J. Highw. Transp. Res. Dev. 2025, 19(4): 55−61

Journal of Highway and Transportation Research and Development

Home page: www.sciopen.com/journal/2095-6215

Quantifying volatility characteristics of passenger flow in metro stations based on rolling-window

Zhao Liu[1*], Dan Ning[2], Wenming Rao[1], Meng Jia[1]

1. School of Traffic Engineering, Nanjing Institute of Technology, Nanjing, Jiangsu 211167, China

2. Department of Electronic Engineering, Nanjing Engineering Vocational College of Jiangsu Province, Nanjing, Jiangsu 210000, China

Abstract: Understanding the characteristics of passenger flows within a metro network is essential to enhance the operational efficiency and safety of urban rail transit system. Most research has focused on investigating trends or patterns of passenger flows, ignoring the volatility inherent in these trends or patterns. This study employs a rolling-window analysis approach, coupled with two nonparametric methods and on statistical method to quantify the volatility in passenger flow patterns across metro stations. The volatility characteristics of passenger flow are captured by the ranges of potential changes in passenger flows, which are measured using three different indices: kickoff, width, and width-to-flow ratio. By employing actual passenger data from six metro stations in Suzhou, China, this study quantifies the volatility characteristics of passenger flows on weekdays and weekends. It reveals that passenger flow volatility on weekends is greater than on weekdays, indicating more erratic passenger flow changes on weekends. Our findings have practical implications for enhancing management's understanding of passenger flow volatility and the provision of reliable metro information to travelers via information systems. Future work is recommended to develop and test more methods for quantifying the volatility of passenger flow in the metro stations, metro lines, and the metro network.

Keywords: transportation engineering; urban rail transit system; passenger flow; volatility characteristics; rolling-window analysis

1 Introduction

The urban rail transit system is becoming increasingly crucial for enhancing residents' travel efficiency and alleviating road traffic congestion in major cities worldwide. Evidence indicates that expanding metro networks have been found to stimulate a sharp rise in ridership [1]. This trend is primarily due to the high speed, safety, and environmental friendliness of the urban rail transit system, which make it a preferred daily travel option for many city dwellers [2, 3]. The rapid increase in passenger flow can create immense pressure on the management of metro stations, resulting in potential safety risks [4]. Therefore, understanding passenger flows is essential for making and coordinating operation plans in a metro system, and the allocation of passenger flows on the metro network plays a pivotal role in this analysis.

Passenger flow data is typically obtained through the use of automatic fare collection (AFC) systems, which are extensively employed in urban rail transit systems to gather information regarding passenger trips, including inbound and outbound passenger flow [5, 6], origin-destination matrices [7], and travel time [8]. Urban metro passenger flow varies across different times and locations, encompassing working days, holidays, residential areas, business centers, climate, transportation modes connected to the metro, and future urban planning. Adhering to some changing trends or patterns, short-term passenger flow generally exhibits a degree of uncertainty or volatility. Very limited research focuses on the volatility of short-term passenger flow. Guo et al. [9] applied the fuzzy information granulation method to obtain the dispersion range of the collected traffic flow time series, and the volatility of traffic flow could be reflected by the traffic flow interval time series. Li et al. [10] employed quantile regression method to generate a prediction interval for strengthening the reliability of travel time prediction through investigating the volatility of travel time. Tsekeris and Stathopoulos [11] investigated the traffic (speed) variability through the conditional variance of traffic flow level, which was described as a latent stochastic (low-order Markov) process. Some studies have directed their attention towards the volatility and uncertainty of short-term traffic flow, which exhibits remarkable similarities with short-term passenger flow. It is worth noting that the previously mentioned methods (i.e., fuzzy information granulation, quantile regression, and conditional variance) for quantifying traffic flow uncertainty and volatility exhibit significant differences. Quantile regression and conditional variance are classical statistic methods. While, fuzzy information granulation is a principle used in fuzzy logic and data analysis to transform raw data into meaningful and manageable granules. Regardless of their computational principles, the indicators of traffic flow volatility are significantly different. However, to the best of the authors' knowledge, little research has been dedicated to examining the volatility characteristics of short-term passenger flow.

One of the primary objectives in the field of passenger flow analysis is the accurate prediction of the short-term passenger flows.

Received: 26 March 2025; **Revised:** 17 June 2025; **Accepted:** 8 October 2025
* *E-mail address*: liuzhao025@njit.edu.cn
DOI: 10.26599/HTRD.2025.9480077

The majority of studies on short-term passenger flow prediction predominantly concentrate on the recurring daily passenger flow patterns [12, 13].When confronted with non-recurrent events such as sporting games, concerts, and extreme weather conditions, accurately predicting the short-term passenger flow would become a more changing task due to their irregular and inconsistent nature. Chen, et al. [2] explored the nonlinear and asymmetric features of passenger flow during special events, taking into account the volatility (heteroscedasticity) of short-term passenger flow. Zheng, et al. [14] proposed a sophisticated network index designed to detect and analyze anomalous and collective patterns of passenger flow across the entire urban metro network, adapting to the unexpected outflow patterns that had not been observed previously. Ni, et al. [15] discovered a positive correlation between passenger flow and social media posts and developed a hashtag-based event detection algorithm to detect high subway passenger flow based on social media data. Rodrigues, et al. [16] presents a Bayesian additive model with Gaussian process components that combines smart card records from public transport with context information about events. Although there have been some studies focusing on the volatility of passenger flow under special events or scenarios, little research investigates the volatility of regular passenger flow or day-to-day recurrent passenger flow. Without the influence of special events, day-to-day recurrent passenger flow also exhibits regular volatility or fluctuation. Therefore, investigating the volatility characteristics of passenger flow in the metro stations or the metro network can provide valuable insight for the subway company to improve operational efficiency and reduce potential security risks.

This research aims to quantify passenger flow volatility by measuring the potential ranges of changes. We define the concept of passenger flow volatility based on real passenger flow data and propose a rolling-window analysis utilizing different methods: fuzzy information granulation and quantile regression. We also introduce three measures, i.e., kickoff, width, and width-to-flow ratio, to quantitatively reflect passenger flow volatility characteristics.

2 Problem description

To depict the volatile nature of passenger flow in metro stations, this study employs six days of real passenger flow data gathered from Leqiao station, which is situated within the Suzhou metro network in China. The specifics of the utilized passenger flow data will be elucidated in Section 5. Fig. 1 (a) and (c) exhibit the inbound and outbound passenger flow on weekdays from Mar. 14, 2016 to Mar. 16, 2016, while Fig. 1 (b) and (d) illustrate the inbound and outbound passenger flow on weekends from Mar. 19, Mar. 20, and Mar. 26, 2016. It is widely acknowledged that passenger flow typically displays a dual-peak pattern on weekdays and a singular-peak pattern on weekends.

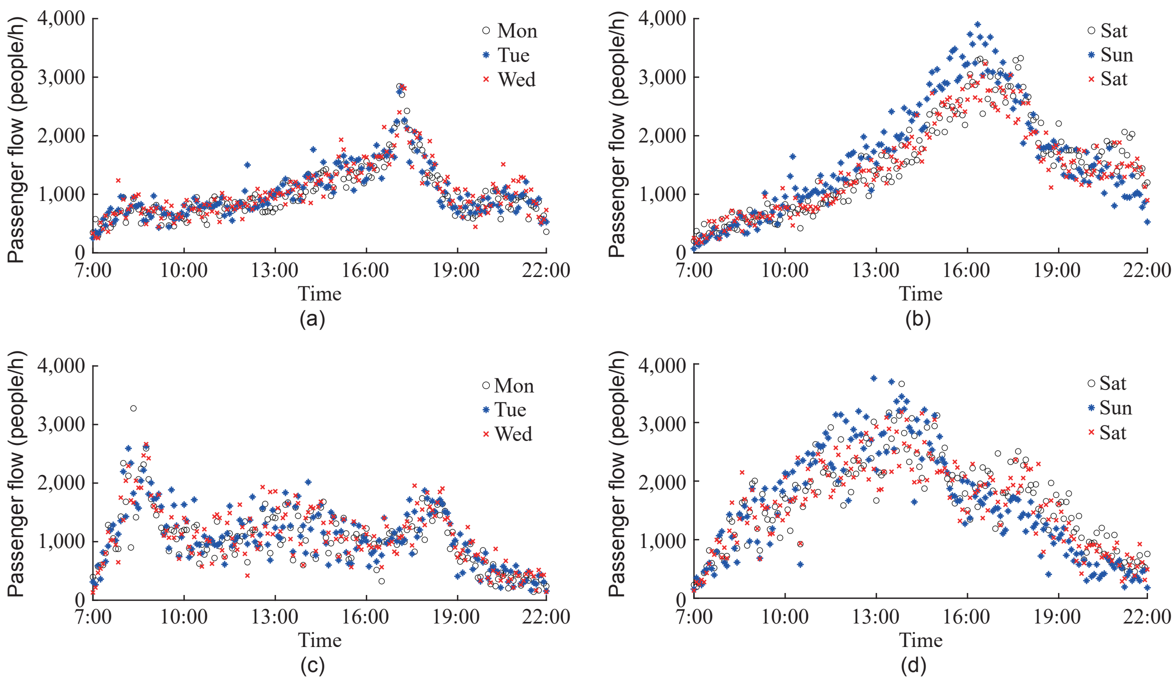

Fig. 1 Scatters of short-term passenger flow in the Leqiao station. (a) Weekdays from Mar. 14, 2016 to Mar. 16, 2016; (b) weekends from Mar. 19, Mar. 20, and Mar. 26, 2016; (c) weekdays from Mar. 14, 2016 to Mar. 16, 2016; (d) weekends from Mar. 19, Mar. 20, and Mar. 26, 2016.

The scatter plots of passenger flow over three consecutive days offer a valuable means of intuitively comprehending the volatility characteristics of passenger flow. Notably, the dispersion of passenger flow differs significantly between inbound and outbound passenger flow on weekdays and weekends. This indicates that the range of passenger flow over three consecutive days at the same time may signify the volatility of passenger flow. For instance, in Fig. 1 (a) and (b), the inbound passenger flow on weekdays between 13:00 and 16:00 exhibits narrower ranges than those on weekends. Conversely, in Fig. 1 (c) and (d), the outbound passenger flow on weekdays and weekends display comparable ranges. Generally, larger changes in passenger flow correspond to greater volatility. Therefore, this study employs dispersion estimation of passenger flow over time as a means of quantifying volatility.

3 Methodology

This section outlines a framework utilizing a rolling-window analysis that combines two nonparametric methods with a conventional statistical method to compute the ranges of potential changes in passenger flow. Three indices are introduced to represent the volatility attributes of passenger flow.

3.1 Overall framework

The rolling-window analysis is employed to construct the overall framework for measuring passenger flow volatility, as illustrated in Fig. 2. Above the dashed line in Fig. 2 lies the discrete time series of passenger flow for a single day. The dynamic process of the rolling-window analysis is depicted below the dashed line. The rolling window size, denoting the consecutive weekdays or weekends per window, is a positive integer greater than one, as exemplified by the three-day window size in Fig. 2. Each rolling window generates a volatility value of passenger flow at a specific time period. Various methods may be employed within this framework to calculate specific volatility values of passenger flow as the rolling window advances.

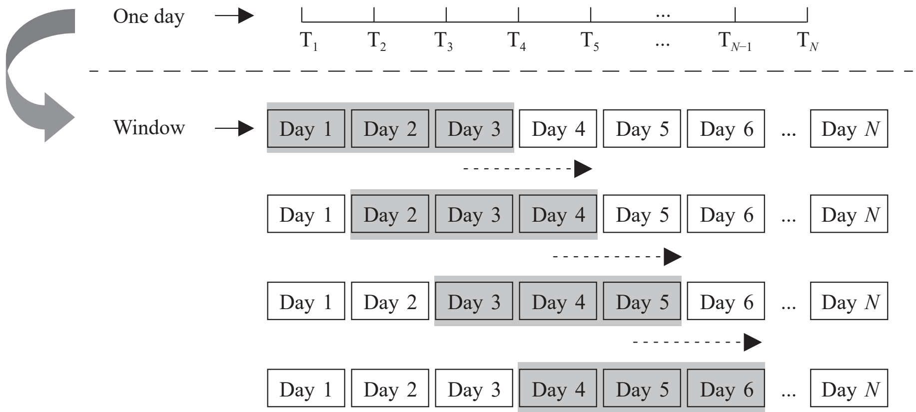

Fig. 2 Rolling-window implementation for quantifying the volatility of passenger flow.

Consequently, the consecutive weekdays or weekends within each rolling window represent a subsample capable of capturing the volatility attributes of passenger flow in metro stations. Notably, Fig. 1 corresponds to Fig. 2 and presents scatter plots of passenger flow over three consecutive days on weekdays and weekends, exemplifying the volatility characteristics of passenger flow.

3.2 Fuzzy information granulation

Fuzzy information granulation has a significant advantage for analyzing passenger flow characteristics compared with the classical statistic methods [17]. We adopt fuzzy information granulation to extract the ranges of possible changes in passenger flow data. Fuzzy information granulation, which is not limited to fuzzy logic, but can also be applied to any concept, method, or theory. Supposing x_t^i is the passenger flow at time t in the i-th metro station. $x_t^i \in U$ and let G^i denote a fuzzy subset of U for the i-th station. The fuzzy granule g can be defined as:

$$g = (x_t^i \text{ is } G^i) \text{ is } \lambda \tag{1}$$

where, G^i is described by its membership function; and λ is fuzzy probability of passenger flow by possibility distribution on the unit interval.

To achieve fuzzy granulation of passenger flow in consecutive days, the core task is to construct a set of fuzzy granules G^i that represent typical passenger flow patterns. Each granule G^i could be defined by a membership function μ_G, which quantifies the degree to which any observed passenger flow. In this study, triangular membership function is selected for its simplicity and interpretability. The parameters of these triangular functions (i.e., the lower bound, peak, and upper bound) are determined from the historical passenger flow data through the fuzzy granulation process, as detailed in Ref. [2, 18]. Triangular fuzzy sets are commonly used as the membership function for granulating passenger flow and consist of three segments:

$$A(x_t,a,m,b) = \begin{cases} 0, \; x_t < a \\ \dfrac{x_t - a}{m - a}, \; a \leqslant x_t \leqslant m \\ \dfrac{b - x_t}{b - m}, \; m < x_t < b \\ 0, \; x_t > b \end{cases} \tag{2}$$

Where, x_t is the specific value of passenger flow at the time t in the consecutive weekdays or weekends; a, b, and m are the minimum value, average value, and the maximum value of the original passenger flow for a single fuzzy particle.

To establish a reliable fuzzy granulation, it is necessary to follow the principles of flexibility and stability. The former ensures sufficient information content, while the latter preserves identity despite small fluctuations in the data. For a given time series, fuzzy set A can be constructed through a dual-objective optimization.

$$\max \sum_{j=1}^{N} A(x_j) \tag{3}$$

$$\min \; |\sup (A)| = b - a \tag{4}$$

where, μ_A denotes the membership degree of data point x_j in fuzzy set A; a and b represent the lower and upper bounds of the support of A, respectively.

By the combined Eq. (3) with Eq. (4), the index Q is introduced as

$$Q_A = \frac{\sum_{j=1}^{N} A(x_j)}{\text{measure}(\text{supp}(A))} = \frac{\sum_{j=1}^{N} A(x_j)}{b - a} \tag{5}$$

Maximize the index Q_A of a and b to determine the parameters of the triangular fuzzy set that satisfy the two principles. This leads to the division of aggregated passenger flow data into up-bound, average bound, and low-bound time series. The up-bound and low-bound time series are selected to reflect the volatility in passenger flow.

$$\bar{v} = \frac{1}{m}\sum_{t=1}^{m} v_t = \frac{1}{m}\sum_{t=1}^{m} (b_t - a_t) \tag{6}$$

where, $\bar{v}$ is the average volatility in the passenger flow; v_t is the specific volatility of passenger flow at the time t; b_t and a_t are the upper-bound and lower-bound value of passenger flow at the time t, respectively; m is the number of discrete time intervals.

3.3 Quantile regression

We attempt to quantify the volatility of passenger flow data with quantile regression. Suppose that the passenger flow data in a metro station for n consecutive days at time t is $\boldsymbol{X}_t = [x_t^1, x_t^2, \cdots, x_t^n]$, arranged in ascending order, and time series data X_t can be transformed into a sequence after rearranging $\boldsymbol{Y}=[y_1, y_2, \cdots, y_n]$, where $y_1 \leqslant y_2 \leqslant \cdots \leqslant y_{n-1} \leqslant y_n$. Then, the positions of the lower and upper quartile in the sequence Y are:

$$\begin{cases} q_1 = \dfrac{n+1}{4} \\ q_3 = \dfrac{(n+1)\times 3}{4} \end{cases} \tag{7}$$

where, q_1 and q_3 are the positions of the lower and upper quartile of the $\boldsymbol{Y}$, respectively; n is the number of elements in the $\boldsymbol{Y}$.

Consequently, the lower quartile and upper quartile of the $\boldsymbol{Y}$ can be calculated as below:

$$\begin{cases} Q_1 = x_{q_{Z1}} + (x_{q_{Z1+1}} - x_{q_{Z1}}) \times q_{X1} \\ Q_3 = x_{q_{Z3}} + (x_{q_{Z3+1}} - x_{q_{Z3}}) \times q_{X3} \end{cases} \tag{8}$$

where, Q_1 and Q_3 are the lower quartile and upper quartile of the $\boldsymbol{Y}$; q_{Z1} and q_{X1} are the integer part and the decimal part of q_1, respectively; q_{Z3} and q_{X3} are the integer part and the decimal part of q_3.

Then, the interquartile range (IQR) of passenger flow can be expressed as:

$$Q_t^r = Q_t^3 - Q_t^1 \tag{9}$$

where, Q_t^r is the interquartile range of passenger flow at the time t; Q_t^1 and Q_t^3 are the lower quartile and upper quartile of the passenger flow at the time t.

Therefore, the quartile-based detection of passenger flow can be expressed as:

$$[X_t^L, X_t^R] = [Q_1 - \alpha IQR_t, Q_3 + \alpha IQR_t] \tag{10}$$

where, X_t^L and X_t^R are the lower and upper limits of the passenger flow at time t; α is the coefficient of the interquartile range.

Based on the Eq. (9), the detection intervals of passenger flow at different times on the same day can be calculated as below:

$$X = [(X_1^L, X_1^R), (X_2^L, X_2^R), \cdots, (X_t^L, X_t^R)] \tag{11}$$

Based on the detected intervals of passenger flow at different times, the volatility in the passenger flow can be obtained:

$$\bar{v} = \frac{1}{m}\sum_{t=1}^{m} v_t = \frac{1}{m}\sum_{t=1}^{m}(X_t^R - X_t^L) \tag{12}$$

where, $\bar{v}$ is the average volatility in the passenger flow; v_t is the specific volatility of passenger flow at the time t; X_t^L and X_t^R are the upper-bound and lower-bound value of passenger flow at the time t; m is the number of discrete time intervals.

3.4 Statistical method

Similar with the quantile regression, suppose that the passenger flow data in a metro station for n consecutive days at time t is $\boldsymbol{X}_t = [x_t^1, x_t^2, \cdots, x_t^n]$. The pattern of the passenger flow can be presented as $\boldsymbol{X}$=[X_1, X_2, X_3,$\cdots$, X_m]. At a specific time t, the variance of the passenger flow for n consecutive days could reflect the dispersion of passenger flow, which can be presented as:

$$\sigma_t = \sqrt{\sum_{i=1}^{n} \frac{(x_t^i - \bar{x}_t)}{n-1}} \tag{13}$$

$$\delta_t = t(n-1, \alpha)\sigma_t \tag{14}$$

where, σ_t is the variance of the passenger flow $\boldsymbol{X}_t$; δ_t is the (1−α) confidence interval at the time t; x_t^i is the passenger flow at the time t on the i-th day; $\bar{x}_t$ is the average value of the passenger flow $\boldsymbol{X}_t$; n is the days of consecutive weekdays or weekends; $t(h,a)$ denotes the a-th quantile of the Student's t-distribution on h degrees of freedom.

Therefore, the volatility in the passenger flow can be quantified as below:

$$\bar{v} = \frac{1}{m}\sum_{t=1}^{m} v_t = \frac{1}{m}\sum_{t=1}^{m} 2\eta\delta_t \tag{15}$$

where, $\bar{v}$ is the average volatility in the passenger flow; v_t is the specific volatility of passenger flow at the time t; δ_t is the (1−α) confidence interval at the time t; η is an adjusted coefficient.

3.5 Performance index

To the best of our knowledge, there is no consensus on a standardized method to evaluate the volatility of passenger flow in metro stations. However, as proposed in Ref. [2, 18], the evaluation of traffic flow volatility can be based on two principles: firstly, real observations should fall within the estimated intervals, which are confidence-level-based intervals; secondly, the estimated intervals should be as narrow as possible to indicate higher informativeness. In this study, we propose three measures, namely K, W, and R, to assess the volatility of passenger flow.

$$K = \frac{\sum_{t=1}^{N} w_t}{N}, \quad w_t = \begin{cases} 0, & x_t^{\text{low}} \leqslant x_t^{\text{ob}} \leqslant x_t^{\text{up}} \\ 1, & x_t^{\text{up}} < x_t^{\text{ob}} \ or \ x_t^{\text{ob}} < x_t^{\text{low}} \end{cases} \tag{16}$$

$$W = \frac{1}{N}\sum_{t=1}^{N} |x_t^{\text{up}} - x_t^{\text{low}}| \tag{17}$$

$$R = \frac{1}{N}\sum_{t=1}^{N} \frac{|x_t^{\text{up}} - x_t^{\text{low}}|}{\bar{x}_t} \tag{18}$$

where, K denotes the exclusion rate, representing the proportion of data points falling outside the constructed interval; W represents the width of the constructed interval; R denotes the ratio of interval width to passenger flow volume, reflecting the relative variability scaled by flow level; x_t^{low} and x_t^{up} are the low and up threshold of passenger flow fluctuation; $\bar{x}_t$ is the average value of consecutive days of N at the time of t; N is the number of days of the selected passenger flow data.

4 Case study

In this section, we conducted an empirical study using real-world data to analyze the volatility characteristics of passenger flow in metro stations.

4.1 Data collection and aggregation

The data used in this study were collected by the AFC system installed in the metro stations of Suzhou in China. The archived information of passenger flow (e.g., entering or leaving the metro station) can be obtained by aggregating passengers' raw transaction data into time series with a 5 min interval. The raw travel information of passengers were collected for 15 h per day and seven days per week from Mar 14, 2016, to May 15, 2016, consisting of 45 working days and 18 weekends. Twelve metro stations were selected from Suzou metro stations, i.e., Mudu station, Xiangmen station, Xinghai Square station, Leqiao station, Nanshi Street station, Guangji South Street station, Suzhoubei Railway station, Shantang Street station, Laodong Street station, Shihu East Road station, Baodaiqiao South station, and Lumu station.

4.2 Impact of the rolling-window size

The selection of an optimal window size plays a crucial role in evaluating passenger flow volatility. A wide range of rolling-window sizes was tested, ranging from 2 to 40 with an increment of 2. For brevity, inbound passenger flow from Leqiao station were used to present the impact of the rolling-window size in terms of the Kickoff and Width, which are shown in Fig. 3.

Fig. 3 (a) and (b) demonstrate the impact of rolling-window size on passenger flow volatility in terms of K and W. As shown in Fig. 3 (a), the K of fuzzy information granulation rapidly drops from 50% to 15% until a rolling-window size of 14, and then gradually

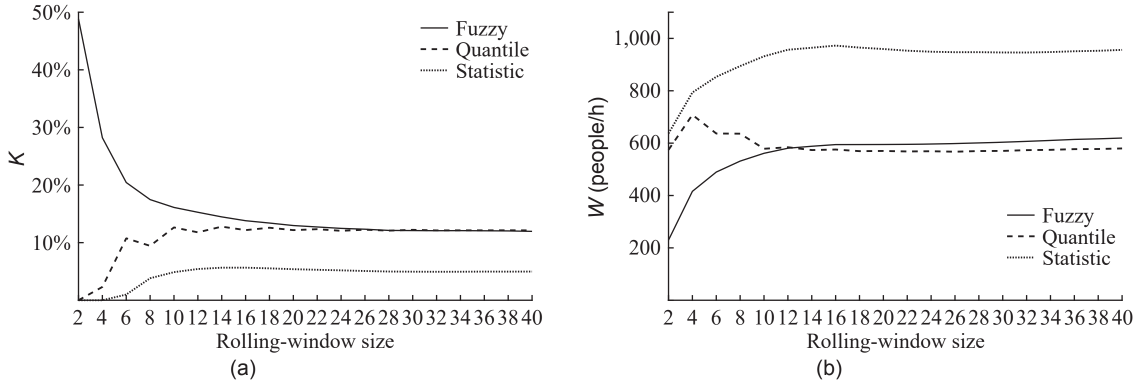

Fig. 3 Impact of rolling-window size on the volatility of the passenger flow. (a) K of passenger flow; (b) W of passenger flow.

declines to about 13% and stabilizes. In contrast, both quantile regression and traditional statistical methods follow a similar pattern where the K rise and then stabilize as the rolling-window size increases. In Fig. 3 (b), both fuzzy information granulation and traditional statistical methods exhibit a consistent trend as their respective W gradually increase as the rolling-window size increases from 2 to 40. Conversely, the W of quantile regression remains stable at around 600 across different window sizes.

Considering the performance of K and W, we set the size of rolling-window as 14 for further investigating the volatility characteristics of passenger flow.

4.3 Comparison of different methods on quantifying the volatility of passenger flow

Once an appropriate rolling-window size has been determined, we will perform a quantitative comparison of the volatility in passenger flow obtained using various methods. This comparison will be based on three metrics, namely the K, W, and R, as illustrated in Fig. 4 (a), (b), and (c), respectively.

Fig. 4 (a) illustrates that the K of fuzzy information granulation, quantile regression, and traditional statistical method remain stable over time, with values of approximately 16%, 12%, and 5%, respectively, indicating a consistent gap between the three methods. Fig. 4 (b) shows that the W of passenger flow exhibit a double-peak pattern over time, with the W of passenger flow from the traditional statistical method being significantly larger than those from the fuzzy information granulation and quantile regression, which have similar W evaluations. In Fig. 4 (c), a small constant gap is observed among the three methods, with the R of fuzzy information granulation being more stable and smaller. Notably, the R of quantile regression fluctuates during morning and evening hours. The tradeoff between the K and W of passenger flow affects the volatility, with higher volatility indicating a wider range of possible changes and a lower K, and vice versa.

4.4 Comparison of weekdays and weekends on the volatility of passenger flow

Comparing the W and R on weekdays and weekends provides more insights into passenger flow volatility. Fig. 5 displays a comparison of the changing trends of passenger flow volatility between weekdays and weekends. Notably, the range of possible changes in passenger flow varies over time, and there is a clear difference between inbound and outbound flows. The R in inbound passenger flow fluctuates between 1 and 3 over time, while the R in out-

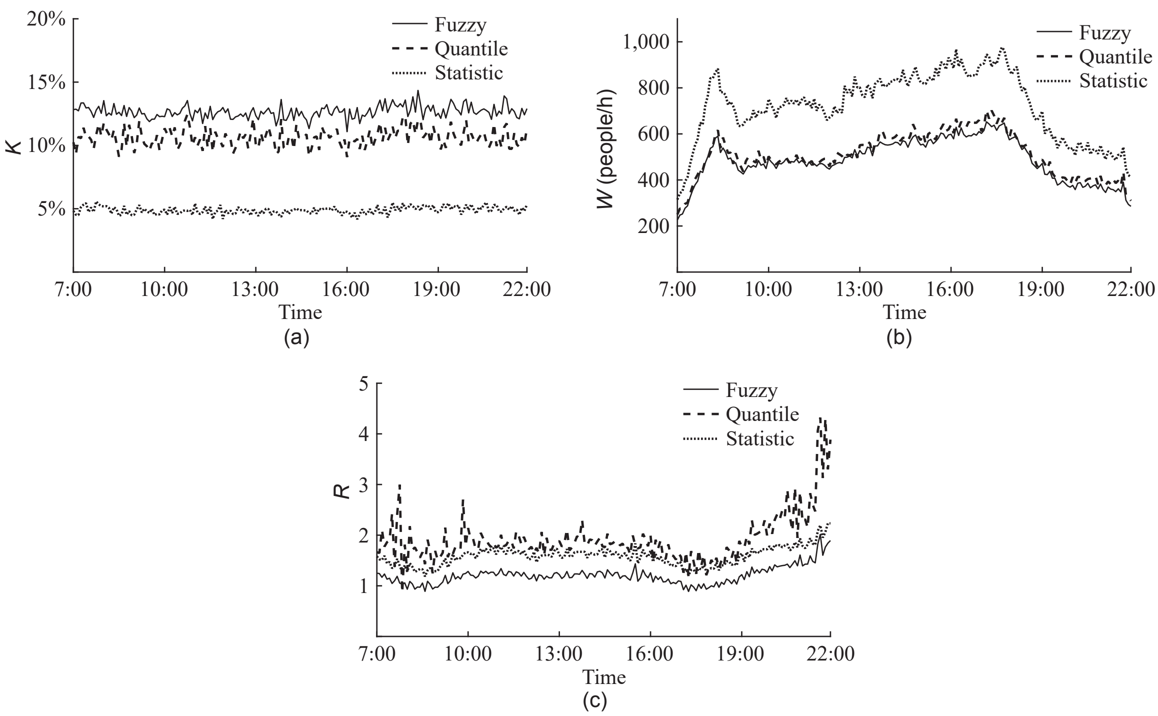

Fig. 4 Comparison of different methods for passenger flow. (a) K of passenger flow; (b) W of passenger flow; (c) R of passenger flow.

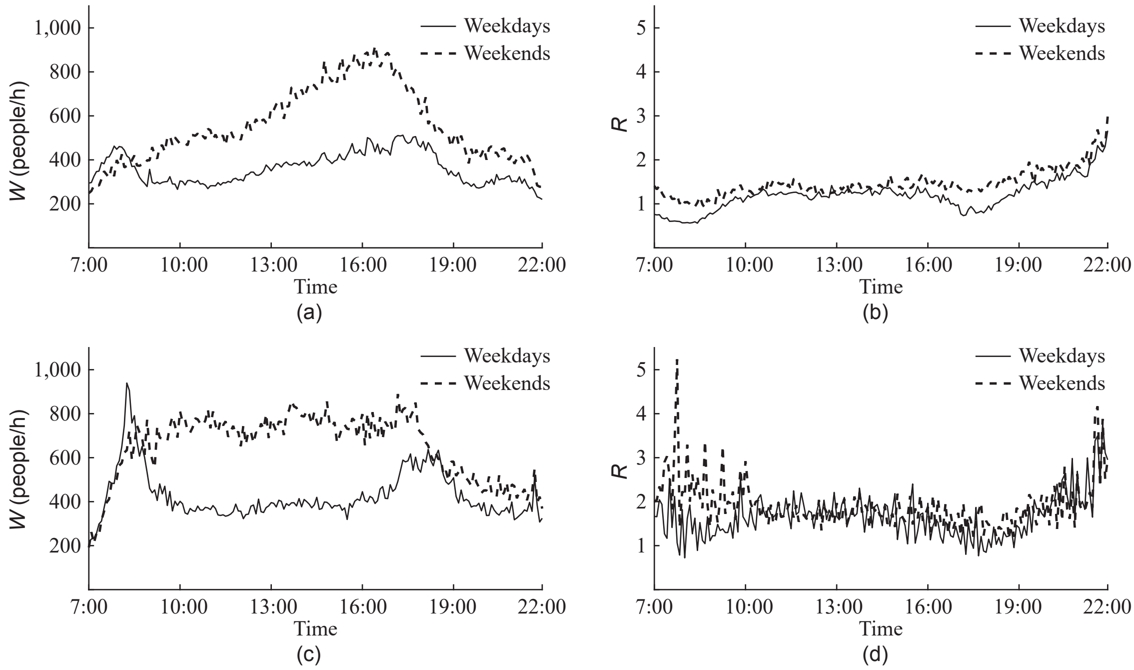

Fig. 5 Comparison of weekdays and weekends on the volatility of passenger flow. (a) The W of inbound passenger flow; (b) The R of inbound passenger flow; (c) The W of outbound passenger flow; (d) The R of outbound passenger flow.

bound passenger flow fluctuates from 1 to over 5.

5 Conclusion

Understanding passenger flow characteristics is essential for improving the operational efficiency and safety of urban rail transit systems. However, existing research on regular or day-to-day recurrent passenger flow patterns in metro networks tends to neglect their inherent volatility. To address this gap, we propose a rolling-window analysis framework to investigate passenger flow volatility based on the three different methods, i.e., fuzzy information granulation, quantile regression, and statistical method. Three different measures were proposed to reflect the volatility of passenger flow. Although the indicators of passenger flow volatility obtained by these three methods were different, they all consistently reveal a similar trend of passenger flow volatility. Using real world passenger flow data from Suzhou metro, we compare the volatility of passenger flow between weekdays and weekends and find that it is reflected by the tradeoff between the K and W of passenger flow. High volatility suggests a wider range of possible changes and a lower K, while lower volatility indicates the opposite. Our results indicate that passenger flow volatility is more significant on weekends than on weekdays.

For future research directions, we can also classify the volatility of short-term passenger flow under different scenarios, including regular conditions and special events. Quantifying the volatility of short-term passenger flow under various scenarios or conditions would be instrumental in identifying the special events that affect passenger flow.

Acknowledgements

The authors would like to thank the Suzhou Rail Transit Group Co., Ltd for providing the necessary data. The authors hold all the responsibility for the analyses and views presented in this work. The authors have no competing interests to declare that are relevant to the content of this article. Neither the entire paper nor any part of its content has been published or has been accepted elsewhere.

Funding

This work is supported by the Natural Science Foundation of the Jiangsu Higher Education Institutes of China under grant No. 23KJB580008.

Declaration of competing interest

The authors have no competing interests to declare that are relevant to the content of this article.

Author contribution statement

Zhao Liu: Conceptualization and writing-original draft. Dan Ning: Data collection and data processing. Wenming Rao: Writing-review. Meng Jia: Data analysis.

Data availability statement

The data used in this paper were provided by Suzhou Rail Transit Group Co., Ltd under a license for this research. The data were not publicly available.

Use of AI statement

During the preparation of this research, we used DeepSeek to improve grammar and readability. After using this tool, we reviewed and edited the content as needed and take full responsibility for content of the publication.

References

[1] Zhang, DP, Wang XC. Transit ridership estimation with network Kriging: A case study of Second Avenue Subway, NYC. Journal of Transport Geography, 2014, 41: 107–115.

[2] Chen, EH, Ye ZR, Wang C, et al. Subway passenger flow prediction for special events using smart card data. IEEE Transactions on Intelligent Transportation Systems, 2020, 21(3): 1109–1120.

[3] Ye, Z, Guan HZ. Fractal quantization of supply and demand of urban rail transit stations. Journal of Highway and Transportation Research and Development, 2013, 30(10): 98–104.

[4] Gu, JJ, Jiang, ZB, Fan, W, et al. Real-time passenger flow anomaly detection considering typical time series clustered characteristics at metro stations. Journal of Transportation Engineering, Part A: Systems, 2020, 146(4): 04020015.

[5] Zhou, F, Xu RH. Model of passenger flow assignmentfor urban rail transit based on entryand exit time constraints. Transportation Research Record: Journal of the Transportation Research Board, 2012, 2284(1): 57–61.

[6] Zhao, JJ, Zhang F, Tu L, et al. Estimation of passenger route choice pattern using smart card data for complex metro systems. IEEE Transactions on Intelligent Transportation Systems, 2017, 18(4): 790–801.

[7] Wang, YH, DE ALMEIDA CORREIA GH, DE ROMPH E, et al. Using metro smart card data to model location choice of after-work activities: An application to Shanghai. Journal of Transport Geography, 2017, 63: 40–47.

[8] Fu, X, Gu Y. Impact of a new metro line: analysis of metro passenger flow and travel time based on smart card data. Journal of Advanced Transportation, 2018, 2018: 9247102.

[9] Guo, JH, Liu Z, Huang W, et al. Short-term traffic flow prediction using fuzzy information granulation approach under different time intervals. IET Intelligent Transport Systems, 2018, 12(2): 143–150.

[10] Li, LC, Ran B, Zhu JS, et al. Coupled application of deep learning model and quantile regression for travel time and its interval estimation using data in different dimensions. Applied Soft Computing, 2020, 93: 106387.

[11] Tsekeris, TS, Stathopoulos A. Short-term prediction of urban traffic variability: Stochastic volatility modeling approach. Journal of Transportation Engineering, 2010, 136(7): 606–613.

[12] Wei, Y, Chen MC. Forecasting the short-term metro passenger flow with empirical mode decomposition and neural networks. Transportation Research Part C: Emerging Technologies, 2012, 21(1): 148–162.

[13] Sun, YX, Leng, B, Guan, W. A novel wavelet-SVM short-time passenger flow prediction in Beijing subway system. Neurocomputing, 2015, 166: 109–121.

[14] Zheng, ZH, Ling XM, Wang, P, et al. Hybrid model for predicting anomalous large passenger flow in urban metros. IET Intelligent Transport Systems, 2020, 14(14): 1987–1996.

[15] Ni, M, He, Q, Gao J. Forecasting the subway passenger flow under event occurrences with social media. IEEE Transactions on Intelligent Transportation Systems, 2016: 1-10.

[16] Rodrigues, F, Borysov SS, Ribeiro B, et al. A Bayesian additive model for understanding public transport usage in special events. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(11): 2113–2126.

[17] Pedrycz, W, Succi G, Sillitti A, et al. Data description: A general framework of information granules. Knowledge-Based Systems, 2015, 80: 98–108.

[18] Sun, Y, Shao C, Ji X, et al. Urban traffic accident time series prediction model based on combination of ARIMA and information granulation SVR. Journal of Tsinghua University (Sci & Technol), 2014, 03: 348–353, 359.