Journal of Scientific Insights
E-ISSN: 3062-8571
DOI: https://doi.org/10.69930/jsi.v2i4.500
Research Article

ScienceTech Group

Vol. 2 (4), 2025
Page: 400-411

Forecasting Soekarno-Hatta Airport Rail Link Passengers using Linear Regression and Exponential Smoothing Methods

Maulida Butar Butar *, Alsen Medikano, Yahya Zulkarnain

Department of Industrial Engineering, Faculty of Engineering, Gunadarma University, Indonesia
*Email (corresponding author): maulida_b@staff.gunadarma.ac.id

Abstract. *This article investigated the passenger demand forecasting for the Soekarno-Hatta Airport Rail Link, utilizing linear regression and exponential smoothing techniques. Given the increased significance of accurate demand forecasts in urban transportation systems, this research aims to enhance operational efficiency and service quality. Historical passenger data was analyzed from January 2024 to April 2025, revealing an upward trend in rid-ership. Three forecasting methods – linear regression, single exponential smoothing, and double exponential smoothing – were employed and compared using key accuracy metrics such as Mean Absolute Deviation (MAD), Mean Squared Error (MSE), and Mean Absolute Percentage Error (MAPE). The results indicate that linear regression yields the lowest error rates, making it the most effective method for forecasting future passenger numbers. En-hanced forecasting models, as proposed in this study, can significantly contribute to urban transportation planning and service optimization. From the research, indicate that the line-ar regression method is better in terms of forecasting accuracy, making it the preferred choice for predicting passenger demand for the airport rail service. The forecasts indicate a consistent and gradual increase in passenger numbers, starting from 770,025 in period 17 and reaching 840,457 by period 24. This steady growth pattern aligns with the upward trend observed in previous historical data, suggesting that increasing access to and aware-ness of the airport rail service will continue to drive ridership. The forecasts assist in opera-tional decision-making, allowing for adequate planning of resources, scheduling adjust-ments, and infrastructure needs to accommodate the projected growth in demand.*

Keywords: *Passenger demand, forecasting methods, linear regression, exponential smoothing, soe-karno-hatta airport rail link*

1. Introduction

The conception of a railway connection to Soekarno-Hatta International Airport (CGK) arose from the urgent need to ease the traffic congestion in Jakarta and enhance access to the airport. The development of urban rail sys-tems is crucial for improving public transportation efficiency and alleviating transportation issues in metropolitan areas. Public transportation is promoted and supported as a means to alleviate traffic issues. In this context, dependable, fast, and comfortable urban rail systems emerge as a leading solution for mass transit. (Abdallah, 2023; Tang et al., 2022). The Soekarno-Hatta Air-port Rail Link plays a critical role in connecting Indonesia's primary interna-tional gateway to the greater Jakarta region, conceived to reduce terrestrial traffic congestion and to enhance accessibility for millions of travelers. The rail link is a vital component of the region's integrated transportation network. Construction for the airport rail link officially commenced in 2015 as part of a comprehensive effort to improve the overall transportation infrastructure in Jakarta. The airport train service was launched on December 2, 2017, facilitat-ing travel between Soekarno-Hatta Airport and

Received: 30 July 2025 / Accepted: 18 August 2025 / Available online: 19 August 2025

Manggarai Station in central Jakarta, spanning approximately 38 kilometers. This train service enables rap-id transportation for passengers heading to and from the city, significantly shortening travel times compared to traditional road transport. Known as Kere-ta Bandara, the railway connection provides a convenient mobility option for travelers. Route for Soekarno-Hatta airport rail line could be seen at Figure 1, start from Manggarai – BNI City – Duri – Rawabuaya – Batu Ceper and end at Soekarno-Hatta airport.

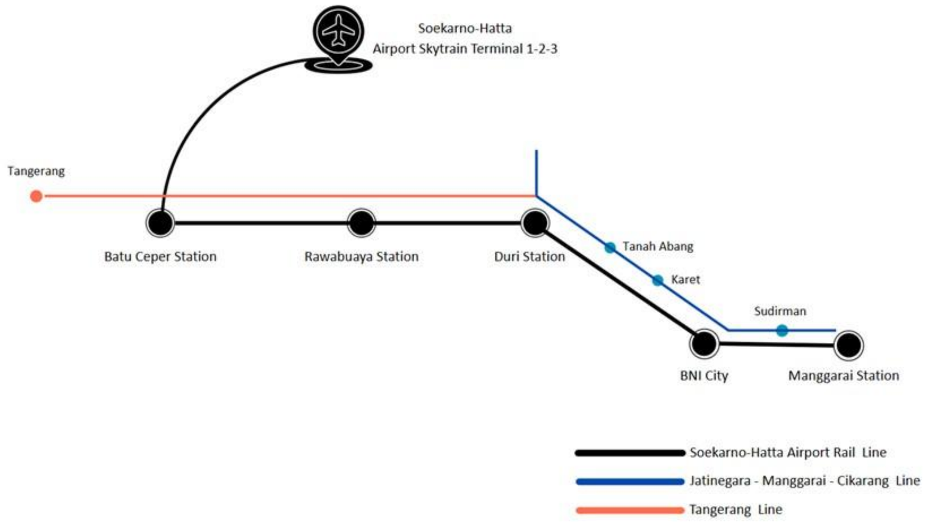

Figure 1. Soekarno-Hatta Airport Rail Route

This research focuses on accurately forecasting passenger demand on the rail link in order to streamline operational efficiency and support future infra-structure planning initiatives in one of Southeast Asia's busiest urban regions. Accurate forecasting of passenger numbers is essential for optimizing opera-tional efficiency and enhancing service quality (Chen, Kockelman, & Zhao, 2015; Jiao, Li, Sun, Hou, & Ibrahim, 2016). Utilizing statistical techniques such as linear regression and exponential smoothing allows for effective anal-ysis and projection of historical data into future trends (Shastri et al., 2018; Zhu, Chen, Wang, Yu, & Tang, 2023). In a dynamic environment such as the Soekarno-Hatta corridor, a miscalculation in demand forecasting can lead to either service underutilization or overcrowding, both of which adversely af-fect passenger satisfaction and operational profitability.

Recognizing these challenges, this study evaluates statistical forecasting techniques—specifically linear regression and exponential smoothing—to project short-term and long-term demand patterns. By benchmarking these methods against the backdrop of previous research in rail passenger forecast-ing systems, the study aims to establish reliable models that improve decision-making processes and long-term planning in a region characterized by rapid urban growth and evolving travel behavior.

The following sections detail the relevant literature, methodology, and the analysis of forecasting techniques. The operational importance of these meth-ods for the Soekarno-Hatta Airport Rail Link, drawing on insights from previ-ous studies that have underlined the benefits of passenger demand forecasting in diverse rail systems around the world will be discussed. This research thus lays a foundation for enhancing rail operations through a unified and system-atic forecasting approach that could lead to significant improvements in schedule planning, cost reduction, and service quality enhancement.

2. Literature Review

The field of passenger demand forecasting for rail systems has evolved significantly over recent years. Earlier research primarily relied on simple statistical techniques, such as linear regression, to analyze the relationship between passenger demand and influencing factors, including economic conditions, population trends, and seasonal variations (Chuwang & Chen, 2022; Nar & Arslankaya, 2022; Yang, Xue, Ding, Wu, & Gao, 2021). However, advancements in data collection and analytical methods have led to the integration of sophisticated approaches that combine traditional time series models with contemporary machine learning algorithms (di Torrepadula, Napolitano, Di Martino, & Mazzocca, 2024; Wan, Cheng, & Yang, 2024).

Traditional forecasting models, including linear regression and exponential smoothing, present distinct advantages for predicting rail demand. Linear regression is instrumental in elucidating the relationship between passenger numbers and explanatory variables, such as population density, ticket prices, and regional economic indicators. In contrast, exponential smoothing effectively manages time series data, prioritizing recent observations to capture abrupt changes or trends in passenger flow (Profillidis & Botzoris, 2018).

Research on urban rail systems supports the efficacy of these classical methods. For instance, early applications of linear regression provided a strong base for forecasting accuracy, subsequently enhanced by integrating diverse data sources and refining models. Simultaneously, exponential smoothing has proven effective in addressing seasonal variability and cyclical peaks in passenger demand, facilitating smoother transitions between peak and off-peak periods (Banerjee, Morton, & Akartunalı, 2020; Doustmohammadi, Sisiopiku, Anderson, Doustmohammadi, & Sullivan, 2016).

2.1. Linear Regression

Linear regression is used to quantify the relationship between passenger demand (dependent variable) and a set of independent variables that influence travel behavior. The general form of the regression model can be represented by:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n + \epsilon$$

where:
Y is the forecasted passenger demand,
$\beta 0$ is the intercept,
$\beta 1, \beta 2, \ldots, \beta n$ are the coefficients for the independent variables X1,X2,…, Xn
ϵ is the error term.

The coefficients are determined using the least squares method, which min-imizes the sum of squared deviations between the observed and predicted values. This approach provides interpretable insights into how each predictor variable impacts overall demand and serves as a basis for validating the fore-casting outcomes (Dougherty, 2011).

2.2. Exponential Smoothing

Exponential smoothing techniques are employed to forecast time-dependent data by assigning exponentially decreasing weights to older obser-vations. The basic formula for simple exponential smoothing is given by:

$$F_{t+1} = \alpha \cdot Y_t + (1-\alpha) \cdot F_t$$

where:
F(t+1) is the forecast for the next time period,
Yt is the actual observation at time t
Ft is the forecast for the current time period, and
α is the smoothing constant, with 0<α<1

This method is ideal for environments with relatively stable trends and is further extended to account for trends and seasonality through Holt's linear trend method and Holt-Winters seasonal methods when necessary. This meth-od is very useful when data maintains a certain pattern over time (Cyril, Mulangi, & George, 2018; Hansun, 2016). The allocation of weights enhances responsiveness to data changes, particularly evident in environments with clear trends (Pang, 2023).

In this study, we will employ both single and double exponential smoothing methods, measuring their respective effectiveness against Common forecast accuracy metrics such as Mean Absolute Deviation (MAD), Mean Squared Error (MSE), and Mean Absolute Percentage Error (MAPE). These metrics serve as fundamental benchmarks for assessing the accuracy and reliability of the linear regression and exponential smoothing models.

3. Methods

This study aims to forecast the number of passengers on the Soekarno-Hatta Airport Train using linear regression and exponential smoothing meth-ods. A comprehensive review of previous studies and relevant literature is conducted to establish the theoretical foundation and explore prior applica-tions of these methods in forecasting transportation demand. The dataset con-sists of time series records with monthly frequency. Historical passenger count data will be obtained from official sources provided by the Central Bureau of Statistics (BPS), covering the period from January 2024 to April 2025.

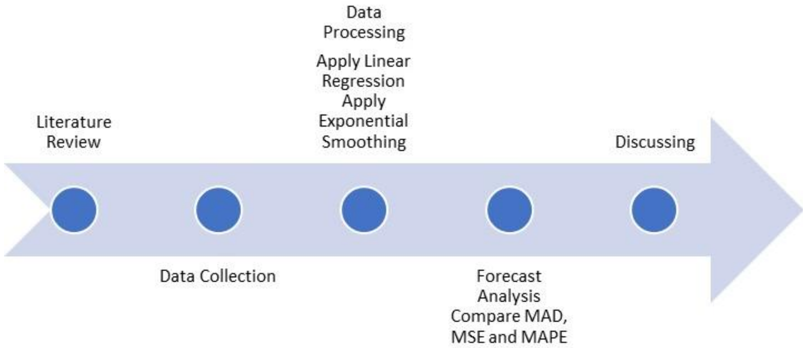

Figure 2. Research Methodology

In this study, we will employ both single and double exponential smoothing methods, measuring their respective effectiveness against forecast accuracy metrics such as Mean Absolute Deviation (MAD), Mean Squared Error (MSE), and Mean Absolute Percentage Error (MAPE). These metrics serve as fundamental benchmarks for assessing the accuracy and reliability of the linear regression and exponential smoothing models. In the final stage, the forecasting results are interpreted. The strengths and limitations of each meth-od are discussed, particularly in terms of their responsiveness to trends and fluctuations in the data.

Based on the findings, recommendations are proposed to guide future forecasting and planning efforts.

4. Results and Discussion

First, the historical data will be presented. The total number of passengers using the Soekarno-Hatta Airport Rail from January 2024 to April 2025 is shown in Table 1. This data was obtained from the Indonesian Central Statistics Agency (Badan Pusat Statistik) and is illustrated in a graph as depicted in Figure 3. Based on the existing graph, it is evident that the number of passengers shows an increasing trend. The data reflects increasing passenger use of the airport rail, demonstrating its growing importance as a transportation mode for airport access. This increasing trend could be attributed to factors such as improved public awareness, expansion of marketing efforts, and ongoing improvements in service reliability and efficiency. Therefore, the use of linear regression and exponential smoothing as forecasting methods for airport train passengers is highly appropriate.

Table 1. Number of Passengers Soekarno-Hatta Airport Train (BPS, 2025)

Period		Number of Passengers
Month	**Year**	
January	2024	618000
February		603000
March		578000
April		704000
May		650000
June		655000
July		711000
Agust		660000
September		698000
October		675000
November		651000
December		775000
January	2025	765000
February		709000
March		662000
April		838000

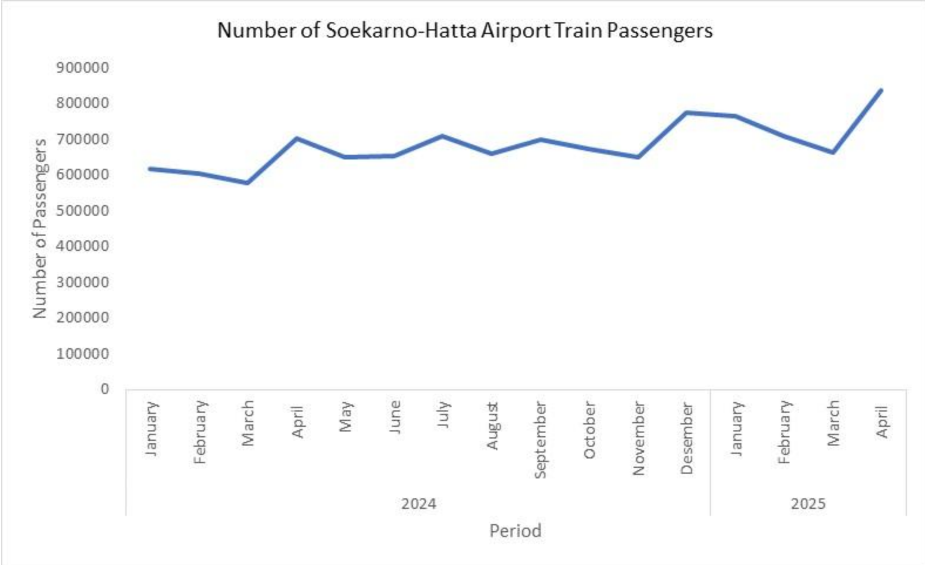

Figure 3. Graph of Passengers Soekarno-Hatta Airport Train

In this paper linear regression, single exponential smoothing and double exponential smoothing methods will be used to forecast the number of airport train passengers. Using the linear regression formula outlined in the literature review; we calculated a slope of 10,061.76 and an intercept of 598,975. The results are graphically represented in Figure 4, while the detailed calculation errors for the linear regression method can be found in Table 2.

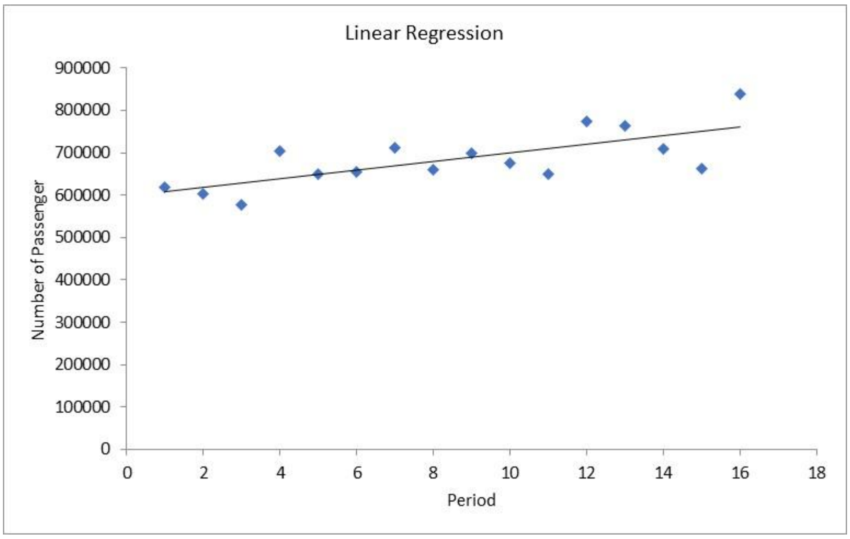

Figure 4. Graph of Linear Regression Method

Figure 4 depicts the linear regression results, showcasing the predicted passenger numbers against the actual observed values. The slope indicates a strong positive relationship between time and passenger numbers, confirming the upward trend observed in Figure 3. The close proximity of the predicted values to the actual figures demonstrates the effectiveness of linear regression in capturing the trend of increasing demand. Table 2 presents the error calculations associated with the linear regression method. The average absolute percentage error of 5.27% indicates a reasonable degree of accuracy.

Table 2. Linear Regression Error Calculation

Period (x)	Number of Passenger (y)	Absolute Deviation	Squared Error	Absolute Percentage Error
1	618000	8963.24	80339586.94	0.015
2	603000	16098.53	259162649.22	0.027
3	578000	51160.29	2617375694.20	0.089
4	704000	64777.94	4196181663.06	0.092
5	650000	716.18	512908.74	0.001
6	655000	4345.59	18884137.11	0.007
7	711000	41592.65	1729948289.36	0.058
8	660000	19469.12	379046541.96	0.029
9	698000	8469.12	71725953.72	0.012
10	675000	24592.65	604798289.36	0.036
11	651000	58654.41	3440340019.46	0.090
12	775000	55283.82	3056301144.03	0.071
13	765000	35222.06	1240593427.77	0.046
14	709000	30839.71	951087458.91	0.043
15	662000	87901.47	7726668531.57	0.133
16	838000	78036.76	6089736645.76	0.093
	TOTAL	586123.52	32462702941	84.29%
	AVERAGE	**36632.72**	**2028918934**	**05.27%**

For the Single Exponential Smoothing method, an α of 0.3 was used. The graph of this single exponential smoothing is shown in Figure 5, and the error calculations can be found in Table 3.

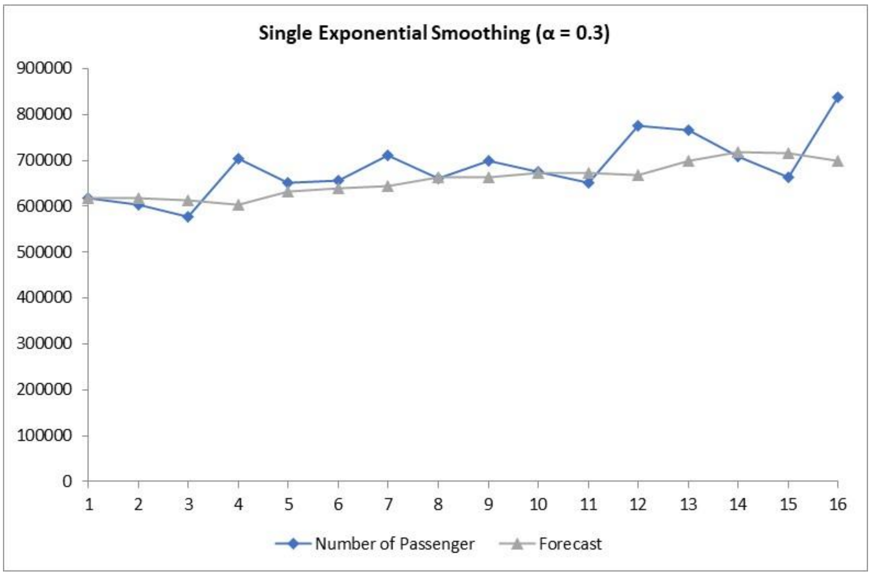

Figure 5. Graph of Single Exponential Smoothing Method

Table 3. Single Exponential Smoothing Error Calculation

Period (x)	Number of Passenger (y)	Absolute Deviation	Squared Error	Absolute Percentage Error
1	618000	0	0	0%
2	603000	15000	225000000	02.49%
3	578000	35500	1260250000	06.14%
4	704000	101150	10231322500	14.37%
5	650000	16805	282408025	02.59%
6	655000	16763.5	281014932.3	02.56%
7	711000	67734.45	4587955717	09.53%
8	660000	3585.885	12858571.23	00.54%
9	698000	35489.881	1259531618	05.08%
10	675000	1842.9163	3396340.673	00.27%
11	651000	22709.959	515742217.6	03.49%
12	775000	108103.03	11686264881	13.95%
13	765000	65672.12	4312827386	08.58%
14	709000	10029.516	100591186.9	01.41%
15	662000	54020.661	2918231820	08.16%
16	838000	138185.54	19095242709	16.49%
	TOTAL	692592.45	56772637905	95.66%
	AVERAGE	**43287.028**	**3548289869**	**05.98%**

Figure 5 demonstrates the Single Exponential Smoothing results, which reflect a more responsive forecasting approach compared to linear regression. The application of an α value of 0.3 allows for a reasonably reactive model, effectively capturing recent trends in passenger demand. The smoother transitions in the graphical representation highlight how this method can track fluctuations better than simple linear models. Table 3 provides a detailed summary of the error calculations for the Single Exponential Smoothing method. The total error, totaling 95.66%, reflects the cumulative deviations from the actual passenger numbers. The average absolute percentage error of 5.98% suggests reasonable prediction accuracy.

Lastly, for the Double Exponential Smoothing method, an α = 0.3 and β = 0.1 were used. The graph of this double exponential smoothing is shown in Figure 6, and the error calculations can be found in Table 4.

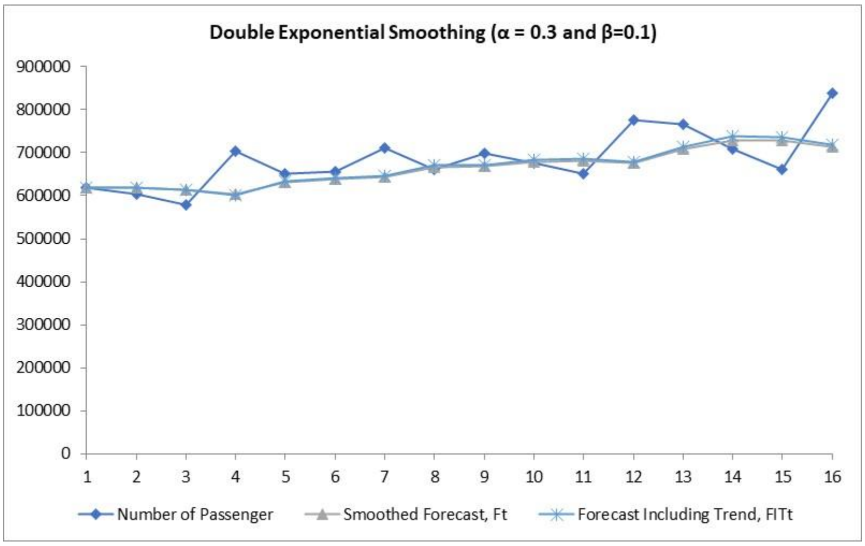

Figure 6. Graph of Double Exponential Smoothing Method

Table 4. Double Exponential Smoothing Error Calculation

Period (x)	Number of Passenger (y)	Absolute Deviation	Squared Error	Absolute Percentage Error
1	618000	0	0	00.00%
2	603000	15000	225000000	02.49%
3	578000	35050	1228502500	06.06%
4	704000	102966.5	10602100122	14.63%
5	650000	16489.055	271888934.8	02.54%
6	655000	14460.17185	209096569.9	02.21%
7	711000	63606.14849	4045742126	08.95%
8	660000	10899.85232	118806780.5	01.65%
9	698000	26272.94269	690267517.4	03.76%
10	675000	9494.28909	90141525.33	01.41%
11	651000	35246.52266	1242317360	05.41%
12	775000	95784.30952	9174633950	12.36%
13	765000	50632.36276	2563636158	06.62%
14	709000	28492.97086	811849388.2	04.02%
15	662000	74025.91526	5479836130	11.18%
16	838000	119321.8011	14237692221	0.14239
	TOTAL	697742.8416	50991511284	97.52%
	AVERAGE	**43608.9276**	**3186969455**	**06.10%**

The results indicate that all three forecasting methods yield substantial insights into passenger demand trends on the Soekarno-Hatta Airport Rail. While linear regression provides a straightforward model for understanding the influencing variables on demand, the exponential smoothing methods offer valuable perspectives on temporal trends and fluctuations. The corresponding errors calculated across the methods demonstrate varying degrees of accuracy and reliability, vital for guiding future operational strategies. Table 5 presents a comparative analysis of forecasting accuracy across three methods: linear regression, single exponential smoothing, and double exponential smoothing. Each method is evaluated using three key error metrics: Mean Absolute Deviation (MAD), Mean Squared Error (MSE), and Mean Absolute Percentage Error (MAPE).

Table 5. Comparison MAD, MSE and MAPE

Methods	MAD	MSE	MAPE
Linear Regression	36632.72	2028918934	5.27%
Single Exponential Smoothing	43287.03	3548289869	5.98%
Double Exponential Smoothing	43608.93	3186969455	6.10%

Based on table 5 it could be seen that linear regression method having the lowest error calculation compare to other methods. Therefore making it the preferred choice for predicting passenger demand for the airport rail service. The insights gleaned from this table emphasize the need for thorough evaluation of forecasting techniques to ensure optimal planning and resource management. Table 6 shown the forecaste number in next period.

Table 6. Forecast Number of Passengers Soekarno-Hatta Airport Train

Period	Number of Passengers
17	770025
18	780087
19	790149
20	800210
21	810272
22	820334
23	830396
24	840457

Number of passengers for the Soekarno-Hatta Airport Train for the upcoming periods, from period 17 to period 24 indicate a consistent and gradual increase in passenger numbers, starting from 770,025 in period 17 and reaching 840,457 by period 24.

Conclusions

This study underscores the importance of accurate passenger forecasting for the Soekarno-Hatta Airport Rail Link, particularly given the significant implications for operational efficiency and service quality. By utilizing linear regression and exponential smoothing, reliable models can be developed that cater to the dynamic nature of urban transportation demands.

Further research could explore the integration of machine learning tech-niques to enhance forecasting accuracy, potentially incorporating real-time data feeds to facilitate adaptive operational strategies. This would ensure that the increasing complexities of urban travel patterns are effectively addressed to meet future challenges in public transportation.

Conflicts of Interest

The authors declare no conflict of interest.

References

Abdallah, T. (2023). Sustainable mass transit: Challenges and opportunities in urban public transportation: Elsevier.

Banerjee, N., Morton, A., & Akartunalı, K. (2020). Passenger demand forecasting in scheduled transportation. European Journal of Operational Research, 286(3), 797-810. https://doi.org/10.1016/j.ejor.2019.10.032

BPS. (2025). Jumlah Penumpang Kereta Api. Retrieved from https://www.bps.go.id/id/statistics-table/2/NzIjMg==/jumlah-penumpang-kereta-api.html

Chen, T. D., Kockelman, K. M., & Zhao, Y. (2015). What matters most in transportation demand model specifications: a comparison of outputs in a mid-size network. Paper presented at the Journal of the Transportation Research Forum.

Chuwang, D. D., & Chen, W. (2022). Forecasting daily and weekly passenger demand for urban rail transit stations based on a time series model approach. Forecasting, 4(4), 904-924. https://doi.org/10.3390/forecast4040049

Cyril, A., Mulangi, R. H., & George, V. (2018). Modelling and forecasting bus passenger demand using time series method. Paper presented at the 2018 7th international conference on reliability, Infocom Technologies and Optimization (Trends and Future Directions)(ICRITO). https://doi.org/10.1109/ICRITO.2018.8748443

di Torrepadula, F. R., 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

Dougherty, C. (2011). Introduction to econometrics: Oxford University Press, USA.

Doustmohammadi, E., Sisiopiku, V., Anderson, M. D., Doustmohammadi, M., & Sullivan, A. (2016). Comparison of freight demand forecasting models. Journal of Traffic Transportation Engineering, 5(1), 19-26.

Hansun, S. (2016). A new approach of brown's double exponential smoothing method in time series analysis. Balkan Journal of Electrical Computer Engineering, 4(2), 75-78.

Jiao, P., Li, R., Sun, T., Hou, Z., & Ibrahim, A. (2016). Three revised Kalman filtering models for short-term rail transit passenger flow prediction. Mathematical Problems in Engineering, 2016(1), 9717582. https://doi.org/10.1155/2016/9717582

Nar, M., & Arslankaya, S. (2022). Passenger demand forecasting for railway systems. Open Chemistry, 20(1), 105-119.

Pang, Y. (2023). Research on market share prediction of highway passenger transport based on exponential smoothing method. J. Traffic Transp. Eng, 8(2).

Profillidis, V. A., & Botzoris, G. N. (2018). Modeling of transport demand: Analyzing, calculating, and forecasting transport demand: Elsevier. https://doi.org/10.1080/01441647.2019.1635226

Shastri, S., Sharma, A., Mansotra, V., Sharma, A., Bhadwal, A., & Kumari, M. (2018). A study on exponential smoothing method for forecasting. International Journal of Computer Sciences Engineering, 6(4), 482-485.

Tang, R., De Donato, L., Besinović, N., Flammini, F., Goverde, R. M., Lin, Z., . . . Wang, Z. (2022). A literature review of Artificial Intelligence applications in railway systems. Transportation Research Part C: Emerging Technologies, 140, 103679. https://doi.org/10.1109/TITS.2021.3131637

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(1), 5259452. https://doi.org/10.1049/2024/5259452

Yang, X., Xue, Q., Ding, M., Wu, J., & Gao, Z. (2021). Short-term prediction of passenger volume for urban rail systems: A deep learning approach based on smart-card data. International Journal of Production Economics, 231, 107920. https://doi.org/10.1016/j.ijpe.2020.107920

Zhu, L., Chen, C., Wang, H., Yu, F. R., & Tang, T. (2023). Machine learning in urban rail transit systems: A survey. IEEE Transactions on Intelligent Transportation Systems, 25(3), 2182-2207. https://doi.org/10.1109/TITS.2023.3319135