (EDF=1.392, P=0.001), University precincts (EDF=0.909, P=0.001), and Agricultural land areas (EDF=0.859, P=0.008). The XGBoost model ranked working professionals (Gain=0.314) and Industrial/business parks (Gain=0.254) as the most influential features, followed by University precincts (Gain=0.139), Agricultural land areas (Gain=0.111), Leisure travellers (Gain=0.080), Work trips (Gain=0.044), Leisure trips (Gain=0.040), and Middle-income individuals (Gain=0.014). Model performance metrics indicate solid predictive capability. The GAM achieved an AUC of 0.859 and F1 Score of 0.785, while the XGBoost model performed similarly with an AUC of 0.858, F1 Score of 0.821. Bootstrap 95% CIs for Accuracy were 0.809–0.901 for GAM and 0.794–0.871 for XGBoost, while for F1 Score, they were 0.819–0.909 for GAM and 0.794–0.880 for XGBoost.

The operational scenario SO6 refers to the integration of autonomous shuttles with other transport offerings such as Mobility-as-a-Service (MaaS), to provide seamless connectivity across modes. GAM identified significant predictors: Work trips (EDF=0.921, P=0.043), Special events trips (EDF=0.786, P=0.014), Senior citizens (EDF=2.674, P=0.031), Leisure travellers (EDF=1.730, P=0.007), and Tourist destinations (EDF=0.807, P=0.014). The XGBoost model ranked Special events trips (Gain=0.290) and Leisure travellers (Gain=0.252) as the most influential features, followed by Senior citizens (Gain=0.164), Work trips (Gain=0.1429), Tourist destinations (Gain=0.067), University students (Gain=0.038), University precincts (Gain=0.023), Low-income individuals (Gain=0.011), and Industrial/business parks (Gain=0.009). The operational scenario for autonomous shuttles to operate 24/7 (SO7), provide continuous service availability to users. GAM identified University precincts (EDF=1.812) as the sole significant predictor. The XGBoost model highlighted University precincts (Gain=0.593) and Working professionals (Gain=0.219) as the most influential features, followed by Leisure trips (Gain=0.110), Agricultural land areas (Gain=0.052), Low-income individuals (Gain=0.012), and Work trips (Gain=0.011).

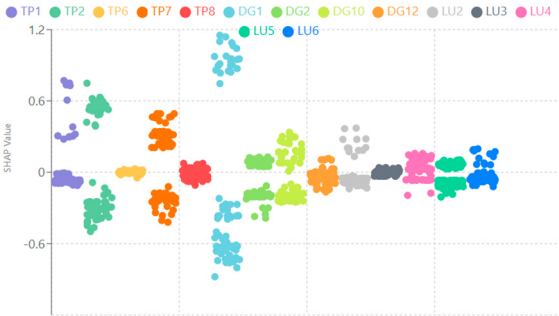
The combined swarm plot from the XGBoost model provides deeper insights into feature contributions of VT1, VT2, VT3, SO1, SO2, SO3, SO4, SO5, SO6 and SO7 (Figures 5(a-j)). For VT1, leisure travellers showed SHAP values ranging from approximately -0.944 to 0.595, indicating a significant and variable impact on predicting small shuttle needs. University precincts exhibited SHAP values from -1.131 to 0.641, with a median around -0.795, suggesting a generally slight negative influence on predictions, though its impact can be positive in some instances. Other features like Work trips and School trips showed SHAP value ranges of -0.187 to 0.147 and -0.158 to 0.147, respectively, indicating moderate and context-dependent influences. For VT2, tourists showed SHAP values ranging from -0.521 to 0.543, indicating a significant and variable impact on predicting minibus shuttle needs. Leisure travellers exhibited SHAP values from -0.846 to 0.663, suggesting a substantial influence that can either increase or decrease the prediction. Shopping trips had SHAP values ranging from -0.462 to 0.474, with a median around 0.133, indicating a generally modest positive influence, though its impact varies. For VT3, school children showed SHAP values ranging from -0.879 to 1.153, indicating a substantial and variable impact on predicting standard-sized autonomous bus needs. School trips exhibited SHAP values from -0.499 to 0.750, with a median around 0.526, suggesting a generally positive influence. Agricultural land areas and Tourist destinations showed SHAP values ranging from -0.194 to 0.145 and -0.208 to 0.101, respectively, indicating a mixed and relatively modest influence.



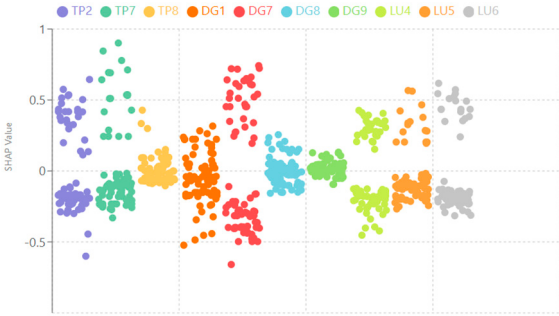
5(a)



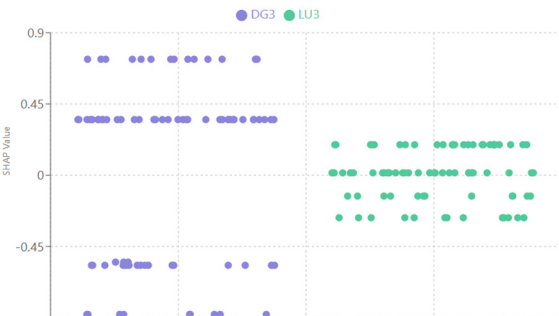
5(b)



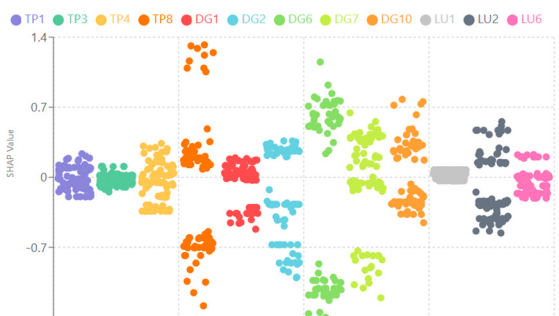
5(c)



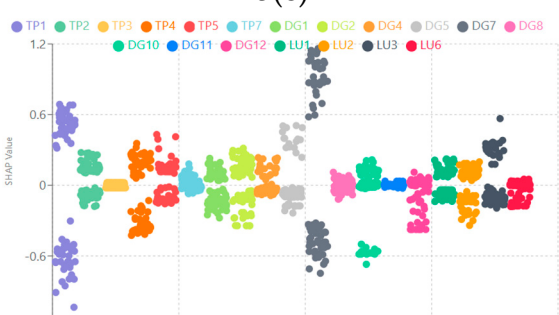
5(d)



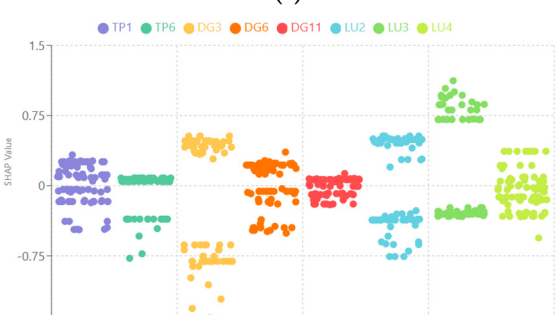
5(e)



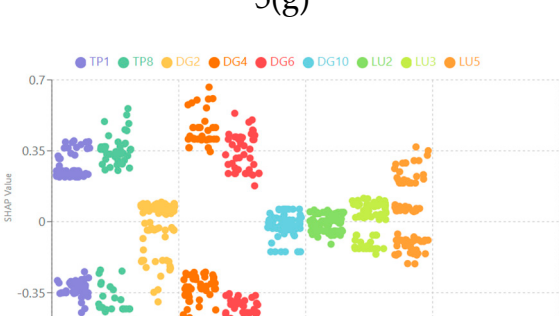
5(f)



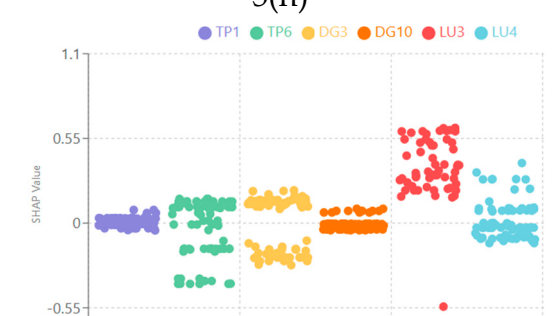
5(g)



5(h)



5(i)



5(j)

Figure 5. Combined swarm plots of SHAP values for XGBoost. (a) VT1. (b) VT2. (c) VT3. (d) SO1. (e) SO2. (f) SO3. (g) SO4. (h) SO5. (I) SO6. (J) SO7.

For SO1, people with physical disabilities showed SHAP values ranging from -0.659 to 0.723, indicating a significant and variable impact. School children exhibited SHAP values from -0.522 to 0.316, suggesting a mixed influence. School trips had SHAP values ranging from -0.600 to 0.644, with a median around 0.417, indicating a generally positive but variable influence. Special events trips ranged from -0.104 to 0.428, showing a modest and variable impact. Agricultural land areas displayed SHAP values from -0.422 to 0.426, indicating a mixed and moderate influence. For SO2, working professionals exhibited SHAP values ranging from -0.879 to 0.732, indicating a substantial and variable impact on predicting the need for connector services. University precincts showed SHAP values from -0.268 to 0.193, with a median around 0.015, suggesting a more modest and mixed influence.

The combined swarm plot for SO3 revealed leisure travellers with SHAP values ranging from -1.399 to 1.151, indicating a substantial and variable impact on the likelihood of needing long-distance connectors. Special events trips showed SHAP values from -1.285 to 1.322, suggesting a significant and highly variable influence tied to event schedules. School children had SHAP values from -0.519 to 0.193, indicating a moderate and mixed effect, with a tendency toward positive values (median around 0.056). The combined swarm plot for SO4 shows work trips with SHAP values ranging from -1.036 to 0.686, indicating a substantial and variable influence on the likelihood of autonomous shuttles serving as private taxis, with both positive and negative contributions depending on work trip patterns. People with physical disabilities has SHAP values from -0.747 to 1.161, underscoring its critical and highly variable role in driving accessibility needs. Low-income individuals show SHAP values from -0.670 to 0.216, suggesting a mixed effect with a tendency toward negative contributions (median around 0.087), reflecting variable demand among this group. Shopping trips ranges from -0.425 to 0.355, indicating a moderate and variable impact.

The combined swarm plot for SO5, working professionals exhibited SHAP values ranging from -1.490 to 0.530, indicating a substantial and variable influence on the likelihood of needing multipurpose services, with both positive and negative contributions depending on professional activity patterns. Leisure trips showed SHAP values from -0.776 to 0.081, suggesting a moderate and variable influence, with a tendency toward negative contributions (median around 0.061). Agricultural land areas had SHAP values from -0.559 to 0.368, indicating a highly variable impact, while University precincts ranged from -0.329 to 1.122, reflecting a significant and diverse effect. For SO6, special events trips with SHAP values ranging from -0.558 to 0.559, indicate a strong and variable influence on integrated transport demand, particularly during events. Leisure travellers have SHAP values from -0.479 to 0.664, reflecting a substantial and variable role in tourist-heavy periods. Senior citizens show SHAP values from -0.541 to 0.607, indicating a highly variable impact, while Work trips ranges from -0.449 to 0.399, suggesting a notable but mixed effect. For SO7, university precincts exhibited SHAP values ranging from -1.063 to 0.618, indicating a strong influence on the demand for 24/7 operations. Working professionals showed SHAP values from -0.273 to 0.208, suggesting a moderate and variable supporting role. Leisure trips had SHAP values from -0.395 to 0.157, indicating a minor and inconsistent impact, while Agricultural land areas ranged from -0.116 to 0.390, reflecting a limited and variable effect.

Model validation metrics for GAM and XGBoost: AUC, F1 score, Bootstrap 95% CI for accuracy and F1 score are presented in Figures 6(a-d). For VT1, Model performance metrics indicate robust predictive capability for both models. The GAM achieved an AUC of 0.909, and F1 Score of 0.816. The XGBoost model performed slightly better with an AUC of 0.913, and F1 Score of 0.838. Bootstrap 95% confidence intervals (CIs) for Accuracy were 0.855–0.947 for GAM and 0.835–0.910 for XGBoost, while for F1 Score, they were 0.821–0.934 for GAM and 0.801–0.894 for XGBoost. For VT2, model performance metrics indicate excellent predictive capability. The GAM achieved an AUC of 0.958 and F1 Score of 0.933. The XGBoost model performed slightly better with an AUC of 0.982 and F1

Score of 0.947. Bootstrap 95% CIs for Accuracy were 0.880–0.958 for GAM and 0.857–0.932 for XGBoost, while for F1 Score, they were 0.870–0.942 for GAM and 0.843–0.929 for XGBoost. For VT3, model performance metrics indicate strong predictive capability. The GAM achieved an AUC of 0.918 and F1 Score of 0.872. The XGBoost model performed slightly better with an AUC of 0.928 and F1 Score of 0.877. Bootstrap 95% CIs for Accuracy were 0.872–0.971 for GAM and 0.840–0.914 for XGBoost, while for F1 Score, they were 0.902–0.978 for GAM and 0.879–0.939 for XGBoost.

For SO1, model performance metrics indicate good predictive capability. The GAM achieved an AUC of 0.843 and F1 Score of 0.860. The XGBoost model performed better with an AUC of 0.909 and F1 Score of 0.899. Bootstrap 95% CIs for Accuracy were 0.789–0.890 for GAM and 0.809–0.886 for XGBoost, while for F1 Score, they were 0.849–0.921 for GAM and 0.863–0.923 for XGBoost. For SO2, model performance metrics indicate moderate predictive capability. The GAM achieved an AUC of 0.855 and F1 Score of 0.727. The XGBoost model showed an AUC of 0.842 and F1 Score of 0.630. Bootstrap 95% CIs for Accuracy were 0.752–0.840 for GAM and 0.739–0.826 for XGBoost, while for F1 Score, they were 0.651–0.807 for GAM and 0.665–0.789 for XGBoost. For SO3, model performance metrics indicate strong predictive capability. Both the GAM and XGBoost models achieved AUC values of 0.902 and 0.907, respectively, F1 scores of 0.750. Bootstrap 95% CIs for accuracy were 0.833–0.927 for GAM and 0.823–0.895 for XGBoost, while for F1 score, they were 0.829–0.927 for GAM and 0.822–0.893 for XGBoost.

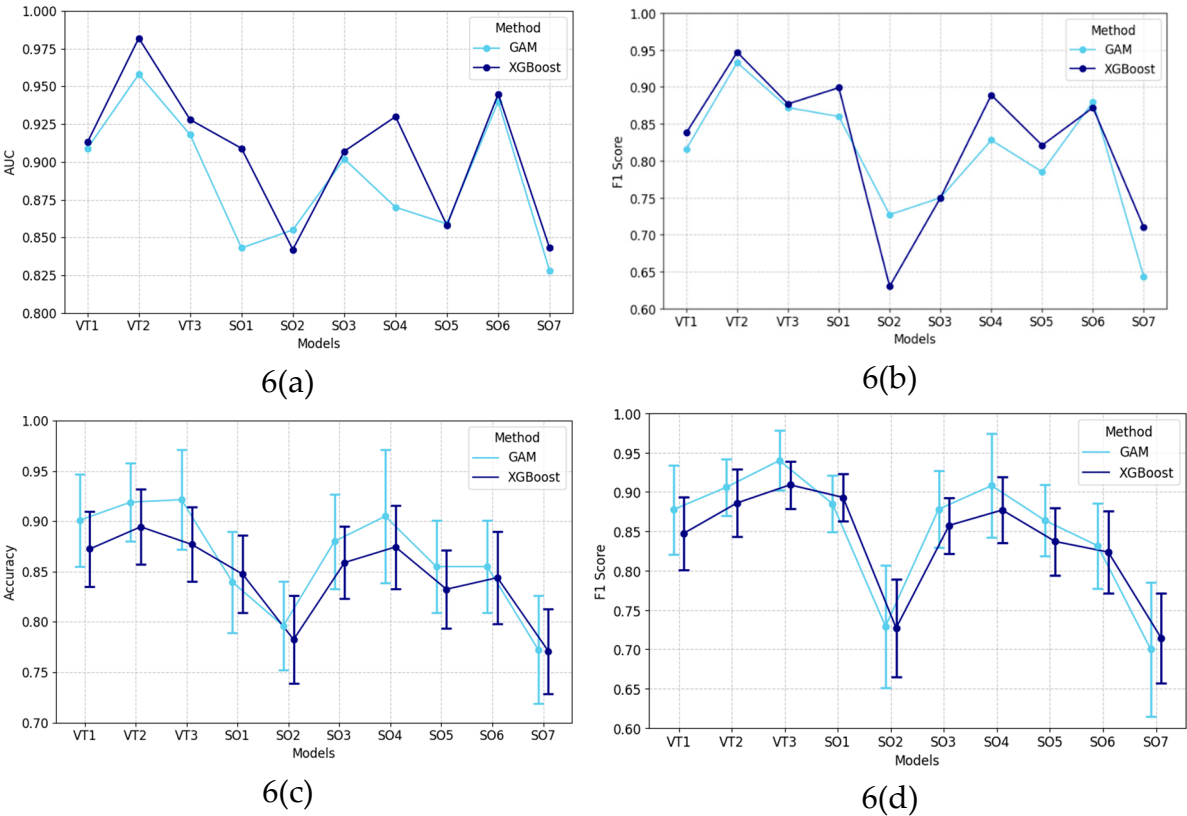


Figure 6. Model validation plots of GAM and XGBoost. (a) AUC. (b) F1 Score. (c) Bootstrap 95% CI for Accuracy. (d) Bootstrap 95% CI for F1 Score.

For SO4, model performance metrics indicate strong predictive capability. The GAM achieved an AUC of 0.870 and F1 Score of 0.828. The XGBoost model performed better with an AUC of 0.930, F1 Score of 0.889. Bootstrap 95% CIs for Accuracy were 0.839–0.971 for GAM and 0.833–0.916 for XGBoost, while for F1 Score, they were 0.842–0.974 for GAM and 0.835–0.919 for XGBoost. For SO5, model performance metrics indicate solid predictive capability. The GAM achieved an AUC of 0.859 and F1 Score of 0.785, while the XGBoost model performed similarly with an AUC of 0.858, F1 Score

of 0.821. Bootstrap 95% CIs for Accuracy were 0.809–0.901 for GAM and 0.794–0.871 for XGBoost, while for F1 Score, they were 0.819–0.909 for GAM and 0.794–0.880 for XGBoost.

For SO6, model performance metrics indicate strong predictive capability. The GAM achieved an AUC of 0.940 and F1 Score of 0.880. The XGBoost model performed similarly with an AUC of 0.945, F1 Score of 0.872. Bootstrap 95% CIs for Accuracy were 0.809–0.901 for GAM and 0.798–0.890 for XGBoost, while for F1 Score, they were 0.777–0.886 for GAM and 0.771–0.876 for XGBoost. For SO7, model validation metrics indicate moderate predictive performance. The GAM achieved an AUC of 0.828 and F1 Score of 0.643. The XGBoost model showed improved performance with an AUC of 0.843 and F1 Score of 0.710. Bootstrap 95% CIs for Accuracy were 0.719–0.826 for GAM and 0.729–0.813 for XGBoost, while for F1 Score, they were 0.615–0.785 for GAM and 0.657–0.771 for XGBoost.

4. Discussion

This section provides a comprehensive discussion on the operational determinants for the different autonomous shuttle vehicle types (section 4.1 - 4.3) and operational scenarios (sections 4.4 - 4.10).

4.1. Autonomous Shuttles as Small-Sized Shuttle

The findings underscore the potential of small autonomous shuttles to address specific mobility needs in targeted contexts. Both the GAM and XGBoost models identified leisure travellers and university precincts as key predictors of small shuttle suitability, with additional significant predictors in the GAM model including work trips, school trips, tourists, and people with physical disabilities. The prominence of leisure travellers in both models, with a high feature importance gain of 0.473 in the XGBoost model and a significant EDF of 0.964 in the GAM model, suggests that small shuttles are particularly well-suited for non-routine, flexible travel demands. This finding is consistent with research by Golbabaee, Yigitcanlar (25), who identified perceived usefulness and ease of use as critical drivers of autonomous shuttle adoption, particularly among users seeking convenient transport options. The variable impact of leisure travellers, as indicated by SHAP values ranging from -0.944 to 0.595, underscores the diverse needs of this group, which may require tailored shuttle services to meet specific demands.

Similarly, the significance of university precincts as a predictor, with a gain of 0.253 in the XGBoost model and an EDF of 0.909 in the GAM model, supports the suitability of small shuttles in campus environments. Iclodean, Cordos (26) note that autonomous shuttles are frequently deployed in university campuses due to their controlled traffic conditions and predictable demand patterns, which facilitate the integration of autonomous technologies. The mixed impact, as shown by SHAP values from -1.131 to 0.641 with a median of -0.795, suggests that while university precincts are generally suitable, their effectiveness may vary depending on specific route characteristics or demand fluctuations. The GAM model's identification of work trips (EDF=0.796) and school trips (EDF=0.999) as significant predictors further highlight the potential of small shuttles for short, routine trips. These findings align with the literature emphasizing the role of autonomous shuttles in reducing urban congestion and enhancing mobility for regular, predictable journeys (27). The moderate SHAP value ranges for TP1 (-0.187 to 0.147) and TP2 (-0.158 to 0.147) in the XGBoost model suggest that while these trip types contribute to shuttle suitability, their impact is context-dependent, possibly influenced by factors such as trip distance or time of day.

4.2. Autonomous Shuttles as Minibus-Sized Shuttles

The findings of the autonomous minibus shuttle underscore its potential as a versatile and inclusive mobility solution, particularly for shopping trips, tourist-related travel, and serving individuals with sensory disabilities, and in university precincts and industrial/business parks. The GAM and XGBoost models demonstrated exceptional predictive performance, with AUC scores of

0.958 and 0.982, respectively, and F1 scores of 0.933 and 0.947, indicating robust and reliable predictions. These results align with the growing body of literature on autonomous shuttle services, which highlights their efficacy in urban and controlled environments for diverse user groups (26, 28). The prominence of shopping trips in both models, with a significant EDF of 1.819 in GAM and a gain of 0.039 in XGBoost, suggests that minibus shuttles are well-suited for urban settings where shopping is a frequent activity. The SHAP values for TP4, ranging from -0.462 to 0.474 with a median around 0.133, indicate a generally positive but context-dependent influence. This aligns with research emphasizing the role of autonomous shuttles in providing last-mile connectivity to shopping districts, reducing reliance on private vehicles and alleviating urban congestion (29-31).

Tourists and leisure travellers emerged as key predictors, with DG5 significant in both models (GAM: EDF=1.107; XGBoost: Gain=0.199) and DG6 leading in XGBoost (Gain=0.275). The wide SHAP value ranges for DG5 (-0.521 to 0.543) and DG6 (-0.846 to 0.663) suggest that their impact varies depending on factors such as tourist volume or leisure travel patterns. This variability is consistent with studies highlighting the flexibility of autonomous shuttles in adapting to sporadic demand (28, 32). The significance of university precincts in both models (GAM: EDF=1.275; XGBoost: Gain=0.058) supports the suitability of minibus shuttles for campus environments, where flexible and efficient transport is critical for students and staff. Autonomous shuttles have been tested in university settings globally, where they have demonstrated effectiveness in controlled environments with predictable demand patterns (33, 34).

A critical finding is the significance of people with sensory disabilities in the GAM model (EDF=2.317), underscoring the importance of accessibility in autonomous shuttle design. However, its lower influence in the XGBoost model (Gain=0.025) suggests that accessibility may not be the primary driver of shuttle suitability in all contexts. This finding is supported by literature emphasizing the transformative potential of autonomous vehicles for individuals with disabilities, particularly those with sensory impairments, but also highlighting the lack of specific accessibility standards (28, 35). The high predictive performance of both models, coupled with stable bootstrap confidence intervals, provides confidence in their ability to guide minibus shuttle deployment. However, the variable SHAP values for key predictors like DG5 and DG6 highlight the need for context-specific deployment strategies to account for fluctuating demand. User acceptance studies suggest that operational improvements, such as increasing travel speeds and reducing abrupt braking, are critical to enhancing passenger satisfaction, particularly for tourists and leisure travellers who prioritize comfort and convenience (25, 36). The findings advocate for deploying autonomous minibus shuttles in mixed-use urban settings, where their higher capacity and flexibility can balance diverse mobility needs (37).

4.3. Autonomous Shuttles as Standard-Sized Conventional Buses

The findings for the standard-sized conventional bus highlight its critical role in public transportation, particularly for school trips and school children. Additionally, the GAM significance of town centres and tourist destinations underscores the pivotal role of standard-sized buses in providing safe, accessible, and efficient transport for large groups, particularly in high-demand scenarios (38-40). The prominence of school trips and school children in both models emphasizes the essential function of standard buses in student transport. The significant influence of DG1 in the XGBoost model, with a gain of 0.427 and SHAP values ranging from -0.879 to 1.153, reflects the variable but substantial impact of school-related demand and parent acceptance. Similarly, TP2's SHAP values from -0.499 to 0.750, with a median around 0.526, indicate a generally positive influence, reinforcing the consistent need for standard-sized autonomous buses for school trips.

The significance of emergency trips in the GAM model (EDF=1.140) and its notable presence in the XGBoost model (Gain=0.064) suggests that standard buses are vital for emergency services, such as community evacuations or rapid response during crises. While not typically designed for emergency medical services, buses are often utilized in large-scale emergencies due to their high capacity and reliability. While, buses have been deployed for disaster management over the years,

evacuation from natural disasters such as hurricanes and bush fires have been proposed using autonomous shuttles (41, 42). However, there should be a careful consideration when deploying autonomous buses, as there may be non-compliance evacuees and other unforeseen technological issues under such critical situations.

The inclusion of agricultural land areas, tourist destinations, and town centres as significant predictors in the GAM model, with modest SHAP values in the XGBoost model, indicates the versatility of standard-sized autonomous buses in serving diverse geographies. In rural areas, where transport infrastructure may be limited, autonomous buses can be utilized to serve various land uses. However, the modest SHAP values for LU4 (-0.194 to 0.145) and LU5 (-0.208 to 0.101) suggest that their influence on individual predictions is less pronounced compared to school-related variables, indicating that deployment strategies should prioritize school transport while considering these areas as secondary contexts. The XGBoost model's emphasis on industrial/business parks (Gain=0.075) and work trips (Gain=0.066) suggests additional suitability for serving employees in areas with limited public transport options, aligning with the role of standard buses in facilitating commuter travel. The presence of low-income individuals (Gain=0.049) as a notable feature further underscores the importance of buses in providing affordable transport solutions, particularly in underserved communities.

4.4. Autonomous Shuttles as Standard-Sized Conventional Buses

The findings suggest that replacing conventional buses with autonomous road-based transit is feasible in specific contexts, particularly in rural town centres, with a strong emphasis on serving school trips and passengers with sensory and physical disabilities. GAM and XGBoost models demonstrated good predictive performance, with AUC scores of 0.843 and 0.909, respectively, and F1 scores of 0.860 and 0.899. The prominence of rural town centres in both models, with a significant EDF of 0.897 in GAM and a gain of 0.142 in XGBoost, underscores the suitability of autonomous road-based transit in urban environments where demand is variable and congestion is a concern (43). The integration of autonomous shuttles in town centres should be integrated through real-time route optimisation and fleet variance to respond to real-time demand fluctuations, improving service efficiency.

Agricultural land areas also emerged as a significant predictor, with an EDF of 0.936 in GAM and a gain of 0.112 in XGBoost. The SHAP values for LU4, ranging from -0.422 to 0.426, suggest a mixed influence, indicating that autonomous road-based transit effectiveness in rural areas depends on specific demand patterns. Literature supports the use of Demand Responsive Transit (DRT) in low-density rural areas, where traditional bus services are often financially unsustainable due to sparse populations (44). The significance of school trips in both models, with an EDF of 0.816 in GAM and a gain of 0.106 in XGBoost, highlights autonomous shuttles potential to serve educational transport needs. The SHAP values for TP2, ranging from -0.600 to 0.644 with a median around 0.417, indicate a generally positive but context-dependent influence.

Accessibility for passengers with disabilities is a critical finding, with People with sensory disabilities significant in GAM (EDF=0.677) and People with physical disabilities leading in XGBoost (Gain=0.208). The wide SHAP value range for DG7 (-0.659 to 0.723) underscores its variable but substantial impact, suggesting that physical accessibility is a key driver for autonomous shuttle adoption in certain contexts. However, the moderate influence of DG8 in XGBoost (Gain=0.057) suggests that sensory accessibility may require further technological advancements to fully meet user needs. Literature emphasizes the transformative potential of autonomous vehicles for individuals with disabilities, provided they incorporate inclusive design features (28, 45).

The robust performance of the XGBoost model, with higher AUC and Specificity (0.700 vs. 0.667 for GAM), indicates its effectiveness in identifying suitable contexts for autonomous shuttle replacement, which is crucial for cost-effective deployment. However, there are potential challenges in distinguishing unsuitable contexts, which could lead to over-reliance on autonomous shuttles in areas better served by conventional buses. This aligns with research advocating for hybrid transit

models that combine conventional transit with autonomous transit (46-48). Additionally, dynamic public transit systems have been proposed by researchers for optimised transit operations (49, 50). The performance of the transit depends on the autonomous shuttle penetration rate, which should be optimised with the real-time passenger demand (51-53).

4.5. Autonomous Shuttles as a Connector to Existing Fixed-Route Bus Services

The findings suggest that autonomous shuttles have significant potential to serve as a connector to fixed-route bus services, particularly for working professionals and in university precincts. GAM and XGBoost models identified these predictors as the most influential, highlighting the specific user group and location where autonomous shuttles can enhance first-mile and last-mile (FMLM) connectivity. The moderate performance metrics, with AUC scores around 0.85 and F1 scores between 0.63 and 0.73, indicate that while autonomous shuttles show promise, challenges remain in accurately predicting its suitability across diverse contexts, likely due to variable demand patterns in university settings.

The prominence of working professionals in both models, with a high gain of 0.829 in XGBoost and an EDF of 0.854 in GAM, underscores their critical role in driving the need for connector services. The SHAP values for DG3, ranging from -0.879 to 0.732, suggest a substantial but variable impact, likely reflecting the diverse commuting patterns of professionals in rural areas. University precincts also emerged as a significant predictor, with an EDF of 0.899 in GAM and a gain of 0.170 in XGBoost. The SHAP values for LU3, ranging from -0.268 to 0.193 with a median around 0.015, indicate a more modest and mixed influence, suggesting that while university precincts are relevant, their impact varies depending on specific campus characteristics, such as size, layout, or transit hub proximity. This aligns with the existing research findings, as many researchers have suggested the applicability of autonomous shuttles to solve the FMLM problem (53-60).

4.6. Autonomous Shuttles as a Connector to Longer Distance Services

GAM and XGBoost models demonstrated strong predictive performance, with AUC scores of 0.902 and 0.907, respectively, and F1 scores of 0.750, indicating reliable predictions for ADRT deployment in this context. The high feature importance of DG6 in the XGBoost model (Gain=0.508) and its substantial SHAP values (-1.399 to 1.151) highlight the critical and variable impact of leisure travel, while TP8's SHAP values (-1.285 to 1.322) emphasize the significant influence of event-related demand. The prominence of leisure travellers in both models suggests that autonomous shuttles are well-suited to meet the flexible and sporadic travel needs of this group, who often require transport to access long-distance services in tourist-heavy areas or during off-peak times. research have clearly identified that passenger quality of service can be significantly improved through autonomous shuttle operations, through improved availability, accessibility, reliability and comfort (51, 61, 62).

Special events trips, with a notable gain of 0.116 in the XGBoost model and highly variable SHAP values, indicate that autonomous shuttles can effectively serve event attendees, whose travel demand is often concentrated around specific times and locations. The moderate influence of school children, with an EDF of 0.856 in GAM and SHAP values ranging from -0.519 to 0.193, suggests that while school-related travel contributes to the need for long-distance connectors, its impact is less dominant compared to leisure and special event travel. The findings advocate for prioritizing autonomous shuttle services in tourist-heavy routes and special events trips to reduce reliance on private vehicles and support sustainable transport goals.

4.7. Autonomous Shuttles as a Private Taxi Service

The findings highlight the versatility of autonomous shuttles when operated as a private taxi service, evidenced by the broad range of significant predictors identified by both the GAM and the XGBoost models. The prominence of Work trips and People with physical disabilities in both models underscores the critical role that autonomous shuttles can play in serving daily commuting needs

and enhancing accessibility for individuals with mobility challenges. The high feature importance of work trips in both models, with an EDF of 1.413 in GAM and a gain of 0.229 in XGBoost, suggests that ADRT as private taxis is particularly well-suited for commuters requiring flexible, on-demand transport. The SHAP values for TP1, ranging from -1.036 to 0.686, indicate a substantial but variable influence, likely reflecting diverse commuting patterns across remote and rural town centres. Accessibility for people with physical disabilities is a critical driver, as evidenced by its high EDF of 2.745 in GAM and gain of 0.206 in XGBoost, with SHAP values ranging from -0.747 to 1.161. This variability underscores the diverse mobility needs within this demographic, which may include requirements for low-floor vehicles, automated boarding systems, or real-time assistance (28).

Low-income individuals also emerged as a significant predictor, with an EDF of 2.397 in GAM and a gain of 0.047 in XGBoost, though their SHAP values (-0.670 to 0.216, median around 0.087) raises concerns on affordability. The mixed influence of low and middle-income individuals indicates that financial barriers may limit adoption unless addressed through subsidies or dedicated services, as suggested by the study's policy recommendations. On a similar note, high-income individuals contribute to demand, with an EDF of 0.999 in GAM and a gain of 0.034 in XGBoost, reflecting their preference for convenient, on-demand transport options, as noted in studies on consumer willingness to adopt autonomous taxis (63, 64). The inclusion of other demographics, such as school children, senior citizens, and trip purposes like university trips and shopping trips, highlights the broad applicability of autonomous shuttles as private taxis. Autonomous taxis will help reduce private car costs by up to 16%, especially for small trip distances and/or in residential areas far from stations, and larger reductions in low-density metropolitan areas (65).

4.8. Autonomous Shuttles as Multipurpose Services

The findings from this study underscore the potential of autonomous road-based transit to operate as multipurpose services, integrating passenger transport and parcel delivery, particularly for working professionals in university precincts, industrial/business parks, and agricultural land areas. GAM and XGBoost models demonstrated solid predictive performance, with AUC scores of 0.859 and 0.858, respectively, and F1 scores of 0.785 and 0.821, indicating reliable predictions for the operational scenario. The prominence of working professionals in both models, with a high EDF of 1.392 in GAM and a gain of 0.314 in XGBoost, suggests the high acceptance among this demographic to have multipurpose services, due to the high acceptance of technology. The wide SHAP value range for DG3 (-1.490 to 0.530) reflects the wider acceptance for this demographic, yet potential discrepancies on prioritisation between passengers and parcels. Thus, the optimisation should be done accordingly, so that the passengers will not experience any extended waiting time. Preference for passenger delivery should always be maintained, while scheduling freight delivery to balance empty vehicle kilometres.

University precincts also emerged as a key predictor, with an EDF of 0.909 in GAM and a gain of 0.139 in XGBoost. The SHAP values for LU3 (-0.329 to 1.122) indicate a significant and diverse impact, with a higher inclination towards the positive, due to the higher acceptability for this land use. Industrial/business parks, with a high gain of 0.254 in XGBoost, highlight the potential for autonomous shuttles to serve both employee commuting and logistics needs in commercial areas. Agricultural land areas, significant in both models (EDF=0.859 in GAM; Gain=0.111 in XGBoost), indicate autonomous shuttles potential to improve rural mobility and logistics. The SHAP values for agricultural land areas (-0.559 to 0.368) suggest a variable impact, likely due to the complex nature in these land uses and consequently, the incapacibilities of autonomous shuttles to navigate effectively.

4.9. Autonomous Shuttles Integrated with Other Transport Offerings

The findings highlight the potential of autonomous shuttles to enhance mobility within integrated transport ecosystems, such as MaaS, particularly for special events, leisure travellers, and senior citizens in tourist destinations. Work trips also play a significant role, reflecting the versatility of autonomous shuttles in supporting diverse mobility needs. Given the prominence of ADRT as a

service type for autonomous shuttles, the further integration to MaaS align with the existing literature to improve connectivity and accessibility. The strong performance of both the GAM and XGBoost models, with AUC scores of 0.940 and 0.945, respectively, and F1 scores around 0.88, indicates reliable predictions for autonomous shuttle deployment in this context.

The high feature importance of special events trips in the XGBoost model (Gain=0.290) and its significant EDF of 0.786 in GAM, combined with SHAP values ranging from -0.558 to 0.559, suggest that autonomous shuttles can be offered with other transport offerings for special events. Leisure travellers are a key driver, with an EDF of 1.730 in GAM and a gain of 0.252 in XGBoost, and SHAP values from -0.479 to 0.664 indicating a substantial and variable impact. This reflects the dynamic travel patterns of tourists, who often require flexible transport to access destinations integrated with broader transport networks. MaaS plays a pivotal role in providing seamless, multimodal travel options for tourists, with ADRT serving as a critical component for last-mile connectivity in tourist-heavy areas. Senior citizens also emerged as a significant predictor, with a high EDF of 2.674 in GAM and a gain of 0.164 in XGBoost, with SHAP values from -0.541 to 0.607 indicating variable but notable influence. The integration of ADRT within MaaS can further support seniors by offering door-to-door services coordinated with other transport modes.

Work trips, with an EDF of 0.921 in GAM and a gain of 0.1429 in XGBoost, suggest that ADRT can support daily commuting within integrated transport systems. The SHAP values for TP1 (-0.449 to 0.399) indicate a mixed effect, reflecting variable commuting patterns. It is proven that integration of ADRT within MaaS can improve service efficiency, reducing congestion and thereby emissions.

4.10. Autonomous Shuttles to Operate 24/7

The findings highlight the potential of autonomous shuttles to provide continuous service, with university precincts emerging as the most significant predictor. The GAM identified university precincts as the sole significant predictor (EDF=1.812), while the XGBoost model reinforced this with a Gain of 0.593, and added working professionals (Gain=0.219) as a secondary factor. The moderate performance metrics, with AUC scores of 0.828 (GAM) and 0.843 (XGBoost) and F1 scores of 0.643 and 0.710, respectively, suggest challenges in predicting continuous operation needs, likely due to fluctuating demand outside peak hours. The dominance of university precincts in both models, with SHAP values ranging from -1.063 to 0.618, underscores the unique suitability of campuses for 24/7 ADRT operations.

Leisure trips (Gain=0.110) and agricultural land areas (Gain=0.052) showed minor and variable impacts, with SHAP values indicating inconsistent demand (TP6: -0.395 to 0.157; LU4: -0.116 to 0.390). This suggests limited need for 24/7 operations in these contexts, likely due to lower nighttime demand. Economically, 24/7 operations face challenges due to high operational costs during low-demand periods. An optimised model, with integrating multipurpose services during low-demand periods is required to maintain effective operational efficiency. However, public acceptance and safety are critical considerations for 24/7 operations, particularly in less controlled environments.

5. Conclusions

This study provides a comprehensive empirical investigation into the operational determinants of autonomous shuttle systems in rural and peri-urban communities of South-East Queensland, addressing a critical gap in the literature on autonomous road-based transit in rural networks. By employing a structured survey and adopting an analytical approach through GAM and XGBoost, the research explains user preferences across diverse demographic groups, trip purposes, land-use contexts, on autonomous shuttle types and service offerings. Small shuttles are optimally suited for flexible, non-routine transport operations, offering adaptability to diverse travel demands. In contrast, minibus shuttles are preferred for their critical role in facilitating FMLM connectivity, effectively bridging gaps in public transport networks. Standard-sized buses remain the preferred choice for high-capacity transport requirements, particularly for school-related trips, due to their

ability to accommodate large passenger volumes efficiently. Furthermore, these buses are recognised for their utility in providing disaster management support in regional areas

While the complete replacement of conventional buses with autonomous shuttles is not widely supported, a hybrid model integrating traditional buses with autonomous shuttles is increasingly advocated. Autonomous shuttles are particularly valued for their capacity to address FMLM connectivity challenges, serving as vital connectors to fixed-route and long-distance transport services. The emergence of autonomous shuttles as private taxi services has been noted, though this development raises equity concerns, particularly for transport-disadvantaged groups, who may face barriers to accessing such services. For cost-effective operations, autonomous shuttles can be deployed as multipurpose services, particularly within industrial and business parks. This approach enables the simultaneous transport of passengers and freight, optimising vehicle utilisation during otherwise empty vehicle kilometres. While cost concerns associated with 24/7 operations persist, these can be mitigated by adopting a multipurpose service framework. Moreover, integrating autonomous shuttles within a MaaS ecosystem enhances operational efficiency and contributes to reduced emissions, aligning with broader sustainability objectives.

6. Policy Implications

Based on the empirical findings from the study, policy recommendations are provided for an effective autonomous road-based transit implementation in rural setting (Table 3).

Table 3. Policy recommendations.

Strategy	Policy Recommendations
Small shuttle	• Partner with universities to implement small shuttle services for campus mobility • Integrate services in tourist destination, ensuring accessibility during peak tourist seasons • Use dynamic routing algorithms to optimize service
Minibus shuttles	• Implement minibus shuttles in town centres and shopping districts to provide FMLM connectivity • Mandate accessibility standards, to accommodate passengers with disabilities • Develop partnerships with tourism boards to deploy shuttles in high-traffic tourist areas • Support deployment in industrial/business parks to serve employees, with schedules aligned to peak commuting hours
Standard-sized conventional bus	• Collaborate with schools to deploy standard-sized buses for student transport, ensuring safety features like real-time monitoring and geofencing • Develop emergency response protocols integrating autonomous buses for large-scale evacuations, particularly in rural areas prone to natural disasters • Promote affordable fares/subsidize services in agricultural land areas to enhance connectivity and ensure equitable access to underserved communities
Completely replace conventional buses	• Initiate pilot programs in town centres with real-time route optimization to manage variable demand • Support school transport with dedicated autonomous shuttle routes, ensuring compliance with safety standards • Evaluate hybrid models combining conventional and autonomous buses to address contexts where full replacement is less viable
Operate as a Connector to Existing Fixed-Route Bus Services	• Develop mobile apps for seamless booking and integration with fixed-route services, improving user experience • Conduct public awareness campaigns to enhance acceptance and trust

	• Conduct regular evaluations to address variable demand patterns
Connector to longer distance services	• Deploy shuttles in tourist destinations to connect with rail or coach services, integrating with event schedules for special events • Offer flexible scheduling to accommodate sporadic demand from leisure travellers and event attendees • Promote integration with regional transport networks to enhance connectivity and reduce private vehicle reliance
Operate as private taxi services	• Introduce subsidies or tiered pricing to ensure affordability for low-income individuals • Deploy in residential and industrial areas to support work and shopping trips • Use on-demand booking platforms to optimize point-to-point services, ensuring flexibility for commuters
Accommodate as a multipurpose service	• Develop dual-purpose shuttle designs to accommodate passengers and parcels, prioritizing passenger demand • Deploy in industrial/business parks to serve employees and logistics, scheduling freight delivery during off-peak hours. • Support rural logistics in agricultural areas with flexible routes to connect farms and markets • Implement real-time optimization to balance passenger and freight demands, minimizing empty vehicle kilometres
Integrated with Other Transport Offerings	• Develop MaaS platforms integrating autonomous shuttles with other modes for seamless connectivity • Prioritize services for special events and tourist destinations, using dynamic scheduling to handle peak demand • Ensure accessibility for senior citizens with door-to-door services and user-friendly booking interfaces
Operate 24/7	• Conduct public awareness campaigns to enhance acceptance and trust in 24/7 autonomous shuttle operations • Base operations in university precincts for vehicle charging during off-peak hours

7. Limitations and Future Research Directions

This study offers a robust methodology to analyse operational determinants for rural autonomous road-based transits, but several methodological limitations warrant consideration. First, the sample size of 357 responses, reduced after cleaning, may not fully represent the diverse preferences of rural populations, and convenience random sampling via Qualtrics risks selection bias, particularly for those with limited digital access. Second, the focus on South-East Queensland restricts generalizability to other rural regions with distinct socio-economic or infrastructural contexts. Third, reliance on self-reported Likert-scale data, binarized at ≥ 4 , may oversimplify nuanced preferences, potentially obscuring subtle variations. Fourth, key predictors like technology trust, safety perceptions, and infrastructure readiness were omitted, despite their relevance to adoption. Fifth, the trade-off between XGBoost’s complexity and GAM’s interpretability may limit stakeholder accessibility, and GAM’s fixed knots may miss complex interactions. Sixth, the cross-sectional design fails to capture dynamic demand or evolving technology acceptance. Seventh, accessibility and equity for people with disabilities and low-income groups were not fully explored, particularly regarding specific technological needs or affordability barriers. Finally, the smaller sample size heightens overfitting risks, especially for XGBoost, despite LASSO and cross-validation efforts.

To address these limitations and advance autonomous public transport in rural settings, future research should pursue several directions. Longitudinal studies are needed to track evolving preferences, capturing seasonal and technological changes through panel surveys or repeated cross-sectional designs. Expanding to diverse global rural regions will enhance generalizability, comparing

areas with varying infrastructure and socio-economic profiles. Incorporating predictors like trust, safety, and infrastructure readiness via mixed-method approaches will deepen understanding of adoption barriers. Detailed analyses of accessibility and equity should explore design requirements for people with disabilities and affordability solutions, including subsidy cost-benefit analyses. Real-world pilots and agent-based simulations can validate findings, optimize fleet management, and test hybrid models. Advanced analytical methods, can improve accuracy and interpretability. Finally, developing policy frameworks with stakeholders to address regulatory, safety, and ethical issues will ensure equitable, scalable deployment, informing sustainable transport solutions for rural communities.

Author Contributions: Conceptualization, S.J., A.B. and J.B.; Methodology, S.J., A.B. and J.B.; Software, S.J.; Validation, S.J. and J.B.; Formal Analysis, S.J.; Investigation, S.J.; Resources, J.B.; Data Curation, S.J.; Writing – Original Draft Preparation, S.J.; Writing – Review & Editing, S.J., A.B. and J.B.; Visualization, S.J.; Supervision, A.B. and J.B.; Project Administration, J.B.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki and approved by the QUT Human Research Ethics Committee (Approval number: 9083-HE09, Date:13/09/2024).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: De-identified data is available upon request.

Acknowledgments: The first author gratefully acknowledges the scholarship support provided by the QUT to carry out this PhD research.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ADRT	Autonomous Demand Responsive Transit
AIC	Akaike Information Criterion
AUC	Area Under the Curve
AV	Autonomous Vehicle
DRT	Demand Responsive Transit
EDF	Effective Degrees of Freedom
EPP	Events Per Parameter
FMLM	First-Mile/Last-Mile
GAM	General Additive Models
GBMs	Gradient Boosting Machines
GLMs	Generalized Linear Models
LASSO	Least Absolute Shrinkage and Selection Operator
MaaS	Mobility-as-a-Service
REML	Restricted Maximum Likelihood
ROC	Receiver Operating Characteristic
SAV	Shared Autonomous Vehicle
SEQ	South-East Queensland
SHAP	Shapley Additive Explanation
TAFE	Technical and Further Education
VIF	Variance Inflation Factors
XGBoost	Extreme Gradient Boosting

References

1. Hanley JA. Characteristic (ROC) curve]. Radiology. 1982;743(2):29-36.

2. Fawcett T. An introduction to ROC analysis. *Pattern Recognition Letters*. 2006;27(8):861-74.
3. Friedman JH, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. *Journal of Statistical Software*. 2010;33(1):1 - 22.
4. Belsley DA, Kuh E, Welsch RE. *Regression diagnostics: Identifying influential data and sources of collinearity*; John Wiley & Sons; 2005.
5. Box GEP, Tidwell PW. Transformation of the Independent Variables. *Technometrics*. 1962;4(4):531-50.
6. Hastie T, Tibshirani R. Generalized additive models. *Statistical science*. 1986;1(3):297-310.
7. Cleveland WS, Devlin SJ. Locally Weighted Regression: An Approach to Regression Analysis by Local Fitting. *Journal of the American Statistical Association*. 1988;83(403):596-610.
8. Wood SN. *Generalized Additive Models: An Introduction with R*, Second Edition (2nd ed.) ed: Chapman and Hall/CRC.; 2017.
9. Hastie T, Tibshirani R, Friedman J. *The elements of statistical learning*. Springer series in statistics New-York; 2009.
10. Wood SN, Pya N, Säfken B. Smoothing parameter and model selection for general smooth models. *Journal of the American Statistical Association*. 2016;111(516):1548-63.
11. Marra G, Wood SN. Practical variable selection for generalized additive models. *Computational Statistics & Data Analysis*. 2011;55(7):2372-87.
12. Friedman JH. Greedy Function Approximation: A Gradient Boosting Machine. *The Annals of Statistics*. 2001;29(5):1189-232.
13. Chen T, Guestrin C, editors. *Xgboost: A scalable tree boosting system*. Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining; 2016.
14. Tibshirani R. Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society Series B (Methodological)*. 1996;58(1):267-88.
15. Hoerl AE, Kennard RW. Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*. 1970;12(1):55-67.
16. Bergstra J, Bengio Y. Random search for hyper-parameter optimization. *The journal of machine learning research*. 2012;13(1):281-305.
17. Kohavi R, editor *A study of cross-validation and bootstrap for accuracy estimation and model selection*. Ijcai; 1995: Montreal, Canada.
18. Prechelt L. Automatic early stopping using cross validation: quantifying the criteria. *Neural networks*. 1998;11(4):761-7.
19. Belkin M, Hsu D, Ma S, Mandal S. Reconciling modern machine-learning practice and the classical bias-variance trade-off. *Proc Natl Acad Sci U S A*. 2019;116(32):15849-54.
20. Lundberg SM, Lee S-I. A unified approach to interpreting model predictions. *Advances in neural information processing systems*. 2017;30.
21. Shapley LS. A value for n-person games. 1953.
22. Powers DM. Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation. *arXiv preprint arXiv:201016061*. 2020.
23. Davis J, Goadrich M. The relationship between Precision-Recall and ROC curves. Proceedings of the 23rd international conference on Machine learning; Pittsburgh, Pennsylvania, USA: Association for Computing Machinery; 2006. p. 233–40.
24. Davison A, Hinkley D. *Bootstrap Methods and Their Application*. *Journal of the American Statistical Association*. 1997;94.

25. Golbabaie, Yigitcanlar T, Paz A, Bunker J. Individual Predictors of Autonomous Vehicle Public Acceptance and Intention to Use: A Systematic Review of the Literature. *Journal of Open Innovation: Technology, Market, and Complexity*. 2020;6(4).
26. Iclodean C, Cordos N, Varga BO. Autonomous Shuttle Bus for Public Transportation: A Review. *Energies*. 2020;13(11).
27. Koskinen K, Mallat N, Raj Upreti B. Shared benefits and sustainable mobility – A case of autonomous bus. *Case Studies on Transport Policy*. 2024;18:101286.
28. Jayatilleke S, Bhaskar A, Bunker J. A Cross-Sectional Study on the Public Perception of Autonomous Demand-Responsive Transits (ADRTs) in Rural Towns: Insights from South-East Queensland. *Smart Cities*. 2025;8(3).
29. Othman K. Exploring the implications of autonomous vehicles: a comprehensive review. *Innovative Infrastructure Solutions*. 2022;7(2).
30. Schippl J, Truffer B, Fleischer T. Potential impacts of institutional dynamics on the development of automated vehicles: Towards sustainable mobility? *Transportation Research Interdisciplinary Perspectives*. 2022;14:100587.
31. Thorhauge M, Fjendbo Jensen A, Rich J. Effects of autonomous first- and last mile transport in the transport chain. *Transportation Research Interdisciplinary Perspectives*. 2022;15.
32. Li H, Jin Z, Cui H, Tu H. An exploration of the preferences and mode choice behavior between autonomous demand-responsive transit and traditional buses. *International Journal of Transportation Science and Technology*. 2024;15:81-101.
33. Xu Z, Zheng N. Integrating connected autonomous shuttle buses as an alternative for public transport – A simulation-based study. *Multimodal Transportation*. 2024;3(2):100133.
34. Golchin M, Grandhi A, Gore N, Pulugurtha SS, Ghasemi A. UNC Charlotte Autonomous Shuttle Pilot Study: An Assessment of Operational Performance, Reliability, and Challenges. *Machines*. 2024;12(11).
35. Golbabaie F, Paz A, Yigitcanlar T, Bunker J. Navigating autonomous demand responsive transport: stakeholder perspectives on deployment and adoption challenges. *International Journal of Digital Earth*. 2023;17(1).
36. Golbabaie F. Challenges and opportunities in the adoption of autonomous demand responsive transit (ADRT) by adult residents of South East Queensland: Queensland University of Technology; 2023.
37. Fadlseed S, Sarhadi P, Ramalingam S, Gan H, Kourtessis P, Jackman G, et al. Recent Advances in Demand Responsive Transport: Opportunities With Autonomous Bus Service—A System-of-Systems Overview. *IEEE Access*. 2025;PP:1-.
38. Sistig HM, Sinhuber P, Rogge M, Sauer DU. Optimizing Fleet Structure for Autonomous Electric Buses: A Route-Based Analysis in Aachen, Germany. *Sustainability*. 2024;16(10):4093.
39. Xu M. Addressing the Fleet Sizing Problem for Shared-and-Autonomous-Mobility Services. 2021.
40. Ciari F, Janzen M, Ziemlicki C. Chapter 8 - Planning shared automated vehicle fleets: Specific modeling requirements and concepts to address them. In: Antoniou C, Efthymiou D, Chaniotakis E, editors. *Demand for Emerging Transportation Systems*: Elsevier; 2020. p. 151-68.
41. Afkham M, Ramezani R, Shahparvari S. Balancing traffic flow in the congested mass self-evacuation dynamic network under tight preparation budget: An Australian bushfire practice. *Omega*. 2022;111:102658.
42. Partheepan S, Sanati F, Hassan J. Autonomous Unmanned Aerial Vehicles in Bushfire Management: Challenges and Opportunities. *Drones*. 2023;7(1).

43. Choi M, Min S, Kim J, Kim S, Kwak J, Lee S. Autonomous vehicle integration with public transit for congestion mitigation and energy efficiency. *Sustainable Energy Technologies and Assessments*. 2025;82:104476.
44. Alonso-González MJ, Liu T, Cats O, Van Oort N, Hoogendoorn S. The Potential of Demand-Responsive Transport as a Complement to Public Transport: An Assessment Framework and an Empirical Evaluation. *Transportation Research Record*. 2018;2672(8):879-89.
45. Golbabaie F, Dwyer J, Gomez R, Peterson A, Cocks K, Bubke A, et al. Enabling mobility and inclusion: Designing accessible autonomous vehicles for people with disabilities. *Cities*. 2024;154:105333.
46. Tang C, Liu J, Ceder A, Jiang Y. Optimisation of a new hybrid transit service with modular autonomous vehicles. *Transportmetrica A: Transport Science*. 2023;20:1-23.
47. Wu J, Kulcsár B, Selpi, Qu X. A modular, adaptive, and autonomous transit system (MAATS): An in-motion transfer strategy and performance evaluation in urban grid transit networks. *Transportation Research Part A: Policy and Practice*. 2021;151:81-98.
48. Maruyama R, Seo T. Integrated Public Transportation System with Shared Autonomous Vehicles and Fixed-Route Transits: Dynamic Traffic Assignment-Based Model with Multi-Objective Optimization. *International Journal of Intelligent Transportation Systems Research*. 2023;21(1):99-114.
49. Abdelwahed A, van den Berg PL, Brandt T, Ketter W. Balancing convenience and sustainability in public transport through dynamic transit bus networks. *Transportation Research Part C: Emerging Technologies*. 2023;151:104100.
50. Rau A, Tiana L, Jain M, Xie M. Dynamic Autonomous Road Transit (DART) For Use-case Capacity More Than Bus. *Transportation Research Procedia* 2019;41:812-23.
51. Leich G, Bischoff J. Should autonomous shared taxis replace buses? A simulation study. *Transportation Research Procedia* 2019;41:450–60.
52. Scheltes A, de Almeida Correia GH. Exploring the use of automated vehicles as last mile connection of train trips through an agent-based simulation model: An application to Delft, Netherlands. *International Journal of Transportation Science and Technology*. 2017;6(1):28-41.
53. Shen Y, Zhang H, Zhao J. Integrating shared autonomous vehicle in public transportation system: A supply-side simulation of the first-mile service in Singapore. *Transportation Research Part A: Policy and Practice*. 2018;113:125-36.
54. Sadler K. First and last mile: emerging autonomous public transport 2017 [Available from: <<https://www.eurotransportmagazine.com/22134/transport-extra/autonomous-public-transport/>>. .
55. Gurumurthy KM, Kockelman KM, Zuniga-Garcia N. First-Mile-Last-Mile Collector-Distributor System using Shared Autonomous Mobility. *Transportation Research Record: Journal of the Transportation Research Board*. 2020;2674(10):638-47.
56. Huang Y, Kockelman KM, Garikapati V, Zhu L, Young S. Use of Shared Automated Vehicles for First-Mile Last-Mile Service: Micro-Simulation of Rail-Transit Connections in Austin, Texas. *Transportation Research Record: Journal of the Transportation Research Board*. 2020;2675(2):135-49.
57. Lau ST, Susilawati S. Shared autonomous vehicles implementation for the first and last-mile services. *Transportation Research Interdisciplinary Perspectives*. 2021;11.
58. Zubin I, Van Oort N, Van Binsbergen A, Van Arem B. Deployment Scenarios for First/Last-Mile Operations With Driverless Shuttles Based on Literature Review and Stakeholder Survey. *IEEE Open Journal of Intelligent Transportation Systems*. 2021;2:322-37.

59. Stevens M, Correia GHdA, Scheltes A, van Arem B. An agent-based model for assessing the financial viability of autonomous mobility on-demand systems used as first and last-mile of public transport trips: A case-study in Rotterdam, the Netherlands. *Research in Transportation Business & Management*. 2022;45.
60. Dong Z, Chen C, Ouyang J, Yan X, Liao C, Chen X, et al. Understanding commuter preferences for shared autonomous electric vehicles in first-mile-last-mile scenario. *Transportation Research Part D: Transport and Environment*. 2025;140.
61. Zhang W, Jenelius E, Badia H. Efficiency of Connected Semi-Autonomous Platooning Bus Services in High-Demand Transit Corridors. *IEEE Open Journal of Intelligent Transportation Systems*. 2022;3:435-48.
62. Hatzenbühler J, Cats O, Jenelius E. Transitioning towards the deployment of line-based autonomous buses: Consequences for service frequency and vehicle capacity. *Transportation Research Part A: Policy and Practice*. 2020;138:491-507.
63. Kalambay P, Kitali A, Ngereza A, Kidando E, Ogungbire A. Autonomous taxis and ride-sharing vehicles: A social construct perspective for future mobility and infrastructure readiness. *Sustainable Cities and Society*. 2025;118:106060.
64. Bi C, Li Y, Gruyer D, Tu M. Who is more willing to use shared autonomous vehicles in first-mile-last-mile? A heterogeneity study on carbon incentive policy from China. *International Journal of Transportation Science and Technology*. 2024.
65. Abe R. Introducing autonomous buses and taxis: Quantifying the potential benefits in Japanese transportation systems. *Transportation Research Part A: Policy and Practice*. 2019;126:94-113.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.