Hindawi
Journal of Advanced Transportation
Volume 2023, Article ID 2661136, 14 pages
https://doi.org/10.1155/2023/2661136

*Research Article*

# An Acceleration Denoising Method Based on an Adaptive Kalman Filter for Trajectory in Merging Zones

**Qiucheng Chen**,[1] **Shunying Zhu**,[1] **Jingan Wu**,[1] **Hongguang Chang**,[1] **and Hong Wang**[2]

[1]*Department of Traffic Engineering, School of Transportation and Logistics Engineering, Wuhan University of Technology, Wuhan, China*
[2]*Department of Road and Bridge Engineering, School of Transportation and Logistics Engineering, Wuhan University of Technology, Wuhan, China*

Correspondence should be addressed to Shunying Zhu; zhusy2001@163.com

Received 24 March 2023; Revised 10 June 2023; Accepted 16 June 2023; Published 27 June 2023

Academic Editor: Zhihong Yao



Vehicle trajectory data can reveal naturalistic driving behaviour trends. However, owing to measurement and processing errors, the trajectory data extracted from videos often contain obvious noise. In merging zones, vehicles tend to accelerate and decelerate frequently, leading to poor denoising performance of the linear Kalman filter (KF). To address this issue, this study proposes a new denoising method based on the adaptive Kalman filter, which automatically switches between KF and Unscented KF to accommodate car-following and merging behaviours, respectively. A merging behaviour detection method was designed based on the PELT method and normalized innovation squared (NIS). The $F1$ score of 92.9% shows the accuracy of behaviour detection. According to our results, the proposed method minimizes the range of jerk compared with other methods, reducing it from −4927.78 to 4960.72 of raw data to −44.92 to 47.14, indicating a significant improvement in denoising and trajectory smoothing. The goal of this study is to achieve high-precision trajectory data under complex real traffic scenarios.

## 1. Introduction

Vehicle trajectory data can be used to calibrate and validate the traffic flow model and provide a broad understanding of driving behaviour[1]. These data are available to the research community and have been widely used in recent years. Among these data, next-generation simulation (NGSIM) data have rich information, becoming the de facto basis of theoretical traffic flow research advances [2]. However, outliers falling out of a limited range cause quality issues in the trajectory data [3]. In addition, processing and measurement errors [4] are evitable [5–7]. Moreover, location and speed amplify noise, negatively affecting the applications of these trajectory data [8]. To achieve high-fidelity trajectory data, researchers have proposed multistep trajectory processing methods [9, 10], denoising speed, acceleration, and space headway [11–13].

Methods for noise removal have improved over the years. Kalman filter (KF) [14, 15] is the most commonly used method. Unlike the feedforward method [16], which uses input signals to predict system response and control, the KF employs a feedback mechanism to continuously update the state estimation and adjust it based on the current observation value. Therefore, KF is an optimal recursive data processing algorithm [17]. Inaccurate or noisy measurements can be used to estimate the state of a linear system with high accuracy. It is also used for trajectory prediction [18, 19]. However, the simple KF is primarily used for constant-acceleration motion and performs well in trajectory noise removal and prediction under car-following scenes [20]. It cannot function correctly if vehicles exhibit complex behaviours, especially when vehicles take two-dimensional movements and interact with other vehicles [21, 22]. For instance, in some complex scenarios, turning

vehicles interact with most traffic flows, which introduce complex features including variation of trajectories [23]. Furthermore, vehicles in merging areas tend to accelerate and decelerate more often owing to the strong competition for space [24]. Therefore, merged areas may be prone to a substantial number of crashes, causing personal injury and property damage. Each merging event has a probability of a near accident [25]. To improve road safety and efficiency levels, a reliable merging area trajectory data are needed to conduct a thorough study on merging behaviour.

To solve these problems, many researchers have used an extended Kalman filter (EKF) [26, 27]. For example, Abbas et al. [28] proposed a multimodel-based EKF used to predict a set of possible scenarios for a vehicle's future position, particularly at an intersection or on a curved path. The EKF uses a low-order Taylor expansion to linearize the nonlinear function; however, the Jacobian matrix of the state vector may be difficult to calculate for complex dynamics. Therefore, the unscented Kalman filter (UKF) has been proposed to solve nonlinear dynamics [29–31]. The UKF performs better than the simple KF and UKF when the system is highly nonlinear [30].

Additionally, researchers have proposed approaches based on maneuver modeling [32, 33]. These approaches meet accuracy requirements with massive data input in the long run; however, they often result in insufficient real-time performance. To achieve a balance between accuracy and real-time performance, the adaptive Kalman filter (AKF) is a relevant option. The term adaptive refers to the ability to adjust the algorithm itself dynamically based on the changing characteristics of the system and shows great robustness in antinoise [34]. For example, Xiao et al. [35] combined the constant turn rate and acceleration (CTRA) motion model and the UKF to predict vehicle trajectory and motion states. Compared with the simple KF and EKF, AKF demonstrates strong adaptability to complicated scenes and reduces the complexity of data processing, in contrast to machine learning methods, which require abundant data and training. However, the aforementioned research does not address how the adaptive process responds to changes in the underlying system, particularly when the system exhibits multiple behaviours that are consistent with complex real-world situations; thus, further research is limited.

This study proposes a method to reconstruct the acceleration data of vehicles in merging zones. First, the constant-acceleration KF and UKF are used to process the acceleration data and calculate the NIS of the KF. Subsequently, a method combined with the pruned exact linear time (PELT) changepoint detection and NIS results is used to identify merging behaviours. Furthermore, outliers are identified and removed during filtering process for noise removal. The outlier-handle method satisfied the Markov assumption made by the KF. Finally, the proposed method is applied to merge zone open data.

The main contributions of this study can be summarized as follows:

(1) A novel adaptive approach to denoise traffic acceleration data is proposed based on the characteristics of car-following and merging behaviours. This approach improves the precision of trajectory data in complex traffic scenarios.

(2) Merging behaviours are identified separately from car-following behaviours. The power of NIS and changepoint detection was leveraged to overcome the limitations of traditional measures such as velocity and acceleration. This method provides a new way to identify different car behaviours with offline data and facilitate more accurate modeling of traffic flow.

(3) Mahalanobis distance is used in this study to measure the distance between an outlier and its distribution, while Chebyshev's theorem is applied to determine the threshold for outliers with raw data. This procedure has significant guiding significance in outlier removal procedures, especially when processing reality data.

The remainder of this paper is organized as follows: Section 2 describes the characteristics of the data and acquisition process. Section 3 describes the proposed method. Section 4 presents the results of the jerk analysis and a comparison with other methods. This paper concludes with a discussion and conclusions in Section 5.

## 2. Trajectory Data

*2.1. Data Description.* Open data downloadable at https://seutraffic.com/ in the ubiquitous traffic eye database [36] were used to verify the proposed method. The video was recorded at the Kazimen interchange in Nanjing City, China (Figure 1) using a DJI Mavic Pro. The unmanned aerial vehicle (UAV) flew for 1184 s at a height of 285 m and covered approximately 352 m. Finally, 2513 trajectories were recorded. After running the image recognition algorithm, the data contained vehicle IDs, timestamps, lane logs, speed, acceleration, and vehicle locations that were saved in the Frenet frame.

The east-west direction, where the merging zone belongs, was selected as the research section. The results of the study by Westphal [37] showed that the merging behaviour of vehicles on ramps only affected the outermost lane of the main lane. Thus, this study focused on lanes 1–3. Figure 2 shows lane IDs and traffic directions.

First, the macroscopic characteristics of traffic flow are described. Figure 3 shows that the velocity distribution is a normal distribution with a 50 km/h mean, consistent with the actual speed distribution. Reasonable ranges in acceleration and deceleration were established as 0 $m/s^2$ to 5 $m/s^2$ and $-6$ $m/s^2$ to 0 $m/s^2$ [38], respectively. Compared to the raw data, the percentage of acceleration $>2$ $m/s^2$ and deceleration $<-2$ $m/s^2$ were 0.05% and 0.01%, respectively, indicating normal driving conditions.

Second, the outliers of acceleration and deceleration with a limit range of $-6$ $m/s^2$ to 5 $m/s^2$ were analyzed, as shown in Figure 4. Most acceleration values in Figure 4 fall within the range of $-6$ $m/s^2$ to 5 $m/s^2$, but some outliers exist, such as maximum deceleration and acceleration values of $-150$ $m/s^2$

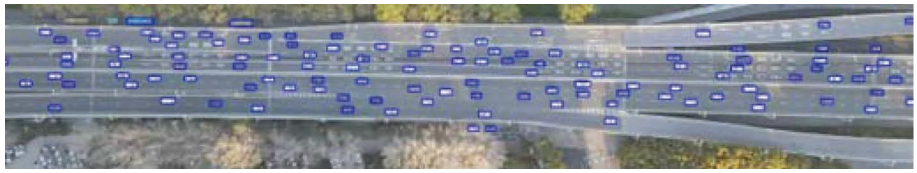

Figure 1: Video recognition of Kazimen interchange.

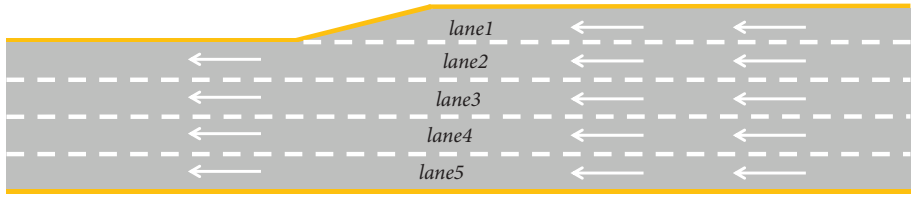


Figure 2: Schematic of the measured road and traffic directions.

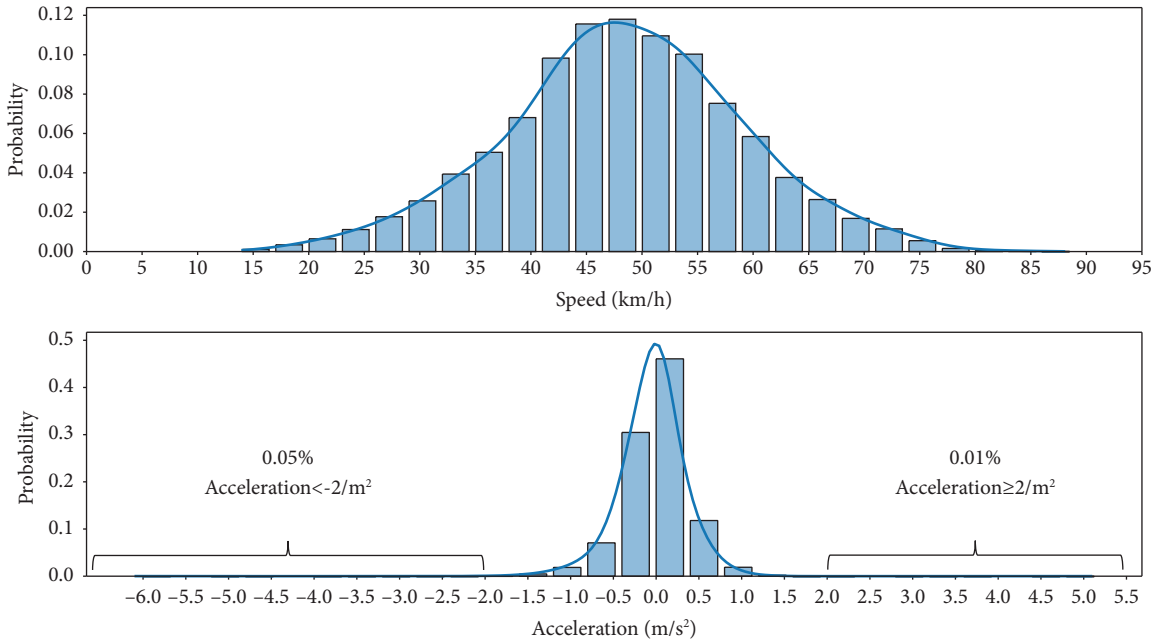


Figure 3: Speed and acceleration distribution histogram.

and 50 m/s$^2$, respectively. Outliers were handled during the filtering process to avoid adverse effects on the final result caused by smoothing them before filtering, which violates the Markov assumption of KF on the observed data.

To determine the frequent interval of merging behaviour, moving least squares (MLS) was used to reconstruct the spatiotemporal evolution of the traffic flow in lanes 1–3. This can fit the instantaneous speed distribution per second for each vehicle. Figure 5 shows the characteristics of the velocity drop and recovery in different spatial and temporal dimensions.

Figure 5 shows that the driving speed was relatively high when most vehicles did not merge, particularly in the first 400 s. During this period, the vehicles primarily moved at a constant velocity or acceleration. By contrast, the acceleration varied, as can be observed in the range of 800 s to 1000 s. Notably, between 200 s and 400 s, lane 1 experienced a drop in speed followed by a recovery. The other two lanes were unaffected by this phenomenon with relatively small changes in speed rate. Note that the dynamics of speed change were similar, and all of them almost dropped to the minimum synchronously in a local space-time region where vehicles merged into main lanes with high probability. This observation provides a preliminary basis for designing the AKF threshold.

*2.2. Data Preprocessing.* The location data in Figure 6 are in the Frenet frame, which is based on a given curve and represents a point's position by displacement along the reference line and lateral offset distance. The Frenet frame is useful because any road can be standardized as a straight tunnel, simplifying trajectory planning for curved roads.

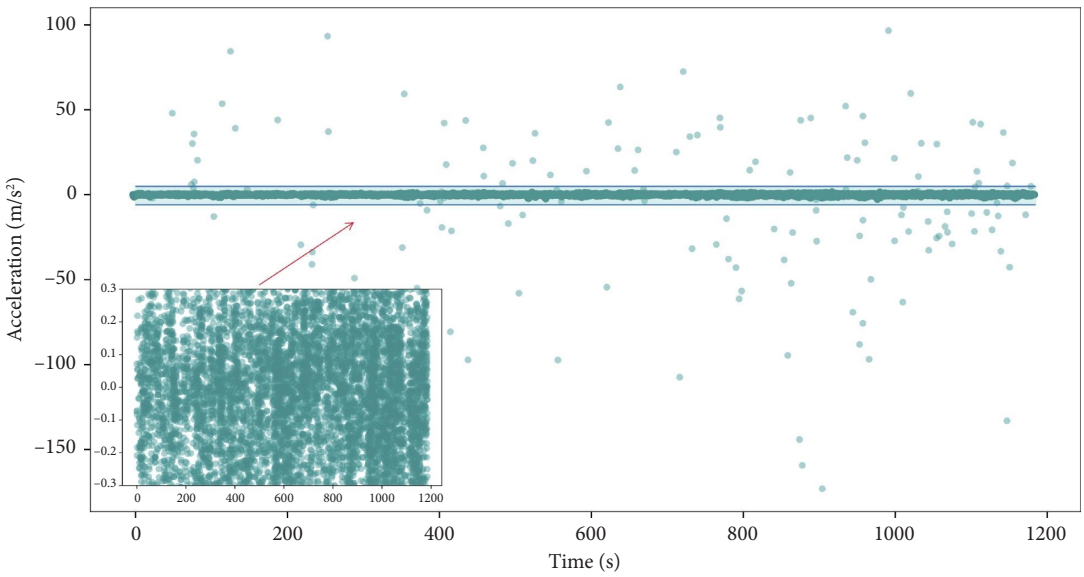


Figure 4: Scatter plot of the acceleration distribution.

However, the Frenet frame has separate coordinate systems for each lane, making it difficult to represent lane-crossing behaviours. Therefore, Frenet data should be converted to a global Cartesian frame.

The process of converting the data from the Frenet frame to the Cartesian frame is divided into the following steps:

*2.2.1. Selection of Reference Lines.* The reference lines are the upper and lower boundary lines of the lane and centerline. The OpenCV library was utilized to identify boundaries and extract reference points from aerial videos, and the resulting selection is shown in Figure 7. Subsequently, the locations of reference points in the Cartesian frame can be determined based on the pixel scale e.g., $(x_1, x_2, x_3, \ldots, x_n)$ and $(y_1, y_2, y_3, \ldots, y_n)$ successively. For the centerline, the location of the reference points can be determined by calculating the median of the corresponding points in the upper and lower boundaries.

*2.2.2. Coordinate Transformation.* For any point in the Frenet coordinate system $\mathbf{P}^i_f$, let $\mathbf{P}^i_f = \begin{bmatrix} s_i \\ d_i \end{bmatrix}$, where $s_i$ is the projected distance, $d_i$ is the lateral offset, and the corresponding point in the Cartesian coordinate system is $\mathbf{P}^i_c = \begin{bmatrix} a_i \\ b_i \end{bmatrix}$, which is between two adjacent reference points; that is, $\mathbf{P}^i_z = \begin{bmatrix} x_i \\ y_i \end{bmatrix}$ and $\mathbf{P}^{i+1}_z = \begin{bmatrix} x_{i+1} \\ y_{i+1} \end{bmatrix}$. Therefore, we have $x_i \le a_i \le x_{i+1}$ and $y_i \le b_i \le y_{i+1}$. Figure 8 illustrates the relationship between the Frenet coordinate system and the Cartesian coordinate system. $\mathbf{P}^i_c$ can be calculated using the following equation:

$$\begin{bmatrix} a_i \\ b_i \end{bmatrix} = \begin{bmatrix} x_i \\ y_i \end{bmatrix} + \Delta s \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} + d_i \begin{bmatrix} \sin\theta \\ \cos\theta \end{bmatrix}, \tag{1}$$

where $\Delta s$ is the distance between $(a_i, b_i)$ and $(x_i, y_i)$ and $\theta$ is the angle between the line passing through points $\mathbf{P}^i_z$ and $\mathbf{P}^i_f$ and the $x$-axis.

Subsequently, the difference in displacement along the reference line between $\mathbf{P}^i_f$ and the nearest reference point can be approximated by

$$\Delta s = s_i - \sum_{i=0}^{n} \left\| \mathbf{P}^{i+1}_z - \mathbf{P}^i_z \right\|_2. \tag{2}$$

Let $\mathbf{P}^1_z = \mathbf{P}^0_z$. Accordingly, $\Delta s = s_i$ if the nearest reference point below is the starting point of the reference line.

Substituting $\Delta s$ into equation (1) gives the Cartesian coordinate of $\mathbf{P}^i_c$:

$$\mathbf{P}^i_c = \begin{bmatrix} x_i \\ y_i \end{bmatrix} + \left( s_i - \sum_{i=0}^{n} \left\| \mathbf{P}^{i+1}_z - \mathbf{P}^i_z \right\|_2 \right) * \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} + d_i * \begin{bmatrix} \sin\theta \\ \cos\theta \end{bmatrix}. \tag{3}$$

## 3. Proposed Methods

*3.1. Adaptive Kalman Filter Design.* The KF was named after Rudolf E. Kalman, one of the primary developers of its theory [39]. It can effectively reduce measurement noise and estimate true values in trajectory processing while remaining robust to model uncertainties. In vehicle dynamics, the variance of the acceleration change is rather small if the vehicle is running at a constant speed or acceleration. Therefore, the constant-acceleration KF is applied with the assumption that the initial acceleration equals zero.

For linear discrete dynamic systems

$$\mathbf{X}(k) = \mathbf{F}\mathbf{X}(k-1) + \mathbf{w}(k), \tag{4}$$

$$\mathbf{Z}(k) = \mathbf{H}(k)\mathbf{X}(k) + \mathbf{v}(k), \tag{5}$$

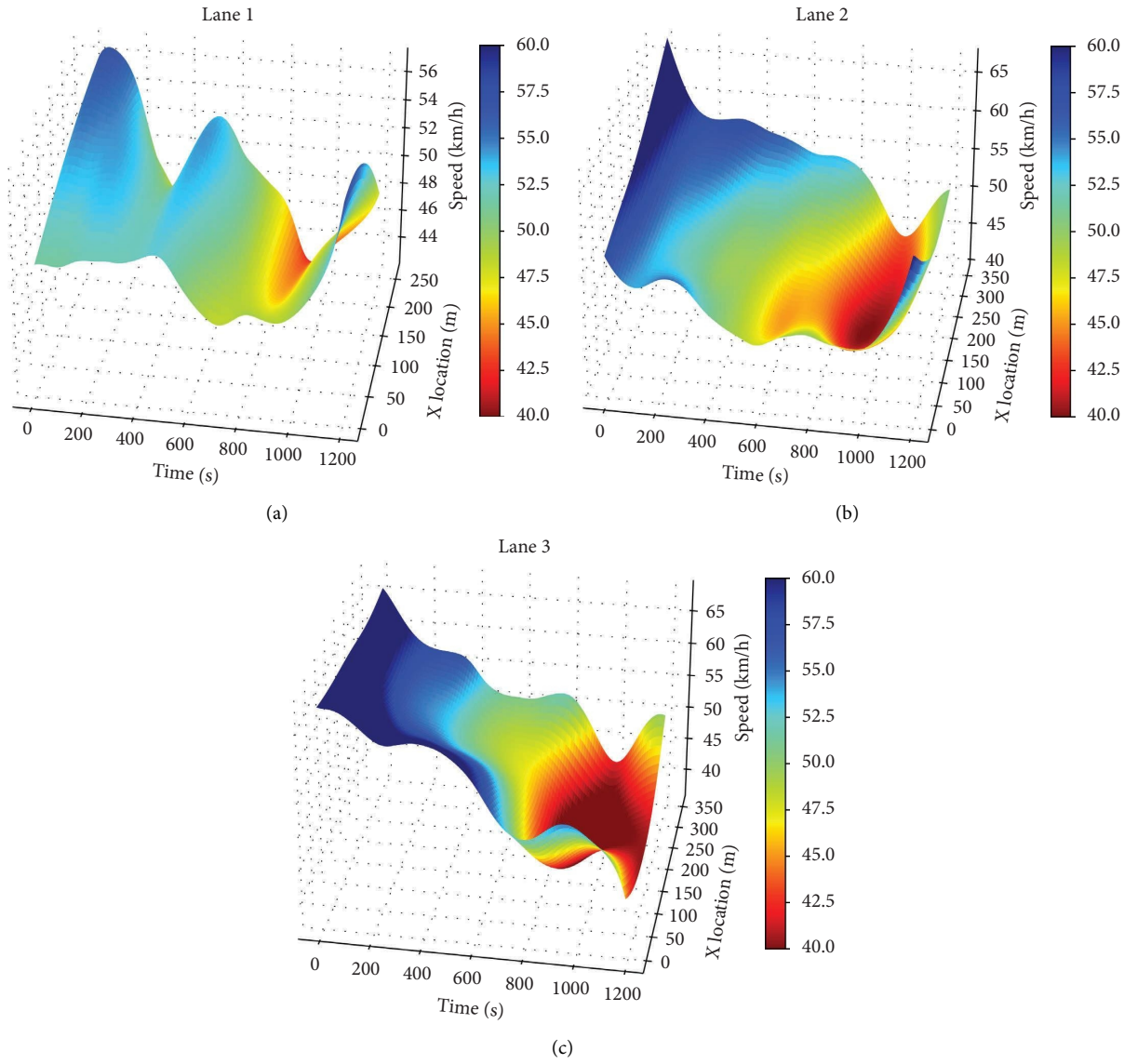


Figure 5: Visualization of traffic spatiotemporal evolution.

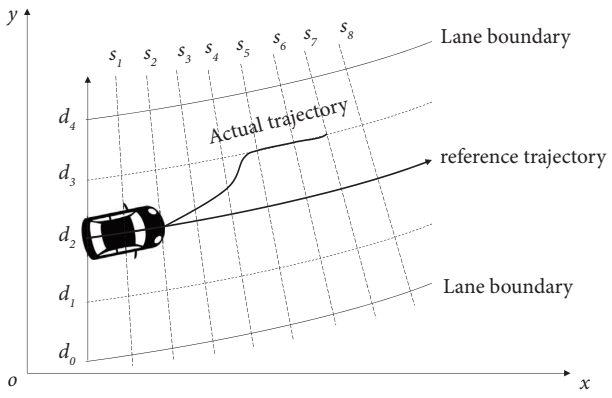


Figure 6: Illustration of the Frenet frame.

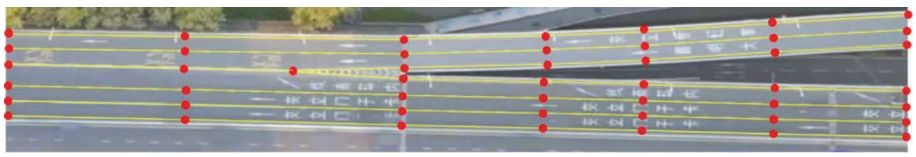

Figure 7: Selection of reference points.

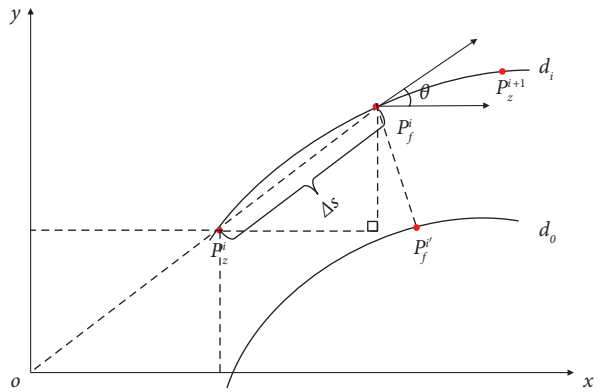


Figure 8: Relationship between the Frenet coordinate system and the Cartesian coordinate system.

where $\mathbf{X}(k)$ is the state vector at time $k$ and $\mathbf{F}$ is the state transition matrix. $\mathbf{w}(k)$ represents the process noise, a multivariate normal distribution with a mean of 0.

In equation (5), $\mathbf{Z}(k)$ is the measurement of the real state $\mathbf{X}(k)$, $\mathbf{H}$ is the observation model that maps the real state space to the observed state space, and $\mathbf{v}(k)$ is the observation noise, assumed to be zero-mean Gaussian white noise with covariance $\mathbf{R}(k)$.

To predict and update the state estimation, KF can be explained as follows:

$$\begin{aligned}
\hat{\mathbf{X}}(k|k-1) &= \mathbf{F}\hat{\mathbf{X}}(k-1|k-1),\\
\mathbf{P}(k|k-1) &= \mathbf{F}\mathbf{P}(k-1|k-1)\mathbf{F}_k^T + \mathbf{Q}(k),\\
\mathbf{K}(k) &= \mathbf{P}(k|k-1)\mathbf{H}^T(k)\left(\mathbf{H}(k)\mathbf{P}(k|k-1)\mathbf{H}^T(k) + \mathbf{R}(k)\right)^{-1},\\
\hat{\mathbf{X}}(k|k) &= \hat{\mathbf{X}}(k|k-1) + \mathbf{K}(k)\left(\mathbf{Z}(k) - \mathbf{H}(k)\hat{\mathbf{X}}(k|k-1)\right),\\
\mathbf{P}(k|k) &= (\mathbf{I} - \mathbf{K}(k)\mathbf{H}(k))\mathbf{P}(k|k-1),
\end{aligned} \tag{6}$$

where $\hat{\mathbf{X}}$ is the a priori state estimate, $\mathbf{P}$ is the a posteriori estimate covariance matrix, and $\mathbf{Q}(k)$ is the covariance matrix of $\mathbf{w}(k)$. $\mathbf{I}$ is the identity matrix and $\mathbf{K}$ is the Kalman gain.

Finally, the state vector $\mathbf{X}(k)$ for constant-acceleration motion is given by

$$\mathbf{X}(k) = \begin{bmatrix} L_k \\ V_k \\ a_k \end{bmatrix} = \begin{bmatrix} x \\ V_k \\ a_k \end{bmatrix}, \tag{7}$$

where $L_k$ is the distance moved, $V_k$ is the velocity, and $a_k$ is the acceleration at time $k$.

However, dealing with nonlinear processes is difficult for a standard KF. When merging behaviour occurs, vehicles are affected by the steering wheel angle and throttle opening; thus, the acceleration distribution exhibits apparent nonlinear characteristics. If a standard KF is applied to these accelerations, the nonlinear process will be considered as large observation noise, leading to a slower prediction changing trend compared with the observation. Trajectory data extracted from aerial videos primarily include location, velocity, and acceleration; thus, they lack microscopic data of steering wheel angle and throttle opening for the standard KF. Therefore, this study used the UKF to simulate the merging process.

As this study is primarily concerned with merging behaviour, the road cross-section was chosen from the nose of the ramp, where lanes 1 and 2 overlapped completely with the specific scene. In addition to the lateral and longitudinal displacements, merging behaviour also involves the heading angle. Therefore, polar coordinates are more suitable to

represent the merging angle between the driving direction of the vehicle at time $k$ and the direction of lateral displacement. This angle will be included in the state vector $\mathbf{X}(k)$ of the UKF.

*3.2. Removal of Outliers.* To avoid conflicting with the Markov assumption, we handle outliers during the filtering process instead of processing them before. This ensures that smoothed outliers do not contain information from previous or subsequent values. To achieve this, we use the Mahalanobis distance to measure the distance between a point and its distribution.

$$D_m = \sqrt{(\mathbf{X} - u)^T \mathbf{P}^{-1} (\mathbf{X} - u)}, \tag{8}$$

where $u$ is the mean of the normal distribution, $\mathbf{X}$ is the state vector, and $\mathbf{P}$ is the a posteriori estimate covariance matrix.

Figure 4 illustrates that the acceleration data had numerous outliers, which can be classified into two categories: those exceed the vehicle's dynamic limit and outliers within a reasonable interval but are isolated and deviated from the overall trend. These two categories can be distinguished by comparing the Mahalanobis distance. Typically, 99.7% of a normal distribution falls within three standard deviations from the mean; however, since this distribution is only approximately normal, Chebyshev's theorem suggests that about 96% of data fall within five standard deviations. Therefore, in this case, five standard deviations are used as a threshold to identify outliers.

*3.3. Merging Behaviour Detection.* In trajectory data, it is commonly assumed that merging behaviour occurs instantaneously, with a vehicle's lane ID changing after a single frame. However, in reality, merging behaviour takes a variable amount of time to complete, making it difficult to identify solely based on measures such as a vehicle's velocity or acceleration. To address this issue, a new method is proposed that combines NIS [40] and changepoint detection to effectively identify merging behaviour.

The NIS is used to check whether a KF is consistent with measured residuals and the associated innovation covariance matrix. By normalizing the residuals using covariance and analyzing the resulting NIS, it is possible to identify situations where the KF assumption of constant car-following behaviour may not hold true. In this way, trajectory data that do not correspond to car-following scenarios can be filtered out based on predictions and measurements. The NIS can be calculated as

$$\begin{aligned} \gamma(k) &= \mathbf{Z}(k) - \mathbf{H}(k)\hat{\mathbf{X}}(k|k-1), \\ \mathrm{NIS} &= \gamma(k)^T \left(\mathbf{H}(k)\mathbf{P}(k|k-1)\mathbf{H}^T(k) + \mathbf{R}(k)\right)^{-1} \gamma(k), \end{aligned} \tag{9}$$

where $\boldsymbol{\gamma}(k)$ is the innovation vector.

However, it should be noted that while a drastic change in the NIS may indicate the emergence of braking in a car-following scenario, it does not necessarily imply that the vehicle is merging. Therefore, in order to accurately identify merging behaviours, it is necessary to look beyond just prediction and measurements.

To effectively utilize trajectory data, we employed the PELT changepoint detection method. This method can be easily integrated into the KF and offers a good balance between algorithm complexity and detection accuracy, especially when the data structure is not overly complex. The PELT algorithm uses observations to identify potential acceleration changepoints in the time series. It further divides the time interval when the NIS fluctuates, which helps capture the variable and complex nature of merging behaviour over time. By examining these results, merging behaviour can be determined from a temporal perspective. The PELT detection method is as follows:

$$\min \sum_{i=1}^{k+1} \left[C\left(o_{\tau_{i-1}+1:\,\tau_i}\right)\right] + \beta f(k), \tag{10}$$

where $o$ denotes acceleration, $k$ is the number of changepoints in the history, and the acceleration time series were split into $k+1$ distinct segments. $\tau_j$ represents the location of the $j^{\text{th}}$ changepoint for $j = 1, \ldots, k$. To simplify the analysis, without loss of generality, $\tau_0 = 0$ and $\tau_{k+1} = n$ were also set. $C$ is the cost function of each segment, while $\beta f(k)$ is a penalty item that prevents overfitting, and $\beta$ is a constant independent of the position and number of acceleration changepoints. As a common choice, $L2$ normalization was selected as the cost function.

For a given acceleration segment, the PELT method assumes a constant $M$ for all $s < m < t$ that satisfies

$$C\left(o_{(s+1):m}\right) + C\left(o_{(s+1):t}\right) + M \le C\left(o_{(s+1):t}\right), \tag{11}$$

where $s$ and $t$ are the start and end points of the segment, respectively.

For $\tau^* = 1, \ldots, t - s + 1$, we have

$$H\left(\tau^*\right) = \min_{\tau \in R_{\tau^*}} \left[H(\tau) + C\left(o_{\tau+1:\tau^*}\right) + \beta\right], \tag{12}$$

where $H(0) = -\beta$, $R_{\tau^*}$ represents the set of possible values for $\tau$, and the number of elements in the set increased with the recursive calculation. The initial value is $R_1 = \{0\}$. Let

$$\tau^1 = \operatorname{argmin}_{\tau \in R_{\tau^*}} \left[H(\tau) + C\left(o_{\tau+1:\tau^*}\right) + \beta\right]. \tag{13}$$

Accordingly,

$$\begin{aligned} \mathrm{cp}\left(\tau^*\right) &= \left[\mathrm{cp}\left(\tau^1\right), \tau^1\right], \\ R_{\tau^*+1} &= \left\{\tau \in R_{\tau^*} \cup \left\{\tau^*\right\} \colon H(\tau) + C\left(o_{\tau+1:\tau^*}\right) + M \le H\left(\tau^*\right)\right\}, \end{aligned} \tag{14}$$

where $\mathrm{cp}(0) = \mathrm{NULL}$ and $\mathrm{cp}(n)$ is the detected acceleration changepoint.

Thus, the merging behaviour is identified further by detecting the acceleration changepoint.

The value of $\beta$ is crucial for the PELT results. Many false changepoints will be detected if the penalty value is too low. By contrast, real changepoints may not be found if the penalty value is too high. In addition to manually selecting parameters with good performance after several trials, parameters can be selected according to the following formula:

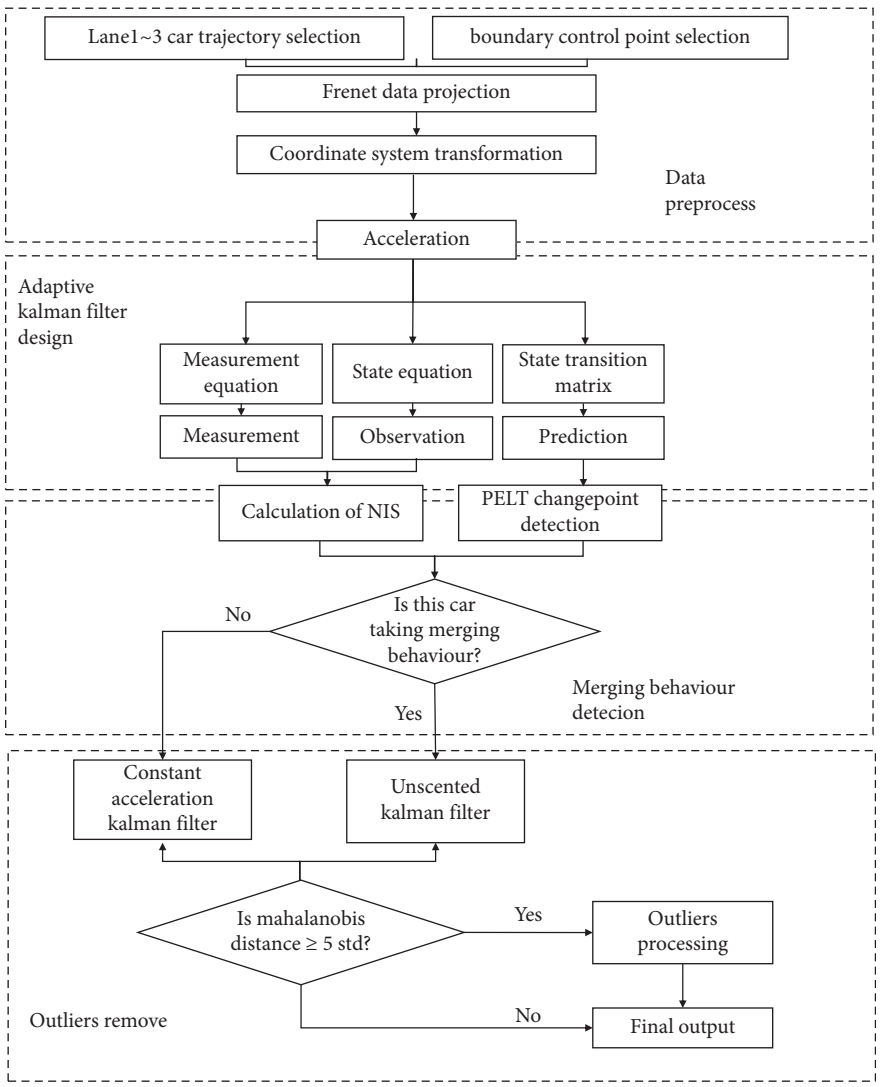


Figure 9: Framework of the proposed method.

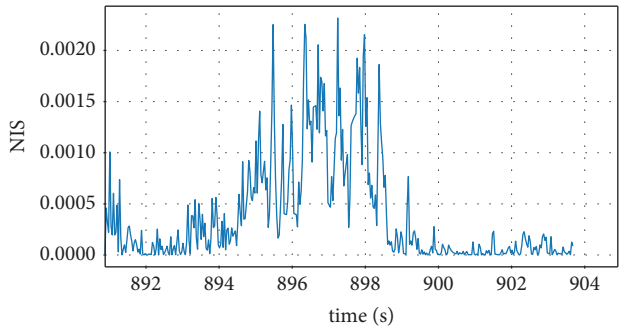


Figure 10: NIS calculation result of No. 173 by constant-acceleration KF.

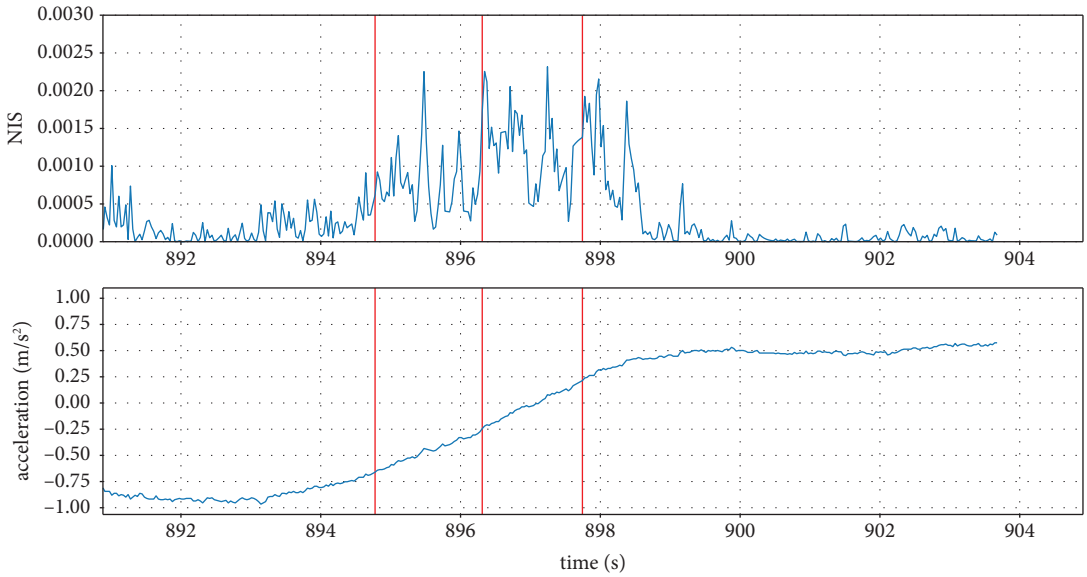


Figure 11: Comparison between the result of acceleration changepoint detection and NIS.


Figure 12: Merging behaviour of Vehicle No. 173.

$$\beta = (0.25n)^{1-\alpha} 2 \ln(n), \tag{15}$$

where $\alpha$ is the sensitivity ranging 0 to 1. When $\alpha$ is 1, it corresponds to minimizing the Bayesian information criterion (BIC).

By combining NIS and PELT changepoint detection, we were able to determine the merging behaviour within a specified space range and period. The framework of this method is illustrated in Figure 9.

## 4. Results

Vehicle No. 173 was chosen as a case study to examine the proposed method's process. The acceleration data were analyzed using a constant-acceleration KF and the NIS was calculated based on measurements and predictions. Figure 10 displays the results, which revealed fluctuations between 894 and 898 seconds. Consequently, PELT changepoint detection was continuously applied to the acceleration data of Vehicle No. 173 (marked in red in Figure 11) based on observation data. As shown in Figure 11, the changepoints' locations fell within an intense growth period of NIS, indicating merging behaviour occurred. The comparison with aerial video (Figure 12) showed that during the merging zone, Vehicle No. 173 crossed longitudinally from lane one to lane three, which means merging behaviour really happened; this result is consistent with our detection outcome. In total, we used 923 vehicle trajectories for the experiment and applied our proposed method to these trajectories. Our method achieved a merging behaviour detection recall rate of 92.3% and precision of 93.5%, resulting in an $F1$ score of 92.9%. These findings indicating that the proposed method performs well in detecting merging behaviours.

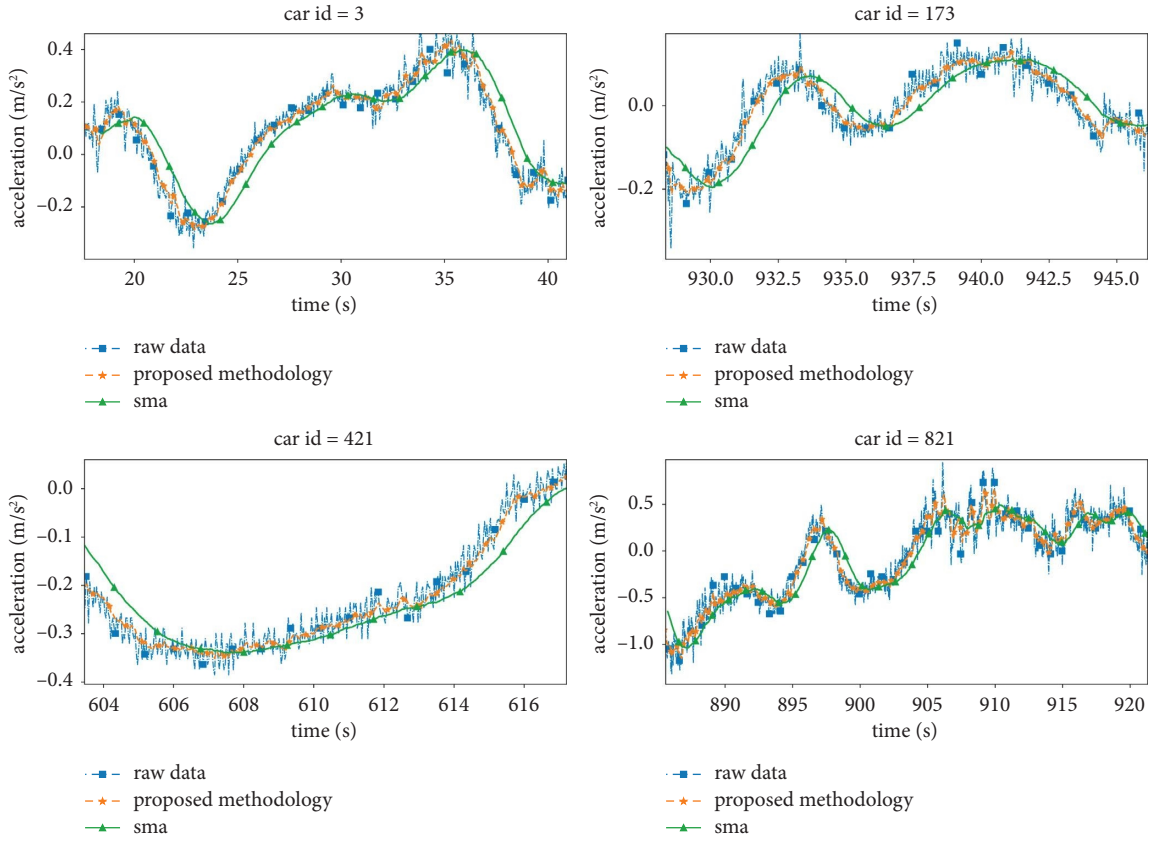


Figure 13: Result of the proposed method compared with the SMA and raw data.

Directly comparing and verifying the improvement effect is not feasible since the ground truth is unknown. Therefore, to evaluate the filtering performance, first, we randomly selected multiple trajectory data from different periods. We compared our method with the simple moving average (SMA) because it can remove high-frequency noise while preserving the underlying trends and patterns and is also robust to outliers. Figure 13 shows the comparison of filtering effect. The SMA processing lags behind the trend of raw data for all vehicles except No. 821. The acceleration data of No. 821 change abruptly both in the overall and local trends, which are well balanced by SMA. However, this may lead to information loss, especially when there is a significant difference between maximum and minimum values.

Secondly, we utilized indirect quantitative indicators, specifically jerk analysis, for comparison. The advantage of jerk analysis in vehicle acceleration judgment lies in providing a more comprehensive and precise understanding of a vehicle's motion dynamics. By analyzing and comparing indicators such as the percentage of jerk sign reversal within a 1-second window, the percentage of the absolute value of jerk above a given threshold, and the range of jerk between different methods, we can better evaluate the mechanical and human-machine feasibility of acceleration data, and consequently, assess the effectiveness of the denoising process.

Table 1 shows that the proposed method significantly minimized the range of jerk compared with other methods, reducing it from −4927.78 to 4960.72 of raw data to −44.92 to 47.14. Additionally, a good reduction effect was achieved in the average value for the percentage of the absolute value of jerk greater than 15 $m/s^3$. For the percentage of jerk sign reversal within 1 s, the degree of average reduction was second to that of the SMA. These findings demonstrate the effectiveness of the proposed method. Note that the maximum sign inversion of jerk was 100% in all columns owing to the characteristics of the data itself. Some vehicles entered lane 3 from lane 4 in a short period of time, and the driving time was relatively short; therefore, the sign reversal was 100%.

To demonstrate the transferability of the proposed method, we conducted experiments using the NGSIM I-80 dataset, which is one of the most popular datasets for studying merging behaviour. The results are presented in Table 2, with Jerk analysis results cited from Rafati Fard et al.

Table 1: Jerk analysis results.

| Variable | | Raw data | Kalman filter | Simple moving average | Proposed method |
|---|---|---|---|---|---|
| Jerk | Std | 28.55 | 10.53 | 1.51 | 0.46 |
| | Mean | 0.01 | 0.01 | 0.00 | 0.00 |
| | Max | 4960.72 | 1475.16 | 198.46 | 47.14 |
| | Min | −4927.78 | −1765.94 | −197.33 | −44.92 |
| Jerk values higher than ± 15 (%) | Std | 0.15 | 0.39 | 0.18 | 0.09 |
| | Mean | 0.05 | 0.13 | 0.06 | 0.02 |
| | Max | 2.35 | 2.26 | 1.05 | 1.96 |
| | Min | 0 | 0 | 0 | 0 |
| 1-s window with more than one sign inversion of jerk (%) | Std | 5.69 | 7.81 | 21.73 | 14.30 |
| | Mean | 99.67 | 88.20 | 70.89 | 73.61 |
| | Max | 100 | 100 | 100 | 100 |
| | Min | 0 | 0 | 0 | 0 |

Table 2: Comparison of different approaches using NGSIM I-80 data.

| Variable | | Raw data | Kalman filter | Simple moving average | Proposed method |
|---|---|---|---|---|---|
| Jerk | Std | 33.40 | 7.98 | 2.95 | 3.41 |
| | Mean | 0.00 | 0.00 | 0.00 | 0.00 |
| | Max | 1640 | 91 | 66 | 53.67 |
| | Min | −1846 | −64 | −59 | −54.33 |
| Jerk values higher than ± 15 (%) | Std | 11.22 | 2.88 | 0.97 | 0.50 |
| | Mean | 22.83 | 0.30 | 0.29 | 0.16 |
| | Max | 69.23 | 22.16 | 11.35 | 4.17 |
| | Min | 4.70 | 0.00 | 0.00 | 0.00 |
| 1-s window with more than one sign inversion of jerk (%) | Std | 10.31 | 12.64 | 14.16 | 4.95 |
| | Mean | 75.23 | 21.22 | 56.10 | 13.54 |
| | Max | 100 | 82.05 | 81.82 | 80.42 |
| | Min | 42.57 | 10 | 6.53 | 7.40 |

[10]. As can be seen, the proposed method achieved good denoising effectiveness in this dataset, reducing the range of Jerk from −1846 to 1640 of raw data to −54.33 to 53.67. Additionally, the proposed method achieves good performance in the other two indicators.

## 5. Summary and Conclusions

This paper presents a novel approach for processing acceleration data in vehicle merging zones, utilizing the unique characteristics of car-following and merging behaviours. The proposed method employs an adaptive mechanism that identifies different vehicle behaviours based on the results of NIS and PELT, taking advantage of the characteristic differences produced by AKF. Outliers are detected using Mahalanobis distance during the filtering process to ensure the accuracy of predictions and maintain data integrity, significantly enhancing the precision of our method.

Compared to the commonly used simple moving average (SMA) method, the proposed approach shows a more accurate filtering result and strong applicability to merging scenes. The resulting trajectory data also suggest a more reliable set of characteristics corresponding with mechanical and human-machine feasibility. However, to improve the generalizability of our approach, further studies are required to investigate the changepoint detection method and parameters in relation to specific traffic scenes and traffic flow characteristics.

In conclusion, our proposed method demonstrates a significant improvement in the processing of acceleration data in vehicle merging zones. By utilizing adaptive mechanisms, our approach offers a more accurate and reliable set of trajectory data, which has promising applications in traffic safety and traffic flow optimization.

## Data Availability

The trajectory data used to support the findings of this study are included within the article.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This study was financially supported by the National Natural Science Foundation of China (No. 52272337).

## References

[1] S. P. Hoogendoorn, H. J. Van Zuylen, M. Schreuder, B. Gorte, and G. Vosselman, "Microscopic traffic data collection by remote sensing," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1855, no. 1, pp. 121–128, 2003.

[2] R. Krajewski, J. Bock, L. Kloeker, and L. Eckstein, "The highD dataset: a drone dataset of naturalistic vehicle trajectories on German highways for validation of highly automated driving systems," in *Proceedings of the 21st International Conference on Intelligent Transportation Systems (ITSC)*, pp. 2118–2125, Maui, HI, USA, November, 2018.

[3] F. Marczak and C. Buisson, "New filtering method for trajectory measurement errors and its comparison with existing methods," *Transportation Research Record*, vol. 2315, no. 1, pp. 35–46, 2012.

[4] P. Chen, L. Wei, F. Meng, and N. Zheng, "Vehicle trajectory reconstruction for signalized intersections: a hybrid approach integrating Kalman filtering and variational theory," *Transportation Business: Transport Dynamics*, vol. 9, no. 1, pp. 22–41, 2020.

[5] X. Y. Lu and A. Skabardonis, "Freeway traffic shockwave analysis: exploring the NGSIM trajectory data," in *Proceedings of the 86th Annual Meeting of the Transportation Research Board*, pp. 21–25, Washington, DC, USA, February, 2007.

[6] V. Punzo, M. T. Borzacchiello, and B. Ciuffo, "Estimation of vehicle trajectories from observed discrete positions and next-generation simulation program (NGSIM) data," in *Proceedings of the TRB 2009 Annual Meeting*, Washington, DC, USA, November, 2009.

[7] B. Coifman and L. Li, "A critical evaluation of the Next Generation Simulation (NGSIM) vehicle trajectory dataset," *Transportation Research Part B: Methodological*, vol. 105, no. 4, pp. 362–377, 2017.

[8] M. Treiber and A. Kesting, "Microscopic calibration and validation of car-following models–a systematic approach," *Procedia-Social and Behavioural Sciences*, vol. 80, no. 10, pp. 922–939, 2013.

[9] M. Montanino and V. Punzo, "Making NGSIM data useable for studies on traffic flow theory: multistep method for vehicle trajectory reconstruction," *Transportation Research Record*, vol. 2390, pp. 99–111, 2013.

[10] M. Rafati Fard, A. Shariat Mohaymany, and M. Shahri, "A new methodology for vehicle trajectory reconstruction based on wavelet analysis," *Transportation Research Part C: Emerging Technologies*, vol. 74, pp. 150–167, 2017.

[11] A. Duret, C. Buisson, and N. Chiabaut, "Estimating individual speed-spacing relationship and assessing ability of Newell's car-following model to reproduce trajectories," *Transportation Research Record*, vol. 2088, pp. 188–197, 2008.

[12] V. Punzo, M. T. Borzacchiello, and B. Ciuffo, "On the assessment of vehicle trajectory data accuracy and application to the Next Generation SIMulation (NGSIM) program data," *Transportation Research Part C: Emerging Technologies*, vol. 19, no. 6, pp. 1243–1262, 2011.

[13] S. Y. Dong, Zhou, T. Chen, S. Li, Q. Gao, and B. Ran, "An integrated Empirical Mode Decomposition and Butterworth filter based vehicle trajectory reconstruction method," *Physica A: Statistical Mechanics and Its Applications*, vol. 583, no. 2, Article ID 126295, 2021.

[14] X. Chen, Z. Li, Y. Yang, L. Qi, and R. Ke, "High-resolution vehicle trajectory extraction and denoising from aerial videos," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 5, pp. 3190–3202, 2020.

[15] J. Mu, Y. Han, C. Zhang, J. Yao, and J. Zhao, "An unscented Kalman filter-based method for reconstructing vehicle trajectories at signalized intersections," *Journal of Advanced Transportation*, vol. 2021, Article ID 6181242, 12 pages, 2021.

[16] S. Wang, P. Takyi-Aninakwa, S. Jin, C. Yu, C. Fernandez, and S. Daniel-Ioan, "An improved feedforward-long short-term memory modeling method for the whole-life-cycle state of charge prediction of lithium-ion batteries considering current-voltage-temperature variation," *Energy*, vol. 254, Article ID 124224, 2022.

[17] G. Welch and G. Bishop, *An Introduction to the Kalman Filter*, University of North Carolina, Chapel Hill, NC, USA, 1995.

[18] J. Schulz, C. Hubmann, J. Löchner, and D. Burschka, "Multiple model unscented Kalman filtering in dynamic Bayesian networks for intention estimation and trajectory prediction," in *Proceedings of the 21st International Conference on Intelligent Transportation Systems (ITSC)*, pp. 1467–1474, Maui, HI, USA, November, 2018.

[19] S. Ammoun and F. Nashashibi, "Real time trajectory prediction for collision risk estimation between vehicles," in *Proceedings of the IEEE 5th International Conference on Intelligent Computer Communication and Processing*, pp. 417–422, Cluj-Napoca, Romania, August, 2009.

[20] V. Punzo, D. J. Formisano, and V. Torrieri, "Nonstationary Kalman filter for estimation of accurate and consistent car-following data," *Transportation Research Record*, vol. 1934, pp. 2–12, 2005.

[21] J. Zhao, V. L. Knoop, and M. Wang, "Microscopic traffic modeling inside intersections: interactions between drivers," *Transportation Science*, vol. 57, no. 1, pp. 135–155, 2023.

[22] Z. Ma, J. Xie, X. Qi, Y. Xu, and J. Sun, "Two-dimensional simulation of turning behavior in potential conflict area of mixed-flow intersections," *Computer-Aided Civil and Infrastructure Engineering*, vol. 32, no. 5, pp. 412–428, 2017.

[23] J. Zhao, V. L. Knoop, and W. Meng, "Two-dimensional vehicular movement modelling at intersections based on optimal control," *Transportation Research Part B: Methodological*, vol. 138, pp. 1–22, 2020.

[24] W. Y. Mergia, D. Eustace, D. Chimba, and M. Qumsiyeh, "Exploring factors contributing to injury severity at freeway merging and diverging locations in Ohio," *Accident Analysis and Prevention*, vol. 55, pp. 202–210, 2013.

[25] H. C. Chin, S. T. Quek, and R. L. Cheu, "Traffic conflicts in expressway merging," *Journal of Transportation Engineering*, vol. 117, no. 6, pp. 633–643, 1991.

[26] C. G. Prevost, A. Desbiens, and E. Gagnon, "Extended Kalman filter for state estimation and trajectory prediction of a moving object detected by an unmanned aerial vehicle," in *Proceedings of the 2007 American Control Conference*, pp. 1805–1810, New York, NY, USA, July, 2007.

[27] G. Rigatos and S. Tzafestas, "Extended Kalman filtering for fuzzy modelling and multi-sensor fusion," *Mathematical and Computer Modelling of Dynamical Systems*, vol. 13, no. 3, pp. 251–266, 2007.

[28] M. T. Abbas, M. A. Jibran, M. Afaq, and W. Song, "An adaptive approach to vehicle trajectory prediction using multimodel Kalman filter," *Transactions on Emerging Telecommunications Technologies*, vol. 31, no. 5, 2019.

[29] E. A. Wan and R. Van Der Merwe, "The unscented Kalman filter for nonlinear estimation," in *Proceedings of the IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control Symposium 00EX373*, pp. 153–158, Lake Louise, Canada, October, 2000.

[30] S. J. Julier and J. K. Uhlmann, "Unscented filtering and nonlinear estimation," *Proceedings of the IEEE*, vol. 92, no. 3, pp. 401–422, 2004.

[31] S. Wang, C. Fernandez, C. Yu, Y. Fan, C. Wen, and S. Daniel-Ioan, "A novel charged state prediction method of the lithium ion battery packs based on the composite equivalent modeling and improved splice Kalman filtering algorithm," *Journal of Power Sources*, vol. 471, Article ID 228450, 2020.

[32] G. Xie, H. Gao, L. Qian, B. Huang, K. Li, and J. Wang, "Vehicle trajectory prediction by integrating physics-and maneuver-based approaches using interactive multiple models," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 7, pp. 5999–6008, 2017.

[33] H. Gao, Y. Qin, C. Hu, Y. Liu, and K. Li, "An interacting multiple model for trajectory prediction of intelligent vehicles in typical road traffic scenario," *IEEE Transactions on Neural Networks and Learning Systems*, 2021.

[34] S. Wang, Y. Fan, S. Jin, P. Takyi-Aninakwa, and C. Fernandez, "Improved anti-noise adaptive long short-term memory neural network modeling for the robust remaining useful life prediction of lithium-ion batteries," *Reliability Engineering and System Safety*, vol. 230, Article ID 108920, 2023.

[35] W. Xiao, L. Zhang, and D. Meng, "Vehicle trajectory prediction based on motion model and maneuver model fusion with interactive multiple models," *SAE International Journal of Advances and Current Practices in Mobility*, vol. 2, no. 6, pp. 3060–3071, 2020.

[36] R. Feng, Z. Li, Q. Wu, and C. Fan, "Association of vehicle object detection and the time-space trajectory matching from aerial videos," *Journal of Transport Information and Safety*, vol. 39, no. 2, pp. 77–69, 2021.

[37] S. Westphal, "Capacity of freeway on-ramps on German motorways," in *Proceedings of the second international symposium on highway capacity*, vol. 2, Sydney, Australia, August, 1994.

[38] C. Thiemann, M. Treiber, and A. Kesting, "Estimating acceleration and lane-changing dynamics from next generation simulation trajectory data," *Transportation Research Record*, vol. 2088, pp. 90–101, 2008.

[39] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Journal of Basic Engineering*, vol. 82, no. 1, pp. 35–45, 1960.

[40] S. Bar, L. Yaakov, X. Rong, and T. Kirubarajan, *Estimation with Applications to Tracking and Navigation: Theory Algorithms and Software*, John Wiley & Sons, New York, NY, USA, 2004.