OPTIMIZATION MODEL FOR URBAN PUBLIC TRANSPORT OPERATIONS USING MACHINE LEARNING-BASED DEMAND PREDICTION

Joshua Banua[1*], Lucia Ingrid Regina Lefrandt[2], Semuel Yacob Recky Rompis[3]
[1,2,3] Graduate Program, Master of Engineering, Universitas Sam Ratulangi, Indonesia
*Corresponding Author: joshua.banua@gmail.com

ABSTRACT

Angkot, a route-based urban paratransit or public minibus service, remains an important component of daily mobility in Manado City. However, its predominantly supply-based operating pattern does not systematically adjust vehicle deployment to hourly passenger demand, which can produce low occupancy during off-peak periods and insufficient capacity during demand peaks. This study integrates machine-learning-based hourly passenger-demand prediction with an integer-constrained fleet-sizing model for the Paal Dua-Pasar 45 route. Direct observations were collected for 21 days, from 25 May to 14 June 2026, through an on-board survey of one angkot and a separate vehicle-headway survey; 1,538 passenger boardings were directly recorded. Random Forest and XGBoost were compared using a chronological 80:20 holdout and Leave-One-Day-Out Cross-Validation (LODO). Random Forest remained the better-performing model after tuning (LODO $R^2 = 0.462$; MAE = 0.76) compared with XGBoost ($R^2 = 0.421$; MAE = 0.83). The selected predictions were transformed from vehicle-level boardings to estimated corridor demand using observed hourly vehicle frequency and were then entered into the fleet-sizing model. The model-derived corridor demand averaged 57 passengers per hour, and the resulting fleet requirement averaged six vehicles, ranging from five to nine vehicles. Because corridor demand was inferred from one directly surveyed vehicle and predictive performance was moderate, these fleet values should be interpreted as route-specific planning estimates rather than externally validated operating prescriptions. The study demonstrates a practical framework for linking demand prediction with adaptive fleet allocation while highlighting the need for multi-vehicle validation and uncertainty-aware deployment.

Keywords: Fleet Optimization, Machine Learning, Passenger Demand, Random Forest, XGBoost

INTRODUCTION

Angkot is a route-based urban paratransit or public minibus service and is one of the principal urban transport modes in Manado, serving daily movement between terminals, residential areas, and activity centers. Unlike a fully scheduled fixed-route bus system, angkot operations commonly exhibit flexible departure behavior and variable spacing between vehicles. The existing operating pattern remains predominantly supply-based and does not directly adjust the number of operating vehicles to hourly demand. This condition can create oversupply on some routes, low vehicle occupancy, unstable driver income, and operational inefficiency (Junior, 2016; Hillary, 2016). Vehicles circulating without demand-responsive control may also contribute to congestion on urban roads (Ritonga, 2015).

Passenger demand changes by time, location, and travel direction. During peak periods, insufficient capacity may reduce service quality, whereas maintaining the same number of vehicles during off-peak periods may create low load factors and empty roaming. In public transport operations, fleet size, service frequency, headway, waiting time, and load factor are interrelated, and the number of operating vehicles should be adjusted to balance capacity and service quality (Vuchic, 2005; Ramanayya, 2014).

Machine learning provides a data-driven approach for capturing nonlinear relationships among temporal, spatial, and external variables. Recent literature shows growing use of machine-learning methods for public-transport demand prediction and emphasizes their potential contribution to service planning, while also identifying persistent challenges related to heterogeneous datasets, limited reproducibility, and operational translation of forecasts (Rocco di Torrepadula et al., 2024). Random Forest combines multiple independently constructed trees through bagging and is relatively stable on tabular datasets, whereas

XGBoost builds trees sequentially to correct previous errors through gradient-based optimization and regularization (Breiman, 2001; Chen & Guestrin, 2016; Grinsztajn et al., 2022). The present study deliberately compares these two ensemble-tree mechanisms because the available data are structured as tabular hour-zone-direction observations and because their outputs can be directly connected to a fleet-allocation rule. The comparison is therefore a focused comparison between two ensemble models and should not be interpreted as evidence that either model is superior to all statistical or machine-learning baselines.

The Paal Dua-Pasar 45 route was selected because it connects the Paal Dua area, which functions as an urban service subcenter and terminal location, with Wenang, which forms part of Manado City center. Restricting the study to one route allowed passenger demand, headway, cycle time, zone, and travel direction to be observed consistently and in detail. The research gap addressed here is not passenger-demand forecasting alone or conventional fleet calculation alone, but the operational linkage between short-horizon, route-level demand prediction and an hourly fleet-sizing rule in a semi-formal angkot context. The contribution is an integrated workflow that (1) compares Random Forest and XGBoost for vehicle-level hourly passenger prediction, (2) transforms the selected prediction into an estimated corridor-demand measure using observed service frequency, (3) determines the minimum integer fleet satisfying capacity and waiting-time constraints, and (4) examines how demand uncertainty can shift integer fleet recommendations. Accordingly, this study aims to compare Random Forest and XGBoost for hourly passenger prediction, determine a route-specific fleet requirement based on predicted demand, and evaluate the resulting implications for headway, waiting time, and load factor.

METHOD

Research Location and Data Collection

The study was conducted on the Paal Dua-Pasar 45 angkot route in Manado City. Field data were collected for 21 days, from 25 May to 14 June 2026. The corridor was divided into three administrative zones: Paal Dua, Tikala, and Wenang. Passenger observations were separated into inbound and outbound directions.

Figure 1. Paal Dua-Pasar 45 Angkot Route

Passenger data were obtained through an on-board survey conducted on one angkot. Surveyors recorded every passenger boarding and alighting by hour, zone, and travel direction. The modeling unit was an hourly zone-direction observation, the response variable was the observed passenger-boarding count on the surveyed vehicle, and the candidate predictors comprised hour, day, holiday status, weather, direction, and zone. A separate headway survey was conducted at one observation point by recording vehicle arrival time and license plate number. These observations were used to estimate hourly vehicle arrivals, headway variability, and a representative cycle time. Importantly, the directly observed passenger counts represent one surveyed vehicle rather than the entire corridor. Consequently, corridor demand used later in the optimization stage is a model-derived estimate based on a scaling assumption and is not an independent corridor passenger count.

Machine Learning Modeling and Validation

The prediction model was developed in Python using Random Forest and XGBoost implementations for tabular regression. Records were organized at the hour-zone-direction level. Categorical variables were encoded using information derived from the training data and the resulting feature structure was aligned with the testing data to prevent information leakage from held-out observations. Random Forest and XGBoost were selected as two contrasting ensemble-tree approaches that can represent nonlinear interactions without requiring the large sample sizes generally associated with deep-learning architectures. A simple statistical baseline was not retained in the original analytical workflow; therefore, the reported results are interpreted as a comparison between the two ensemble methods rather than a comprehensive benchmark against all possible forecasting alternatives.

Initial evaluation used a chronological 80:20 split, preserving the temporal order of the observations so that earlier records were used for model development and later records for testing. This evaluation was complemented by Leave-One-Day-Out Cross-Validation (LODO), with survey date as the grouping variable. Because the survey covered 21 days, LODO was conducted in 21 iterations: one date was held out for testing and the remaining 20 dates were used for training. Out-of-fold predictions from all iterations were combined to assess generalization to days not used during each training iteration. Hyperparameter tuning produced the final configurations reported in Table 1. The complete machine-readable tuning search log and exact software-package versions were not preserved in the archived manuscript materials, and a fully nested cross-validation design was not implemented. Consequently, tuned LODO scores are treated as model-selection/generalization evidence within this dataset rather than as an unbiased estimate of performance on an external population. Future replication should archive the complete search space, random seeds, package versions, and use nested or rolling-origin validation.

Table 1. Hyperparameter Configurations Used in the Final Models

Model	Hyperparameter	Initial	Tuned
Random Forest	n_estimators	200	300
Random Forest	max_depth	Unlimited	4
Random Forest	min_samples_leaf	1	8
XGBoost	n_estimators	300	200
XGBoost	max_depth	6	2
XGBoost	learning_rate	0.10	0.05
XGBoost	min_child_weight	1	5
XGBoost	subsample	0.80	0.80
XGBoost	colsample_bytree	0.80	0.80

Corridor Demand Transformation and Fleet Optimization

The initial machine-learning output represented predicted passenger boardings on the single surveyed vehicle rather than directly observed total corridor demand. For hour t, let $\hat{p}(t,z,d)$ denote the predicted number of boardings for zone z and direction d on the surveyed-vehicle scale, and let $F_{obs}(t)$ denote the observed number of angkot arrivals per hour from the headway survey. Estimated hourly corridor demand was calculated as $D(t) = F_{obs}(t) \times \Sigma_{z,d}\, \hat{p}(t,z,d)$. This transformation assumes that the surveyed vehicle is broadly representative of other vehicles operating during the corresponding hour. Because passenger loads can vary across vehicles, drivers, departure positions, and traffic conditions, this scaling assumption is a principal source of uncertainty. The resulting D(t) values are therefore described throughout the paper as estimated or transformed corridor demand, not as directly measured corridor totals.

Fleet size was determined using a one-variable integer-constrained optimization that minimizes the number of physical vehicles while satisfying capacity and waiting-time constraints. Let N(t) be the integer number of vehicles, C the representative cycle time in minutes, K the passenger capacity per vehicle, L_{max} the planning load-factor ceiling, cv the coefficient of variation of headway, and W_{max} the maximum planning waiting-time tolerance. Service frequency is $f(t) = 60N(t)/C$ vehicles per hour, scheduled-equivalent headway is $h(t) = C/N(t)$ minutes, and corrected expected waiting time is approximated by $W(t) = [h(t)/2] \times [1 + cv^2]$, which accounts for irregular headways (Esfeh et al., 2021). The optimization objective is min N(t), subject to $D(t) \leq f(t) \times K \times L_{max}$ and $W(t) \leq W_{max}$, with N(t) restricted to positive integers. After algebraic rearrangement, the constraints are equivalent to $N(t) \geq D(t)C/[60KL_{max}]$ and $N(t) \geq C(1 + cv^2)/(2W_{max})$. Thus, the operational recommendation is the smallest integer satisfying both lower bounds. This formulation makes explicit the distinction between physical fleet size and hourly service frequency.

The parameters in Table 2 are scenario inputs derived from the survey and the study's operational planning assumptions. The capacity of nine passengers represents the vehicle capacity adopted in the analysis. The maximum load factor of 0.85 is used as a planning ceiling to preserve approximately 15% reserve capacity rather than as a universal regulatory threshold, while the 10-minute waiting-time value is the service-quality tolerance adopted for this optimization scenario. The 41-minute cycle time is a representative observed cycle value and is treated as constant in the current model; time-of-day variation in cycle time is therefore not explicitly modeled. The headway coefficient of variation (cv = 1.12) was derived from observed vehicle arrivals and represents highly irregular spacing between vehicles.

Table 2. Operational Parameters of the Fleet Optimization Model

Parameter	Value
Vehicle capacity	9 passengers/vehicle
Maximum planning load factor	0.85
Maximum planning waiting-time tolerance	10 minutes
Representative cycle time	41 minutes
Headway coefficient of variation	1.12

The headway coefficient of variation of 1.12 indicates that the standard deviation of headway exceeded the mean headway, reflecting highly irregular vehicle arrivals. Under the waiting-time approximation $W = (h/2)(1 + cv^2)$, cv = 1.12 produces a waiting-time multiplier of approximately 2.25 relative to h/2 under perfectly regular service. Demand sensitivity was evaluated using multiplicative scenarios of ±10%, ±25%, ±50%, and ±100% around the transformed demand. This sensitivity analysis does not constitute a formal probabilistic prediction interval, but it provides a deterministic assessment of how forecast or scaling error may alter the integer fleet recommendation.

Research Procedure

The following flowchart summarizes the main steps of the research:

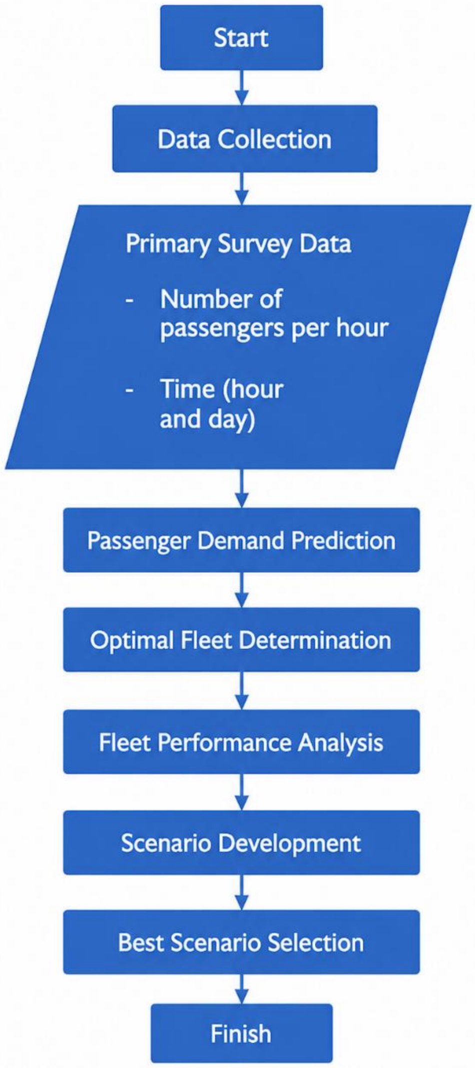


Figure 2. Research procedure flowchart

RESULTS AND DISCUSSION

Passenger Survey Results

The on-board survey directly recorded 1,538 passenger boardings on the surveyed vehicle, with an average of approximately 74 directly observed boardings per survey day. These figures are empirical vehicle-level observations and should be distinguished from the corridor-demand estimates presented later, which are derived by combining model predictions with observed vehicle frequency. Among the directly observed boardings, the highest contribution originated from the Paal Dua zone, followed by Wenang, while Tikala functioned mainly as a transition zone. Passenger movement was generally characterized by travel between Paal Dua and Wenang.

Figure 3. Research Zoning on the Paal Dua-Pasar 45 Corridor

Table 3. Passenger Boardings by Research Zone

Zone	Passengers	Share
Paal Dua	795	51.69%
Tikala	122	7.93%
Wenang	621	40.38%
Total	1,538	100%

Paal Dua recorded 795 boarding passengers, or 51.69% of the directly observed sample, and was dominated by outbound movement toward Tikala and Wenang. Tikala recorded 122 passengers, or 7.93%, confirming its role as a transition zone with relatively limited boarding and alighting. Wenang recorded 621 passengers, or 40.38%, consistent with its concentration of commercial activities around Pasar 45. The rounded percentages sum to 100.00%. Repeated observations of no pickup or no boarding, particularly during several midday and afternoon periods, indicate that passenger demand was not evenly distributed throughout the operating day.

Machine Learning Performance

The chronological 80:20 holdout evaluation showed that Random Forest outperformed XGBoost on the later held-out observations. Random Forest generated lower MSE and RMSE and a higher coefficient of determination. Its RMSE of 1.070 corresponds to a typical error of approximately one passenger for an observation unit defined by hour, zone, and direction. The moderate R^2 values for both models indicate that a substantial share of variation remains unexplained; therefore, the comparison supports selection of the better of the two tested models but does not imply near-complete predictability of passenger demand.

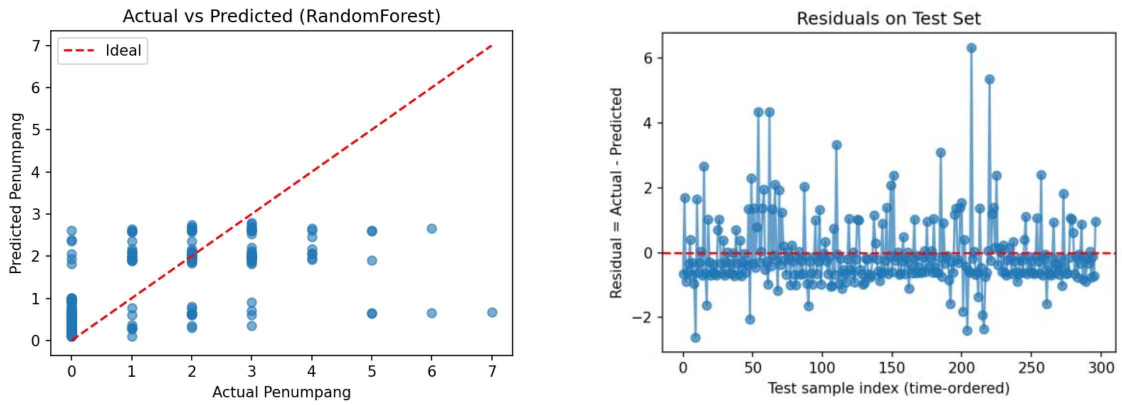
Table 4. Model Performance Using the 80:20 Split

Model	MSE	RMSE	R^2
Random Forest	1.146	1.070	0.431
XGBoost	1.215	1.102	0.397

After hyperparameter tuning, LODO evaluation produced an R^2 of 0.462 and an MAE of 0.76 for Random Forest, compared with an R^2 of 0.421 and an MAE of 0.83 for XGBoost. Random Forest therefore remained the selected model because it maintained higher explanatory performance and lower absolute prediction error on days excluded from each training fold. Nevertheless, an R^2 of 0.462 represents moderate predictive performance rather than high certainty. Because the tuning procedure was not fully nested within an external validation design, these results should be interpreted as within-study model-selection evidence and should be externally validated before operational deployment.

Table 5. Leave-One-Day-Out Cross-Validation Results

Model	Configuration	R^2	MAE
Random Forest	Initial	0.252	0.84
Random Forest	Tuned	0.462	0.76
XGBoost	Initial	0.186	0.89
XGBoost	Tuned	0.421	0.83

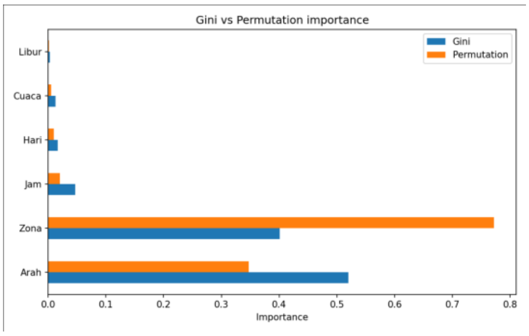


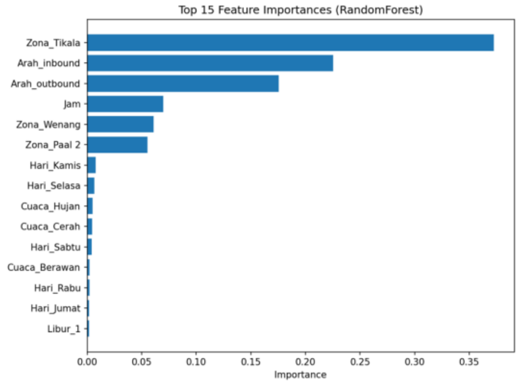
(a) (b)

Figure 4. Random Forest Model Performance: (a) Actual versus Predicted Values and (b) Residuals

The actual-versus-predicted plot shows a positive relationship toward the ideal prediction line, although the model tended to underestimate several high passenger observations and overestimate some very low observations. Most residuals were concentrated within a relatively small passenger-count range, consistent with the reported MAE and RMSE. The tuned LODO model explains 46.2% of observed passenger-demand variation, leaving 53.8% unexplained by the included predictors. Potential omitted influences include traffic conditions, special events, passenger-waiting behavior, competition with other modes, vehicle-specific differences, and day-to-day operational disruptions. From an operational perspective, the model is therefore more appropriate as a decision-support input that should be updated with current counts than as a stand-alone deterministic scheduling rule.

Feature Importance and Temporal Response





(a) (b)

Figure 5. Random Forest Feature Importance: (a) Gini versus Permutation Importance and (b) Ranked Features

The feature-importance analysis indicated that spatial variables, particularly zone and travel direction, received most of the importance weight in the fitted Random Forest, while hour and day received smaller weights. This pattern is consistent with the concentration of observed boardings in Paal Dua and Wenang. However, impurity-based Random Forest importance can be biased by predictor scale, category structure, and correlation among encoded variables (Strobl et al., 2007). Accordingly, the importance results are interpreted only as model-specific associations, not causal percentages or definitive evidence that spatial factors explain 90% of passenger behavior. The permutation-importance visualization is used as a qualitative robustness check; future work should report repeated permutation importance with uncertainty intervals and assess stability across validation folds.

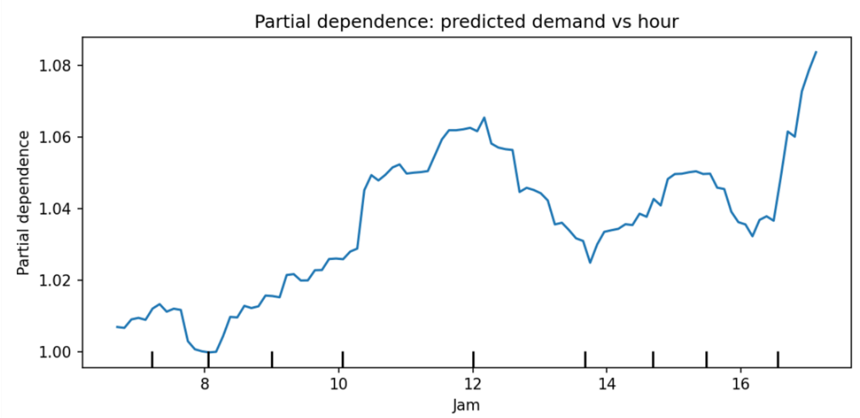


Figure 6. Partial Dependence of Predicted Passenger Demand on Hour of Day

The partial-dependence result shows that predicted demand changes nonlinearly by hour. The response decreases around 08.00, rises toward midday, decreases after the midday period, increases again around 15.00, and rises sharply toward the late-afternoon period. This pattern demonstrates why a constant fleet size throughout the day does not reflect the observed temporal variation in passenger demand.

Fleet Optimization Results

After transforming the selected vehicle-level predictions into estimated corridor demand using observed hourly vehicle frequency, the model-derived average demand was 57 passengers per hour and the maximum was 86 passengers per hour. These values are transformed estimates rather than directly counted corridor totals. When entered into the integer fleet-sizing model, the average requirement was six physical vehicles, the maximum requirement was nine vehicles, and the waiting-time constraint retained a minimum of five vehicles during low-demand hours. Higher-demand periods were governed mainly by the capacity constraint. Because the demand input is predicted and scaled, the resulting fleet values should be interpreted as scenario-based planning outputs.

Table 6. Selected Hourly Optimization Results for an Illustrative Post-Survey Model Scenario on 8 July 2026

Time	Demand (pass./h)	Fleet	Headway (min)	Waiting Time (min)	Load Factor
07.00	81.36	8	5.13	5.78	0.77
15.00	54.78	5	8.20	9.24	0.83
18.00	15.01	5	8.20	9.24	0.23

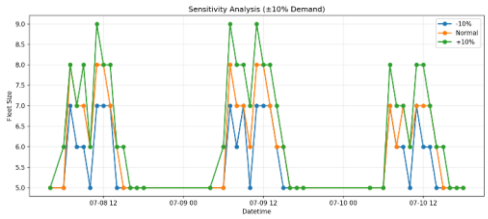
Table 6 presents an illustrative post-survey model scenario dated 8 July 2026; it is not an additional field-survey date and therefore falls outside the 25 May-14 June 2026 observation period. At 07.00, an estimated demand of 81.36 passengers per hour resulted in eight vehicles, a service frequency of 11.71 vehicles per hour, a headway of 5.13 minutes, a corrected waiting time of 5.78 minutes, and a load factor of 0.77. At 15.00, five vehicles serving 54.78 passengers per hour yield a headway of 8.20 minutes and a corrected waiting time of 9.24 minutes, with a load factor of 0.83. At 18.00, the same five-vehicle fleet serves only 15.01 passengers per hour, also producing a headway of 8.20 minutes and a corrected waiting time of 9.24 minutes but a much lower load factor of 0.23. The comparison demonstrates that the same physical fleet can satisfy the waiting-time constraint at very different occupancy levels.

The representative cycle time of 41 minutes and headway coefficient of variation of 1.12 are central to distinguishing physical fleet size from service frequency. With five vehicles, $f = 60(5)/41 = 7.32$ vehicles per hour and $h = 41/5 = 8.20$ minutes. Applying the irregular-headway correction gives $W = (8.20/2)(1 + 1.12^2) = 9.24$ minutes. This result illustrates that service improvement does not always require additional vehicles: departure control, restrictions on excessive ngetem time, and better spacing between vehicles can reduce actual headway variability and passenger waiting. However, the current model uses a single representative cycle time and does not explicitly model congestion-induced cycle-time variation.

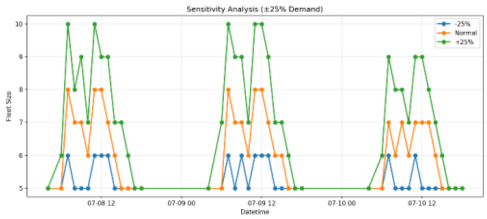
Uncertainty and Fleet-Threshold Analysis

Because fleet size is an integer decision, prediction or scaling errors matter most when estimated demand lies close to a capacity threshold. Under the adopted parameters, the maximum demand served at a 0.85 load-factor ceiling is $D_{max}(N) = [60N/41] \times 9 \times 0.85$. This corresponds to approximately 55.98, 67.17, 78.37, 89.56, and 100.76 passengers per hour for fleets of five through nine vehicles, respectively. The 15.00 estimate of 54.78 passengers per hour is only about 1.20 passengers per hour below the five-vehicle capacity threshold; an upward demand error of approximately 2.2% would therefore change the recommendation from five to six vehicles. Similarly, the average transformed demand of 57 passengers per hour lies just above the five-vehicle threshold and consequently produces a six-vehicle recommendation. These threshold effects show why point forecasts should not be interpreted as exact operational requirements. In practice, the recommended fleet should be expressed as a decision band and updated using current passenger counts, especially when predicted demand is near an integer-fleet threshold.

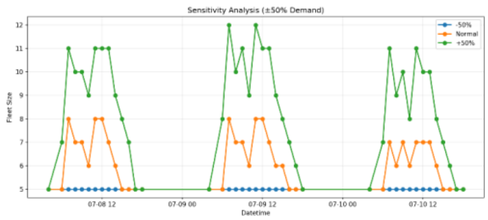
Demand Sensitivity and Comparison with Previous Studies



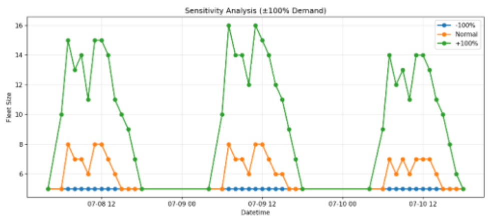
(a)



(b)



(c)



(d)

Figure 7. Passenger-Demand Sensitivity Analysis: (a) ±10%, (b) ±25%, (c) ±50%, and (d) ±100%

The deterministic sensitivity analysis confirms that the fleet recommendation responds to changes in transformed passenger demand while maintaining the adopted waiting-time standard. A 50% increase in demand requires 10 vehicles, while a 100% increase requires up to 16 vehicles. When demand decreases, the fleet does not fall below five vehicles because the waiting-time constraint becomes binding. The ±10%, ±25%, ±50%, and ±100% scenarios should be interpreted as stress-test bands for demand or scaling uncertainty rather than formal statistical prediction intervals. Future work should estimate probabilistic prediction intervals or bootstrap distributions and propagate them through the fleet-sizing model so that operators can evaluate the probability that a given fleet size will violate capacity or waiting-time constraints.

The survey findings support earlier local studies reporting excess angkot supply on Manado routes. Hillary (2016) reported a low existing load factor on the Paal Dua-Lapangan route, while Junior (2016) identified oversupply on the Paal Dua-Politeknik route. The present study differs by producing hourly, demand-responsive fleet estimates rather than one aggregate fleet value for an entire route and by explicitly linking predicted demand with capacity, cycle time, headway irregularity, waiting time, and load factor. At the broader international level, a recent systematic review shows that machine-learning-based public-transport demand prediction is increasingly used to support planning and management, but also highlights heterogeneity in data sources and a need for more reproducible operational applications (Rocco di Torrepadula et al., 2024). More data-rich studies, including multimodal and time-series approaches, demonstrate the potential value of incorporating richer temporal and external information (Wan et al., 2024). In contrast, the present contribution is a small-data, route-level case for a semi-formal minibus context; its practical relevance lies in the transparent connection between prediction and fleet thresholds, while its generalizability is necessarily more limited.

Random Forest performed better than XGBoost in this 21-day survey dataset. This differs from the Transjakarta study reported by Wijaya (2025), in which a gradient-boosted approach performed better on a substantially larger historical dataset. The difference is plausibly related to data scale and structure: the present study uses hour-zone-direction observations over a limited survey period, whereas larger automated datasets can support more complex temporal relationships. The result is consistent with the practical suitability of tree-based methods for tabular data (Grinsztajn et al., 2022), but the absence of a retained simple statistical baseline means that the study can only conclude that Random Forest was the better of the two tested ensemble models. Broader benchmarking against persistence, historical-mean, linear, and time-series baselines is required before claiming general predictive superiority.

Limitations and Future Research

This study has several limitations that directly affect interpretation and transferability. First, passenger boardings were directly observed on only one angkot over 21 days on a single route; total corridor demand was therefore estimated by scaling vehicle-level predictions with observed service frequency rather than validated against independent corridor counts. Second, the representativeness assumption may be affected by vehicle-specific passenger loads, driver behavior, departure position, traffic, and competition among vehicles. Third, the predictive model achieved only moderate LODO explanatory performance, and the original analytical workflow did not retain a simple forecasting baseline, a fully nested validation procedure, the complete hyperparameter search log, or exact software-package versions. Fourth, the optimization uses a representative 41-minute cycle time and deterministic operational parameters, whereas actual cycle time, headway variability, and capacity conditions can change by time of day. Fifth, the demand-sensitivity analysis provides deterministic scenario bands rather than formal probabilistic uncertainty intervals. Future research should collect simultaneous data from multiple vehicles and/or terminal or automated passenger-counting points, integrate GPS and traffic conditions, benchmark ensemble models against simple statistical and time-series baselines, use nested time-aware validation, estimate probabilistic demand intervals, and incorporate those intervals into robust or stochastic fleet optimization. External validation on additional Manado routes and in other paratransit systems is required before broader policy implementation.

CONCLUSIONS

Within the two ensemble models evaluated, Random Forest provided better passenger-demand prediction than XGBoost. In the initial chronological 80:20 evaluation, Random Forest produced an MSE of 1.146, an RMSE of 1.070, and an R^2 of 0.431, compared with an MSE of 1.215, an RMSE of 1.102, and an R^2 of 0.397 for XGBoost. After tuning and Leave-One-Day-Out Cross-Validation, Random Forest achieved an R^2 of 0.462 and an MAE of 0.76, compared with an R^2 of 0.421 and an MAE of 0.83 for XGBoost. These results support Random Forest as the preferred model among the two tested approaches, while the moderate R^2 indicates that substantial demand variation remains unexplained.

After scaling vehicle-level predictions by observed hourly vehicle frequency, the estimated corridor demand averaged 57 passengers per hour and reached a maximum of 86 passengers per hour. The resulting integer fleet-sizing model produced an average requirement of six vehicles, a maximum of nine vehicles, and a minimum of five vehicles under low-demand conditions because of the waiting-time constraint. Sensitivity analysis indicated that a 50% demand increase would require 10 vehicles and a 100% increase up to 16 vehicles. These are model-derived route-specific scenario outputs rather than directly observed corridor requirements.

Demand-based hourly fleet allocation has the potential to reduce oversupply during low-demand periods while retaining capacity during high-demand periods. For the specific Paal Dua-Pasar 45 scenario and parameter assumptions studied here, the model generally retains at least five vehicles and increases the fleet when estimated demand crosses capacity thresholds. These values should not be adopted as fixed prescriptions without external validation because demand was scaled from one surveyed vehicle and the forecast contains material uncertainty. Operational use should therefore combine the model with current passenger counts, departure control, restrictions on excessive ngetem time, and monitoring of vehicle spacing. The main contribution is the integrated and reproducible decision logic linking predicted demand to capacity and waiting-time constraints; broader implementation requires multi-vehicle data, time-aware external validation, uncertainty quantification, and testing on additional routes.

REFERENCES

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

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 785-794. https://doi.org/10.1145/2939672.2939785

Esfeh, M. A., Wirasinghe, S. C., Saidi, S., & Kattan, L. (2021). Waiting time and headway modelling for urban transit systems—A critical review and proposed approach. Transport Reviews, 41(2). https://doi.org/10.1080/01441647.2020.1806942

Grinsztajn, L., Oyallon, E., & Varoquaux, G. (2022). Why do tree-based models still outperform deep learning on typical tabular data? Advances in Neural Information Processing Systems, 507-520.

Hillary, L. (2016). Analisa kebutuhan angkutan umum perkotaan di Manado (Studi kasus: Trayek Paal Dua-Lapangan). Jurnal Sipil Statik, 283.

Hillier, F. S., & Lieberman, G. J. (2015). Introduction to operations research. New York: McGraw-Hill Education.

Junior, N. P. (2016). Analisis kebutuhan angkutan umum di Kota Manado (Studi kasus: Paal Dua-Politeknik). Jurnal Sipil Statik, 373.

Osuna, E. E., & Newell, G. F. (1972). Control strategies for an idealized public transportation system. Transportation Science, 6(1), 52-72. https://doi.org/10.1287/trsc.6.1.52

Ramanayya, A. V. (2014). Public transport planning and management in developing countries. Taylor & Francis Group.

Ritonga, D. (2015). Analisa biaya transportasi angkutan umum dalam Kota Manado akibat kemacetan lalu lintas (Studi kasus: Angkutan umum trayek Pusat Kota 45-Malalayang). Jurnal Sipil Statik, 66-67.

Strobl, C., Boulesteix, A.-L., Zeileis, A., & Hothorn, T. (2007). Bias in random forest variable importance measures: Illustrations, sources and a solution. BMC Bioinformatics, 8, 25. https://doi.org/10.1186/1471-2105-8-25

Vuchic, V. R. (2005). Urban transit: Operations, planning, and economics. John Wiley & Sons.

Wijaya, I. M. (2025). Pemodelan prediksi volume penumpang Transjakarta menggunakan regresi pada algoritma machine learning. Journal of Informatics, Electrical, and Electronics Engineering, 105.

Rocco di Torrepadula, F., Napolitano, E. V., Di Martino, S., & Mazzocca, N. (2024). Machine Learning for public transportation demand prediction: A systematic literature review. Engineering Applications of Artificial Intelligence, 137, 109166. https://doi.org/10.1016/j.engappai.2024.109166

Wan, L., Cheng, W., & Yang, J. (2024). Optimizing metro passenger flow prediction: Integrating machine learning and time-series analysis with multimodal data fusion. IET Circuits, Devices & Systems, 2024. https://doi.org/10.1049/2024/5259452