nature communications

Article

https://doi.org/10.1038/s41467-025-57845-z

# Predicting driving comfort in autonomous vehicles using road information and multi-head attention models

Received: 29 December 2023

Accepted: 4 March 2025

Published online: 19 March 2025

Check for updates

Zhengxian Chen[1], Yuqi Liu[2], Wenjie Ni[1], Han Hai[1], Chaosheng Huang[1] ✉, Boyang Xu[3], Zihan Ling[1], Yang Shen[1], Wenhao Yu[1], Huanan Wang[1] & Jun Li[1]

Driving comfort is a crucial consideration in the automotive industry. In the realm of autonomous driving, comfort has always been a factor that requires continuous improvement. A common approach to improving driving comfort is through the optimization of local path planning. Nevertheless, it is imperative to recognize that macroscopic factors, including traffic flow and road conditions, wield a substantial influence on comfort. For instance, complex traffic scenarios increase the possibility of emergency braking, thereby affecting comfort. Consequently, investigating the intricate interplay between comfort and global path planning becomes essential. This paper introduces a methodology and framework for predicting driving comfort by leveraging road information. The study established a road information-driving comfort dataset and devised prediction models using multi-head attention mechanism. The ensuing discussion elucidates the practical application of the model in path planning through examples and tests. Following the path optimized by the model, the vehicles exhibited a reduction in jerk. This research predicted driving comfort based on road information and integrated it with global path planning, which holds significant implications for autonomous driving navigation systems and provides a valuable reference for related research.

Autonomous driving vehicles have emerged as a focal point in the realm of transportation research, garnering attention for their pivotal role in accident prevention, emissions reduction, and mitigating traffic congestion[1]. A critical element integral to autonomous driving is its motion planner, tasked with orchestrating the vehicle's trajectory and path, enabling it to navigate the road judiciously[2]. The process of path planning can be dichotomized into global path planning and local path planning[3,4]. For instance, when utilizing a map for navigation, the recommended route is derived from global path planning. Key considerations in global path planning encompass factors such as path length, time cost, and other pertinent data, providing drivers with the ability to select the optimal path to traverse.

In recent years, the emphasis on the safety and comfort of autonomous driving has become prominent, leading to the integration of these crucial dimensions into related studies on path planning and trajectory tracking. Safety, recognized as a fundamental prerequisite for autonomous driving, has garnered substantial attention, prompting numerous studies in this domain[5]. On the other hand, the comfort of autonomous driving, being an advanced requirement, also merits attention[6,7]. Several researchers have investigated the comfort of autonomous driving in local path planning and vehicle control[8–15]. This includes integrating acceleration or jerk values as constraints during control to enhance lane-change and braking smoothness, which abrupt changes in acceleration are prone to occur during lane changes and braking, that can affect comfort. Sara Luciani et al.[11] conducted a

[1]School of Vehicle and Mobility, Tsinghua University, 100084 Beijing, China. [2]Graduate School of Natural Sciences, Utrecht University, 3584CS Utrecht, Netherlands. [3]College of Design and Engineering, National University of Singapore, 117575 Singapore, Republic of Singapore.
✉e-mail: huangchaosheng@tsinghua.edu.cn

comfort assessment utilizing metrics such as equivalent acceleration and motion sickness dose value (MSDV). The design of model predictive control (MPC) weight parameters was guided by assumed values of these metrics, aiming to enhance passenger comfort in trajectory tracking. In a study by ref. 16, segmented quadratic Bézier curves, grounded in safe lane-change distances, were employed in local path planning to optimize the ride comfort of autonomous vehicles. Meixin Zhu et al.[12] introduced a speed control model for following vehicles based on reinforcement learning. This model optimized the speed of the following vehicles, enhancing the X-direction ride comfort to some extent. It is noteworthy, however, that its primary emphasis was on safety rather than comfort. Even with substantial advancements in local path planning to enhance the comfort of steering and other behaviors, autonomous driving vehicles still lag behind traditional driving modes in terms of comfort[11]. This phenomenon arises from the inherent contradiction between safety and comfort in autonomous driving vehicles[6], coupled with their decision-making and judgment capabilities that remain inferior to those of human drivers. In this context, focusing solely on local path planning falls short of achieving a sufficient comfort level for autonomous vehicles. Hence, global path planning also holds significant importance. Factors such as the number of traffic lights, intersections, required turns, and even the weather over extended distances can significantly impact the comfort of autonomous driving.

However, comfort considerations are often overlooked in global path planning and selection. To incorporate the comfort factor into global path planning, obtaining relevant information in advance is essential. The fundamental basis for global path planning relies on road information, including the length of a specific path and the degree of road congestion, both of which are easily accessible. Unlike these factors, comfort is difficult to calculate or predict based on intuitive road information. Current researches predominantly center on forecasting human comfort by other information. For instance, ref. 17 delved into the feasibility of predicting passenger seat comfort and discomfort by analyzing human, environmental, and seat characteristics. Their research emphasized quantifying these characteristics to serve as input for predictive models. The findings underscored a correlation between anthropometric variables and interface pressure variables, with body posture further influencing this relationship. The study is more of a review study, with specific models and simulations being more abbreviated. Endeavor, S. Lerspalungsanti et al.[18] devised an artificial neural network-driven method for predicting ride comfort. This method employed design parameters associated with driveline comfort, such as the friction coefficient gradient of the clutch friction pair and the inertial mass and damping of the dual mass flywheel, as inputs for subjective comfort prediction. The article amalgamated an objective measurable dataset, representing the input data, with a subjective evaluation set, representing the target data, to formulate the model. The resultant output was a 10-digit quantity scale denoting NVH (Noise, Vibration, and Harshness). While this method demonstrated efficacy in predicting comfort, it did not incorporate road information into its predictive framework.

Motion sickness is the manifestation of comfort, and with the driver of an autonomous vehicle relieved of the driving task, they transition into a passive role akin to a regular passenger. Consequently, the susceptibility to developing motion sickness (MS) symptoms is notably heightened. Numerous current studies delve into the prediction of motion sickness, with a predominant focus on the swift prediction of motion sickness (MS) from the driver's biosignals, given the substantial variation in MS onset observed among individuals. While the majority of this research[19] has concentrated on virtual reality or simulator environments, ref. 20 introduced a method for predicting MS in a realistic driving setting. This method utilized a practical and easy-to-wear dry EEG device, relying solely on EEG signals for MS prediction. The approach can adapt the planning control of an autonomous vehicle to mitigate motion sickness through an assisted driving system. However, it is very difficult to apply it to global path planning. Although some researchers currently enhanced driving comfort by evaluating motion sickness or predicting motion sickness[21,22] based on partial human body movements[23], and adjusted driving modes accordingly, this approach lacks integration of real road information and did not focus on predicting future variations in driving comfort. Therefore, it is challenging to consider this type of comfort prediction with global path planning, which is based on road information. Hence, it is necessary to explore a method for calculating or predicting comfort based on road information, Autonomous driving is enabling this possibility.

In contrast to autonomous driving, the perception of driving comfort in conventional driving is intricately linked to individual factors[24]. For the same person driving a car at different times, even with consistent driving habits, the level of fatigue and mood can significantly impact their driving comfort. For example, due to distractions, a driver may fail to notice the vehicle ahead, leading to sudden braking. Similarly, because of a bad mood or an argument with passengers, the driver might suddenly accelerate by pressing the gas pedal harder. This makes driving comfort difficult to predict. Therefore, the driving comfort in traditional driving cannot be solely obtained from road information, nor can it be considered exclusively in the route selection process. However, in the realm of autonomous driving, the driving habits and patterns of autonomous vehicles are explicitly defined by code. Autonomous driving, on the other hand, does not involve such subjective factors; it does not experience the same variability as humans do, thereby minimizing the impact of subjective factors. With a predetermined autonomous driving algorithm, it becomes feasible to predict macroscopic driving comfort, including parameters like jerk or acceleration, based on road information. Nevertheless, if some personal data sampling is added, it may be able to predict the comfort of traditional driving mode, but the effect may have a certain gap with automatic driving, the paper has also carried out relevant tests as shown in the result section.

In this work, a prediction method that forecasts the macroscopic comfort indicators of the vehicle based on road information is systematically presented, enabling the selection of a relatively optimal path according to the result. The related process is shown in Fig. 1. This research holds significant importance in enhancing the comfort of autonomous driving, promising a notably improved riding experience for both passengers and drivers. Furthermore, it possesses high commercial potential, with the ability to positively impact major navigation map software in adapting to the era of autonomous driving. Moreover, it introduces a novel perspective to the domain of global path planning for autonomous driving. The primary contributions of this paper are outlined as follows:

1. The problem of driving comfort prediction is systematically proposed. Develops autonomous driving comfort prediction (ADCP) model based on multi-head attention and XGBoost. This model is capable of driving comfort prediction from road information.
2. Gives a method of constructing the road information-driving comfort dataset by real vehicle collection and simulation. The dataset describes the relationship between road information and comfort in the context of an autonomous driving solution.
3. A scientific evaluation index of driving comfort is proposed by combining human experiment and actual measurement.
4. Real vehicle tests were undertaken to validate and ascertain the significance of the model. The results show that the path planning scheme using the ADCP model can effectively reduce the jerk value and improve human comfort.

## Results

### Road information-driving comfort dataset

After simulation and real vehicle data collection, we selected four typical environmental scenarios and four road scenarios for constructing the dataset as shown in Fig. 2.

https://doi.org/10.1038/s41467-025-57845-z

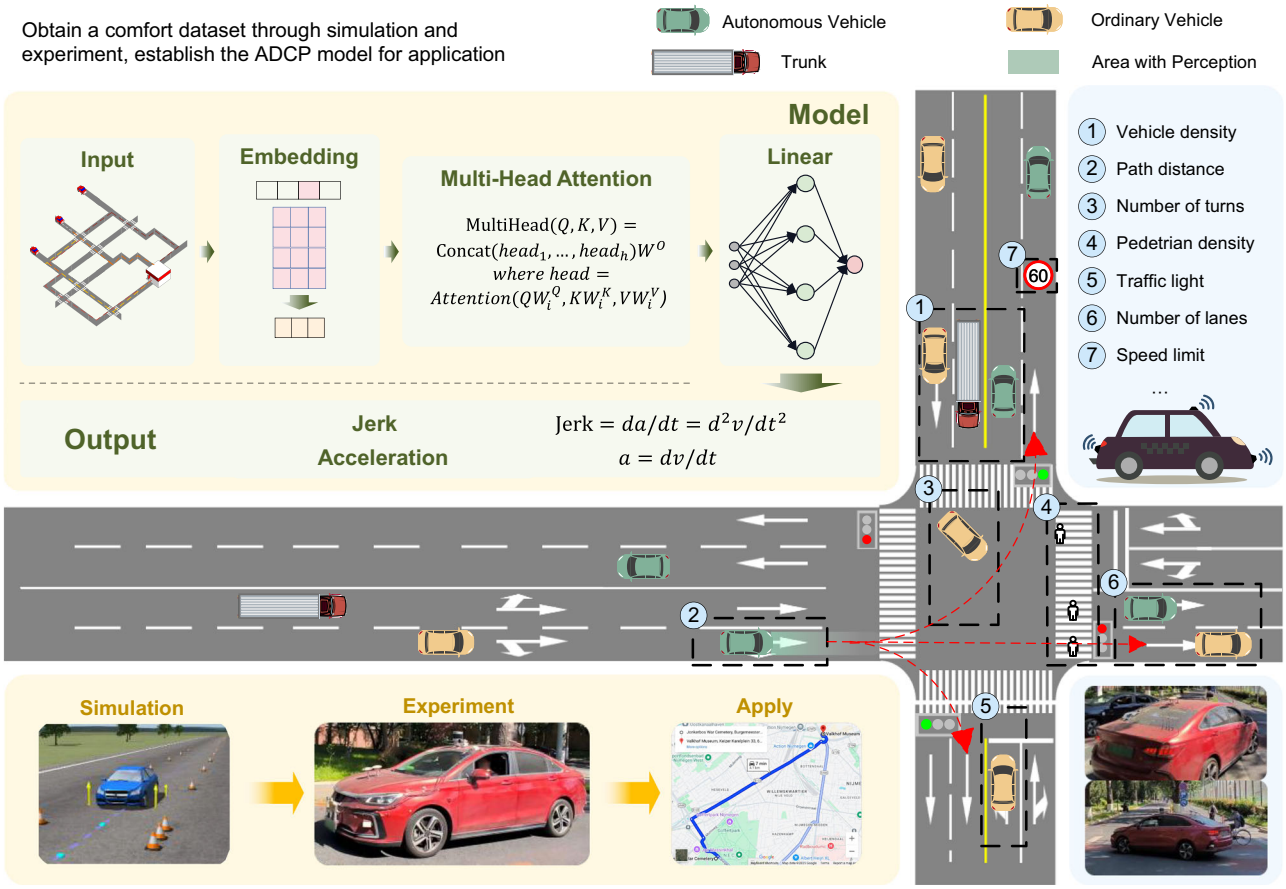


**Fig. 1 | Frame diagram of driving comfort prediction.** Through simulation and real vehicle data collection, various kinds of information were obtained, which were used for training the ADCP model for global path planning.

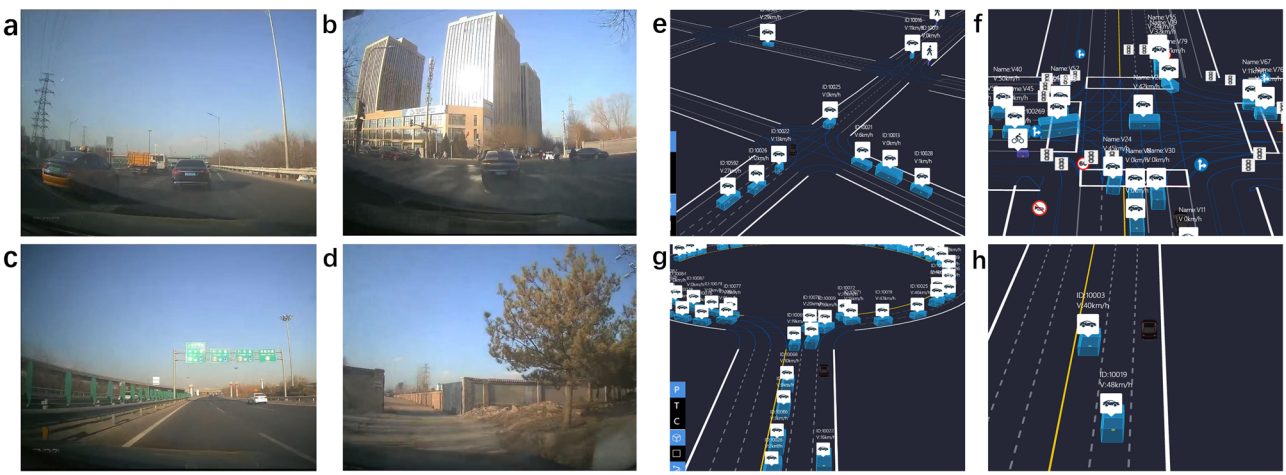


**Fig. 2 | Dataset scenarios of real vehicle collection and simulation. a** Urban fast roads. **b** Urban roads. **c** Highways. **d** Rural ways. **e** Simulated urban roads. **f** Intersections. **g** Traffic circles. **h** Straight roads.

Four environmental scenarios: urban fast roads, urban roads, highways, and rural roads were covered in the dataset. The urban fast roads are shown in Fig. 2a, representing a fast road in the city without traffic lights, but always with heavy traffic flow and numerous entries and exits. Conversely, the rural roads, are characterized by narrow lanes and generally poor road conditions. These scenarios essentially cover different driving conditions and driving routes. The traffic flow information was obtained in real-time through the API of commercial mapping software, and the information includes real-time road traffic congestion levels, semantic descriptions of congestion on specific road sections, average travel speeds, congestion distances, and more. Relevant scenarios are shown in Fig. 2a–d.

Four micro-view road scenarios, straight roads, intersections, traffic circles, and micro-urban roads, were simulated. Vehicle data was collected at intervals of 0.016 ms by configuring various traffic flow scenarios, and the dataset was constructed accordingly. The simulation map and environment are depicted in Fig. 2e–h.

**Table 1 | Dataset structure example**

| Features | Path distance(m) | Entrance (same side) | Entrance (opposite) | Turns | Traffic lights | Vehicle density | Left turns | Turning degree | Straight intersection | Lanes | Speed limit | Vehicle density | Weather |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1300 | 1 | 0 | 2 | 2 | 10 | 2 | 150 | 2 | 3 | 60 | 30 | 0.2 |
| Labels | Jerk x | | | Jerk y | | | | Acceleration x | | | Acceleration y | | |
| | 0.35151313991958205 | | | 0.1841569799715372 | | | | 0.3501355696717623 | | | 0.233740259764202 | | |

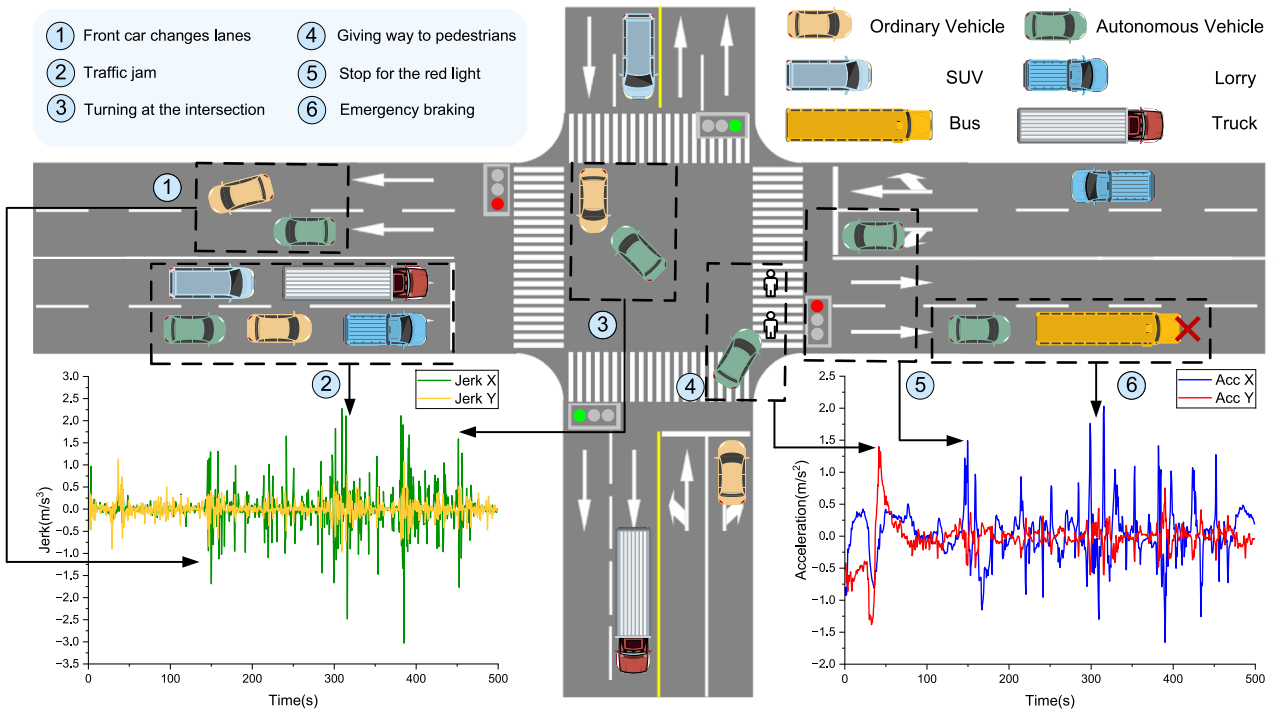


**Fig. 3 | Uncomfortable driving phenomenon.** Six typical uncomfortable behaviors that may lead to rapid increase or decrease of jerk and acceleration are presented: front car changes lanes, traffic jam, turning at the intersection, giving way to pedestrians, stop for the red light, emergency braking and starting.

**Table 2 | MAE result**

| Model | Real vehicle dataset | | Simulation dataset | |
|---|---|---|---|---|
| | Jerk | Acc | Jerk | Acc |
| BP | 0.1191 | 0.1345 | 0.0925 | 0.1339 |
| XGBoost | 0.0767 | 0.1060 | 0.0503 | 0.0664 |
| Attention | 0.0908 | 0.1297 | 0.0619 | 0.0696 |

In total, 274 sets of real vehicle data and 252 sets of simulation data are included in the dataset. Initial data is illustrated in Supplementary Table 1. The key information of the dataset is shown in Table 1. The preprocessing methods are presented in the Methods section.

The collected data has a strong correlation with the characterization data of comfort (jerk and acceleration), as shown in Fig. 3.

The model input takes the number of traffic lights and the number of intersections as examples. When the vehicle encounters more traffic lights on the path, the driver must frequently shift down and accelerate. This will cause significant changes in the acceleration of the vehicle multiple times. Every stop and restart means a change from zero acceleration to positive acceleration, and then to zero acceleration again. This frequent change has led to an essential correlation between the number of traffic lights, acceleration, and jerk. At every crossroad, there may have the phenomenon of vehicles meeting, and the amount of entry on the road or traffic flows merging will prominently increase. Therefore, drivers need to slow down to observe the traffic situation. Deceleration and subsequent acceleration can cause changes in jerk and speed. So more crossroads mean that drivers need to perform acceleration and deceleration operations more continually, which directly affects the frequent changes in acceleration and jerk, bringing an uncomfortable feeling.

## Prediction

The model was trained and tested using both the real vehicle dataset and the simulation dataset, yielding comparable results. The models were compared and evaluated using MAE (mean absolute error), as shown in Table 2 below.

The effect of the model trained with the real vehicle collected dataset is shown in Fig. 4.

Figure 4 illustrates a comparison between the true values and the predicted values for 42 test set data points. It can be observed that the real vehicle collected data exhibits some randomness, but most of the predicted values are concentrated within a small range near the true values, which effectively reflects the trend of comfort. Among these, both the Attention and XGBoost models have lower MAE values than the BP model. XGBoost performs better in regions close to the average value, but its prediction performance significantly decreases under relatively extreme conditions, with predicted values concentrated within a smaller range. In contrast, the Attention model performs relatively better in this regard, being able to predict some extreme cases, and to some extent, responding to the randomness of the actual driving environment. The description of the error and the values of lateral jerk is provided in Supplementary Note 1.

https://doi.org/10.1038/s41467-025-57845-z

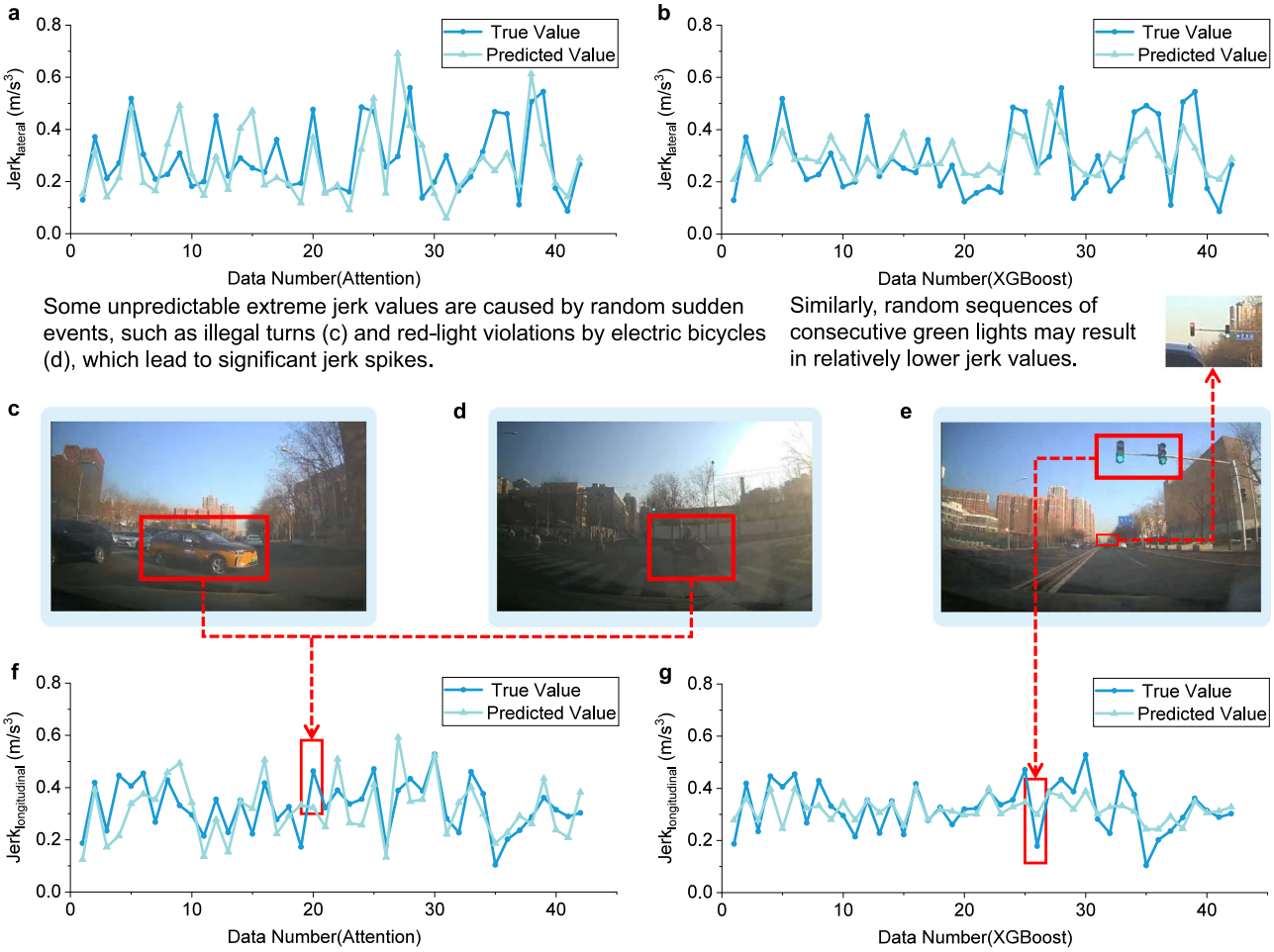


**Fig. 4 | Result of models trained and tested by vehicle collection. a** Comparison of predicted and target values of lateral jerk y (Attention). **b** Comparison of predicted and target values of lateral jerk (XGBoost). **c** Emergency situation 1: A sudden U-turn. **d** Emergency situation 2: Run a red light. **e** Continuous green light. **f** Comparison of predicted and target values of the longitudinal jerk (Attention). **g** Comparison of predicted and target values of the longitudinal jerk (XGBoost).

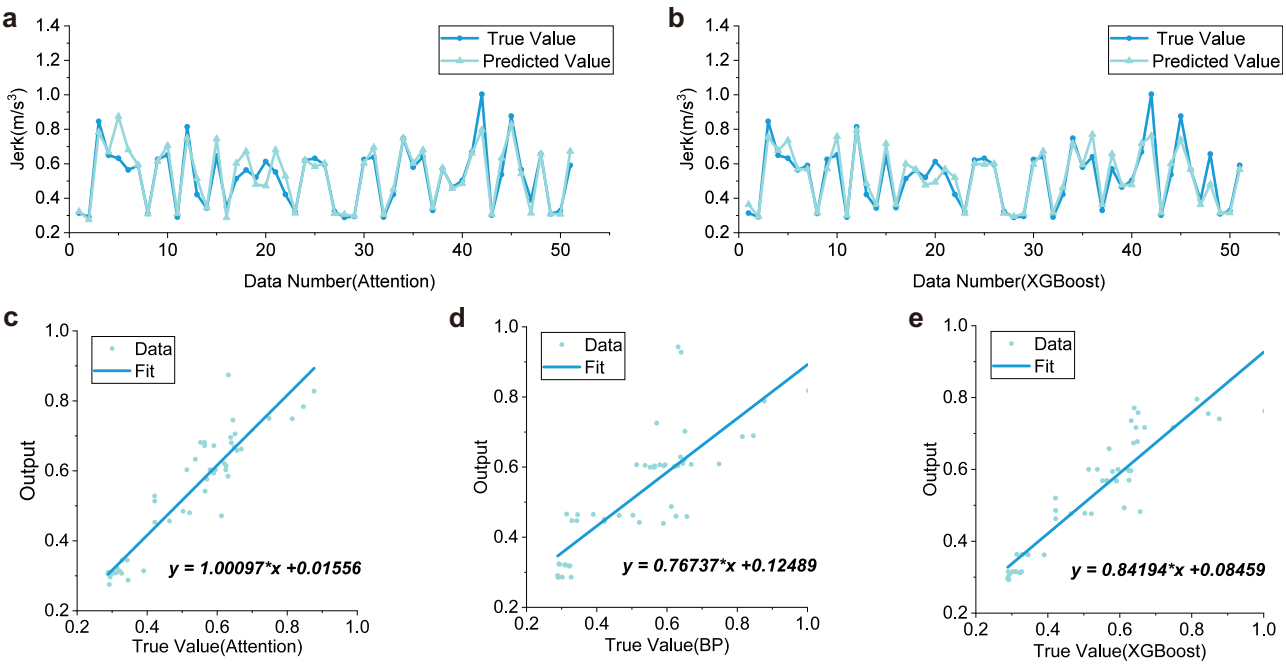


**Fig. 5 | Result of models trained and tested by simulation. a** Comparison of predicted and target values of Jerk (Attention). **b** Comparison of predicted and target values of Jerk (XGBoost). **c** Regression plots of Jerk (Attention). **d** Regression plots of Jerk (BP). **e** Regression plots of Jerk (XGBoost).

The model performance trained on the simulation dataset is shown in Fig. 5:

Overall regression results demonstrate that the regression for jerk and acceleration by Attention and XGBoost appears relatively sound, which proves the effectiveness of both models in predicting driving comfort. The overall prediction accuracy of the simulation dataset is better than that of the real vehicle dataset. This is because both the training and test sets of the simulation dataset are generated through simulation, which, to some extent, represents a more ideal traffic flow. However, driving comfort has a high degree of randomness, as unexpected phenomena on the road—such as accidents, drunk driving behavior from other vehicles, or sudden braking caused by distractions—cannot be accurately simulated in the simulation. As a result, the error in the simulation model is relatively smaller, as it excludes the interference of many random behaviors.

Predicting the absolute value of driving comfort is quite challenging. This situation is similar to some related problems, such as predicting the trajectory of adjacent vehicles, where there is high uncertainty due to the inability to predict the behavior of other vehicles. However, judging the relative comfort of choosing driving paths based on predicted values, or predicting the range of driving comfort for future paths and using this information for global path planning, can help select a more comfortable path to some extent. This has significant implications for improving driving comfort.

We conducted a path selection comparison using random inputs, comparing the model trained on the simulation dataset with the one trained on the real vehicle dataset. Most of the results were similar. When selecting different paths, if the difference in predicted values between two routes is greater than 20%, 94% of the model choices are the same. When the predicted value difference is less than 20%, 61% of the choices are the same. This essentially indicates that the actual driving comfort of the two selected routes at this time is quite similar, and at this point, randomness becomes a key factor in determining the driving comfort of these two segments. For example, a single-lane change could lead to significantly more discomfort.

### Path selection

The relationship between comfort and both global and local path planning is mentioned in the introduction section. Braking, lane changing, and steering affect driving comfort, as reflected in the abrupt changes in lateral and longitudinal acceleration, as shown in Fig. 3. Global path planning can, to some extent, reduce the probability of braking, lane changing, and steering, as illustrated in Fig. 6. Reasonable path planning can reduce the frequency of steering and lane changes, as well as avoid roads that may lead to discomfort, thereby reducing the frequency of braking and acceleration. This requires integrating comfort prediction to achieve. Regarding path planning and selection, this paper presents cases and results as follows:

This work selects a real scenario through commercial mapping software and performs path selection considering comfort based on the models, as shown in Supplementary Fig. 1.

Feeding the data of three paths into the model can get the results, as shown in Supplementary Table 2.

The path A (Predicted Jerk = 0.44) has the lowest traffic flow density, resulting in fewer instances of merging in front of vehicles or similar scenarios, leading to a lower number of emergency brakes and, consequently, the lowest jerk value. The predicted results align with the actual situation, choosing path A is deemed optimal.

The specific decision on which path to take must be considered in conjunction with the particular situation, and further investigation into the impact of jerk and acc values on the human body will be a focal point of future research. While the prediction model aligns with the true value, the disparity in comfort is not particularly pronounced over short distances when comparing similar path data for different routes. However, over longer distances, the difference becomes more evident. For instance, in the comparison of a highway and an ordinary road with the same destination and departure, the predicted jerk of the highway is much lower than the ordinary road.

### Real car test

The change of jerk value before and after loading the model into Global Path Planning is shown in Fig. 7a and Supplementary Fig. 2.

Following data processing, when the autonomous vehicles utilize the path optimized by the model, the average jerk is reduced by ~15% in the X-direction and around 9% in the Y direction. This indicates a notable improvement in the average comfort level facilitated by the model. Notably, areas with heavy traffic demonstrated relatively poor performance. Subsequent research endeavors will prioritize enhancing the model's generalizability.

From the results of the human experiment scoring, it can be clearly seen that after deploying the ADCP model, the scores of the relevant testers have improved to a certain extent. Using the comfort evaluation method proposed in this article, the comprehensive evaluation results also made some progress, with a total score increase of about 13%.

Through testing in manual driving mode, we found that the ADCP model can also be applied to traditional driving, the human experiment result is shown in Fig. 7d. But there are certain differences for different people, which is due to driving habits. If the driving habits of individual drivers differ greatly from the autonomous driving strategy used in training, the path planning strategy based on the ADCP model will fail to select the most comfortable path that matches it. In the current path planning of commercial map apps, there are similar situations. For example, if a "shortest time" path it plans requires passing through a long highway section, and some drivers are accustomed to driving at the standard speed on the highway section (as novices or for safety reasons). Therefore, the "shortest time" path may not be the fastest. In summary, the ADCP model has shown good performance when trained using its own driving data for a specific autonomous driving solution. However, if used in other autonomous driving solutions or manual driving, it also has the same effect, but it is not perfectly suitable for all.

## Discussion

This paper systematically proposes the problem of driving comfort prediction. The study was initiated by constructing a framework for predicting driving comfort based on road information. A comprehensive dataset consisting of both real vehicle collection and simulation data, has been provided to fulfill the requirements for prediction and training. This paper summarizes the relevant evaluation indexes of driving comfort, and puts forward the evaluation method of comfort. Moreover, a comfort prediction model based on multi-head attention is provided. The results demonstrate that the model's efficacy meets the fundamental prediction requirements. Following this, the paper elucidates the application of the model in path selection and planning. The validity and reasonableness of the theory and method were verified through real vehicle tests, the autonomous vehicles exhibited an approximate 15% reduction in the mean longitudinal jerk and a 9% reduction in the mean lateral jerk, and with a total score increase of about 13%. Given the pioneering nature of this study, the richness of its content has led to a selective focus, with certain simpler and fundamental elements not elaborated upon in detail due to space constraints. This model holds significant implications for path planning and autonomous driving comfort, achieving the integration of autonomous driving comfort and global path planning.

This paper focuses on the problem of comfort prediction, while providing a well-performing model and baseline for this issue. This paper has not conducted a more detailed study of the variable weight values in path planning or used more advanced optimization algorithms for iterating the weight values in this paper. Future research will focus on the weight optimization of comfort path planning under

https://doi.org/10.1038/s41467-025-57845-z

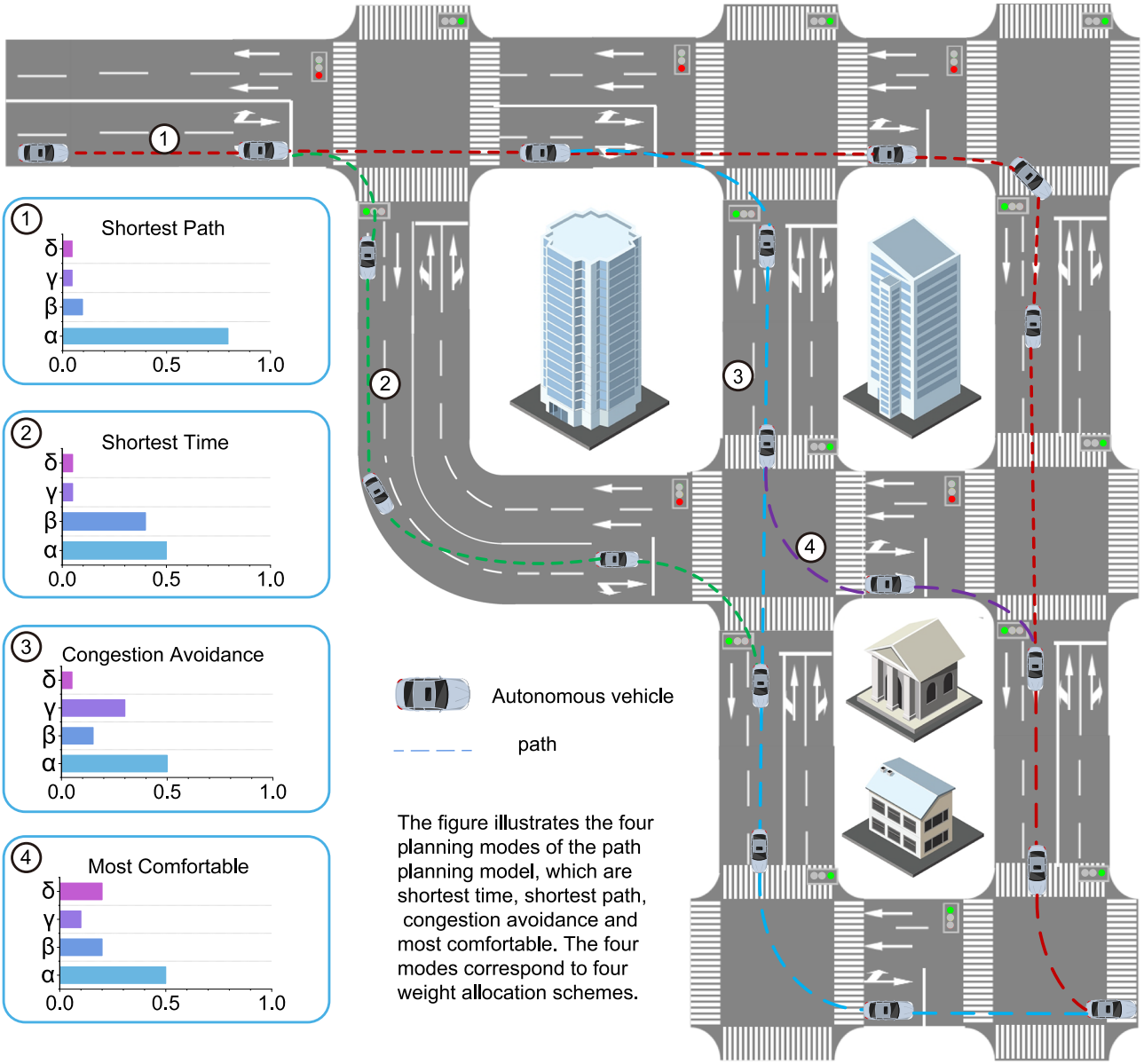


**Fig. 6 | Path-planning method diagram.** Weight allocation in four path planning methods (shortest path, shortest time, congestion avoidance, most comfortable).

different driving scenarios, and will also be directed towards enhancing the overall effectiveness of the model and lowering the value of jerk through the methods, with the hope of advancing progress in the relevant field.

## Methods

### Dataset construction

Integrating global path planning with comfort considerations, and predicting comfort based on road information has emerged as a promising approach. The development of road information-comfort prediction datasets holds paramount importance in this regard. Presently, datasets pertaining to road information can be categorized into three main types. Firstly, there are macroscopic traffic flow information datasets coupled with road databases[25]. Secondly, microscopic vehicle datasets are available[26]. Lastly, there are vehicle-road cooperative datasets that amalgamate information from both vehicles and roads. Road databases primarily encompass road network data. Common examples of such databases include Nature Earth, VMap0[27–29], and the GRIP database[30]. On the other hand, traffic flow information datasets provide insights into the traffic conditions across specific sections or regions. For instance, the Xi'an traffic dataset from the DiDi platform offers a comprehensive view of historical and real-time road conditions, encompassing parameters such as traffic density, road dimensions, lane count, traffic direction, speed limits, functional classification, and road hierarchy. Notably, these datasets do not capture the individual travel trajectories of vehicles. Micro-vehicle datasets, meanwhile, capture the driving trajectories of numerous individual vehicles. Examples include the Beijing T-driver dataset and the New York TLC dataset. Both traffic flow datasets and individual vehicle datasets present limitations when applied in scenarios that necessitate the integration of traffic flow with individual vehicles. This limitation is somewhat mitigated by vehicle-road cooperative datasets, such as the NGSIM dataset[31], developed under the auspices of the Next Generation Simulation project. In this initiative, researchers post-processed raw video data to derive trajectory information for each vehicle within the traffic flow, thereby combining traffic flow data with individual vehicle data. However, it's essential to clarify that this dataset contains information from regular vehicles rather than autonomous vehicles. Consequently, it cannot be utilized in the context of this study.

https://doi.org/10.1038/s41467-025-57845-z

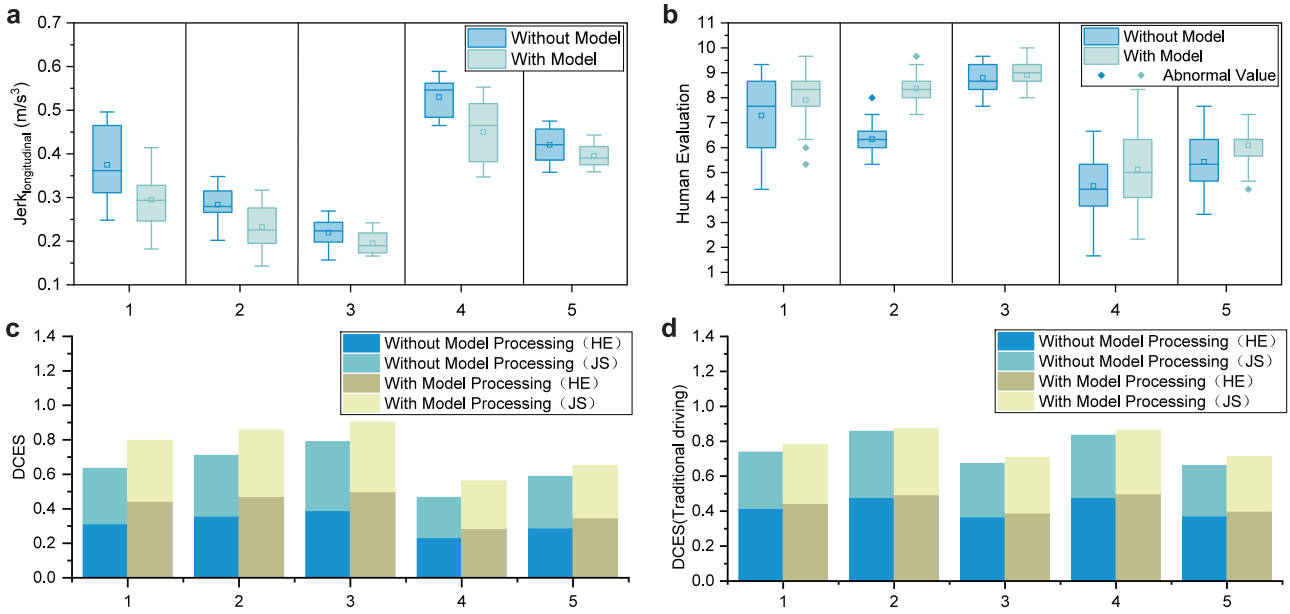


**Fig. 7 | Result of real car test. a** Longitudinal jerk. **b** Human evaluation. **c** DCES. **d** DCES of traditional driving.

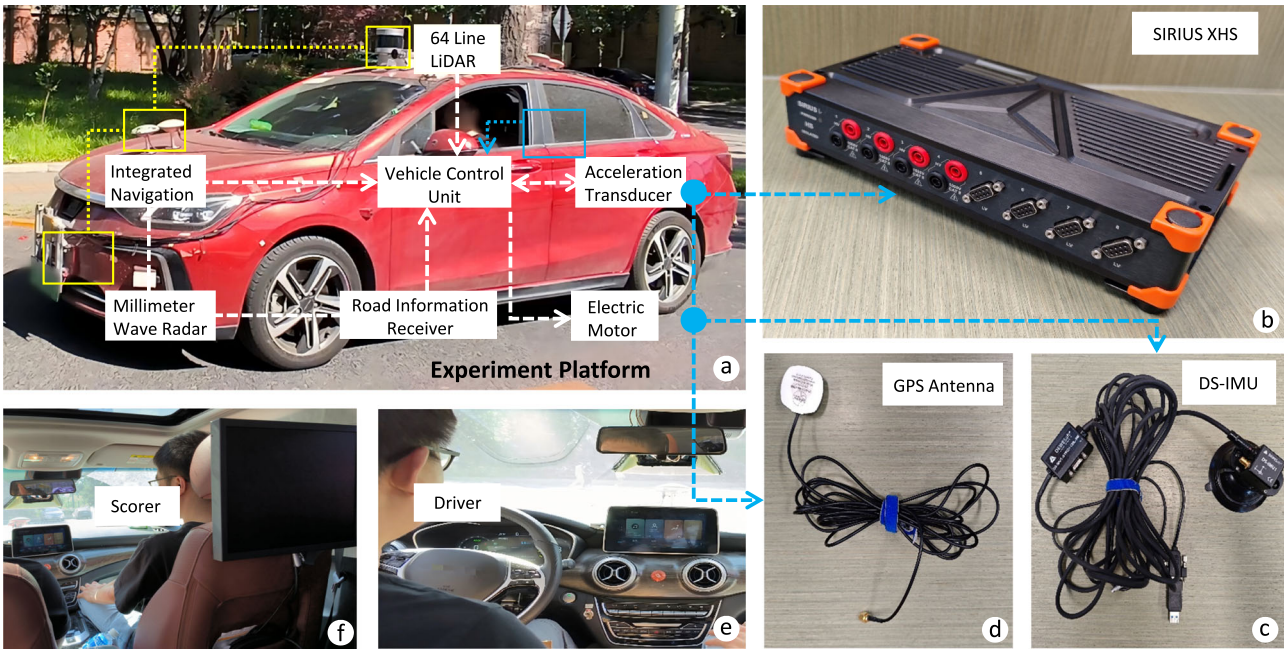


**Fig. 8 | Experiment platform. a** Vehicle structure. **b** SIRIUS XHS. **c** DS-IMU. **d** GPS antenna. **e** Driver (Security). **f** Scorer.

Hence, this paper introduces a novel method for constructing datasets for the ADCP model.

The dataset provides diverse data, including data collected from actual measurements, as well as data obtained through simulations. Both sources contribute extensive driving comfort data across multiple scenarios.

The real vehicle data was collected using Dewesoft's SIRIUS XHS and DS-IMU systems, the relevant acquisition vehicles and equipment are shown in Fig. 8.

In the simulations, traffic flow density was configured across three dimensions: vehicles, pedestrians, and bicycles. Additionally, three distinct vehicle types with varying volumes can be specified. Based on actual road conditions, vehicle types were quantified with a ratio of 7:2:1 for vehicles, buses, and engineering vehicles, respectively. Throughout the simulation, lighting remained constant, while weather and other variables were manipulated separately to conduct simulations using a specified autonomous driving algorithm. The study also considered road roughness as an input; However, due to challenges in obtaining precise road roughness values for the future path in real environments, it was approximated using road grade. In this paper's simulation, the pavement roughness for short-distance urban roads was assumed to be relatively constant.

The speed information obtained from the 51sim-one simulation was in the global coordinate system and needs to be converted.

The above data were processed according to the theory of vehicle kinematics and dynamics, the global coordinate system is transformed

into the vehicle coordinate system as follows:

$$\begin{cases} V_x = \dot{X}\cos\psi + \dot{Y}\sin\psi \\ V_y = \dot{Y}\cos\psi - \dot{X}\sin\psi \end{cases} \tag{1}$$

Where $V_x$ indicates the x-directional velocity. $V_y$ denotes y-directional velocity, $\dot{X}$ and $\dot{Y}$ denote the derivative of $X$ and $Y$, $\psi$ denotes the heading angle.

Calculate the vehicle acceleration based on vehicle dynamics:

$$\begin{cases} a_x = \dot{V}_x - V_y\omega_z \\ a_y = \dot{V}_y + V_x\omega_z \end{cases} \tag{2}$$

Where $a_x$ indicates x-directional acceleration,$a_y$ is y-directional acceleration, $\dot{V}_x$ and$\dot{V}_y$ are derivative of $V_x$ and $V_y$, $\omega_z$ indicates yaw rate.

At the same time, part of the real-time traffic information obtained by commercial map software is calculated as a road traffic congestion index to unify with the dimension of simulation data:

$$P = 25\frac{\sum_0^L S\Delta L}{L} - 0.2V + C_t \tag{3}$$

In the equation, $P$ represents the traffic congestion index, ranging from 0 to 100, consistent with the simulation dataset. $L$ is the path length, and $S$ is the average congestion level of the segment, which includes four levels: 1–4. $V$ is the average traffic speed of the segment, measured in kilometers per hour. $C_t$ represents the congestion trend, where if the congestion has eased compared to 10 min ago, $C_t$ is defined as −5; if it has worsened, $C_t$ is defined as +5; and if it remains the same, $C_t$ is defined as 0.

The autonomous driving system adopts relatively traditional and universal methods for acceleration and jerk control strategies. Under the condition of meeting safety requirements, the system includes functions for comfortable braking and acceleration, where jerk and acceleration (or deceleration) thresholds are set. For example, the jerk threshold is set to $1m/s^{-3}$. However, in emergency situations, the comfort braking function will not be activated, and the vehicle will directly brake using the ideal deceleration.

### Driving comfort evaluation

In a moving vehicle, we may experience sudden impacts, which are actually the result of changes in vehicle speed or acceleration. Frequent impacts can lead to discomfort and even endanger the safety of drivers and passengers. Therefore, it is necessary to control these impacts in an acceptable way. However, which kind of impact is acceptable and how to effectively quantify it are issues that need to be considered. At present, there is no clear method for measuring driving comfort. Some researchers use the car ride comfort index (CI) to represent riding comfort, but this is different from driving comfort. Riding comfort does not entirely depend on driving, but largely on the design of seats, suspension, and other factors. More researchers[32,33] are using jerk (a derivative of acceleration) to evaluate driving comfort. SAE summarized early research and demonstrated two methods for quantifying jerk[34], namely (a) peak jerk and (b) root mean square of jerk. For method (a), Ge A[35] believes that the upper limit of jerk that humans can tolerate is 10 m/s³, and exceeding this value will not cause discomfort. However, the frequency is related to human physiological sensation, so using only the peak jerk does not seem comprehensive. Shouren Z[36] believes that at frequencies below 3 Hz, the acceptable upper limit for jerk is 2.6 g/s, which is 25.5 m/s³. There are also some doubts about the method (b), as the guidelines provided by SAE suggest that using the root mean square of a jerk as a measure of comfort may overlook the peak jerk, which cannot effectively evaluate comfort. Afterward, SAE proposed a method to evaluate the impact of acceleration on ride comfort by combining the peak and root mean square values of jerk, which is represented as:

$$\mathrm{AJV} = 0.004\mathrm{hpj}^2 + \mathrm{rmsj} \tag{4}$$

$$\mathrm{hpj} = \max(\mathrm{jerk(t)}) \tag{5}$$

$$\mathrm{rmsj} = \sqrt{\frac{1}{T}\int_0^T \mathrm{jerk}^2(t)\,dt} \tag{6}$$

Among them, hpj represents the highest peak value of jerk, and rmsj represents the root mean square value of jerk. This formula calculates AJV by taking the square of the peak jerk and adding it to the root mean square value of the jerk, and then multiplying it by an empirical coefficient of 0.004. It can be said that the method provided by SAE is a mixed model that integrates multiple factors, and it has a certain representative significance. However, ref. 33 investigated the relationship between vehicle motion attributes and passenger comfort through experiments. The experimental results show that an increase in jerk amplitude is positively correlated with an increase in discomfort level, and this effect is influenced by the direction of motion. Compared to forward motion, backward motion has lower comfort, while lateral motion has lower comfort. In addition, they believe that higher jerks (shorter duration pulses) are considered more compatible. This conclusion is very novel and differs from the methods provided by SAE and previous studies. It takes into account the impact of frequency and believes that a larger jerk in a shorter period of time does not necessarily mean more uncomfortable. In summary, a standardized system for evaluating Jerk values and comfort has not yet been established. The simple use of measured values cannot indicate the magnitude of driving comfort. Therefore, this article integrates previous research and proposes a driving comfort evaluation method that combines measured values with human experiments.

Among the measured parameters, we use acceleration and jerk to evaluate driving comfort. The focus of this article is on the overall path planning of vehicles, while the smoothness evaluation method provided by ISO 2631-1[37] focuses on the riding comfort of a single vehicle. The comfort index (CI) value largely depends on the design of the car seat and suspension, which is not the content of this article. Therefore, the acceleration evaluation method used in this article is the vehicle body acceleration, which is commonly used in enterprises, rather than the acceleration on the seat (CI), this needs to be clearly pointed out. According to the evaluation of vehicle body acceleration used by relevant enterprises in vehicle production, there will be no weighted summary in the X, Y, and Z directions. Vehicle acceleration can also be used for qualitative evaluation of comfort on different paths. The root mean square value CA and jerk value CJ of acceleration are calculated using the formulas:

$$\mathrm{CA} = \sqrt{\frac{\sum_{i=1}^N a_i^2 \cdot \Delta t_i}{\sum_{i=1}^N \Delta t_i}} \tag{7}$$

$$\mathrm{CJ} = \sqrt{\frac{\sum_{i=1}^N (a_i - a_{i+1})^2/\Delta t_i}{\sum_{i=1}^N \Delta t_i}} \tag{8}$$

In addition to using objective parameter evaluations, evaluating comfort should also consider the subjective feelings of passengers. Subjective feelings refer to the personal feelings and emotional reactions experienced by passengers during the driving process, which often reflect comfort more accurately than objective parameters.

Through subjective evaluation, testers can understand passengers' true feelings and expectations for driving comfort. The comprehensive evaluation of objective parameters and subjective feelings can provide a more comprehensive understanding of comfort. The subjective evaluation method is shown in Supplementary Fig. 3, the relevant testing score according to Supplementary Table 3.

As shown in Supplementary Fig. 3, the human evaluation score (HES) is weighted and added to the measurement values after adjusting the importance parameters of the three evaluations to obtain the driving comfort evaluation score (DCES). The jerk value and the score of the human body have been standardized and normalized through normal distribution to ensure that the obtained trend and correlation are reasonable. The calculation formulas are as follows:

$$\mathrm{Jerk_{nd}} = \frac{(\mathrm{Jerk} - \mu)}{\sigma} \tag{9}$$

$$\mathrm{Jerk_{sd}} = \frac{\mathrm{Jerk_{nd}} - MIN(\mathrm{Jerk_{nd}})}{MAX(\mathrm{Jerk_{nd}}) - MIN(\mathrm{Jerk_{nd}})} \tag{10}$$

$$\mathrm{HES} = k_1\mathrm{BE} + k_2\mathrm{AE} + k_3 SE \tag{11}$$

$$\mathrm{HES_{nd}} = \frac{(\mathrm{HES} - \mu)}{\sigma}\Big) \tag{12}$$

$$\mathrm{HES}_{sd} = \frac{\mathrm{HES_{nd}} - MIN(\mathrm{HES_{nd}})}{MAX(\mathrm{HES_{nd}}) - MIN(\mathrm{HES_{nd}})} \tag{13}$$

$$\mathrm{DCES} = k_4(1 - \mathrm{Jerk_{sd}}) + k_5\mathrm{HES_{sd}} \tag{14}$$

In Formula 9, $Jerk_{nd}$ is the standardized jerk value, μ is the target expectation (mean),σ is the standard deviation of jerk, $Jerk_{sd}$ is the normalized final score, BE is the brake evaluation score, AE is the acceleration evaluation score, SE is the shake evaluation score, $k_1, k_2, k_3, k_4, k_5$ are the importance correction coefficients. For the values of these parameters, we refer to the weight values of each direction in ISO 2631 for ride comfort evaluation. Brake and Acc are more sensitive to X-direction, which is also the reason why related driving comfort studies[24,34,38,39] mainly focus on longitude direction. Therefore, we choose the positive and negative scores in the X-direction, and each of them has the same weight as the shake score. The three scores are averaged to get the final score. When the sample size is small, the jerk value can also be directly put into Formula 14 without pretreatment, while the HES needs to be normalized.

### Driving comfort prediction model

Due to the complexity of predicting driving comfort and the fact that it is a topic without reference from other methodological materials, it is necessary to try different types of machine learning methods and then comprehensively consider the effects of multiple methods, and then select the best prediction method from them. Therefore, from two representative machine learning methods, neural network[40] and binary tree[41], we select the multi-head Self-Attention mechanism and XGBoost for comfort prediction according to the matching degree with the data and the relevance of the research content. In addition, we will use the BP[42] as a baseline to compare the performance of the Attention and XGBoost mentioned above.

XGBoost[43–46], a highly efficient boosting algorithm, demonstrates outstanding performance in regression tasks. It is comprised of numerous weak classifiers, implemented using classification and regression trees (CART). These trees can be viewed as decision rules derived from multiple features, enabling them to capture cross information. Residuals serve as the learning objective for each decision maker, and XGBoost grows a tree by continuously adding trees and performing feature splitting. Each time a tree is added, it actually learns a new function $f(x)$ to fit the last predicted residual, which can make the loss function more accurate and better fit the data. In addition, XGBoost incorporates regularization to avoid overfitting of trees and improve the model's generalization ability. It also adopts a block structure and parallel computing to improve training speed and efficiency, enabling it to handle large-scale datasets. The objective function for the overall model is illustrated in formulas as follows:

$$L = \sum_{i=1}^{n} l(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k) \tag{15}$$

$$L = \sum_{i=1}^{n} l\left(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)\right) + \sum_{j=1}^{t} \Omega(f_j) \tag{16}$$

$$l\left(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)\right) = l\left(y_i, \hat{y}_i^{(t-1)}\right) + g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i) \tag{17}$$

As shown in Formula 15, $l(y_i, \hat{y}_i)$ represents the loss function of an individual sample, while $\Omega(f_k)$ denotes the complexity of the ensemble of all trees, serving as a regularization term. For each weak classifier, the preceding classifier results are known, thus enabling the expansion of the overall loss function as depicted in Formula 16. In this formula, $\hat{y}_i^{(t-1)}$ represents the known predicted results of all trees before the t-th tree, and $f_t$ represents the predicted residuals. Expanding the loss function further leads to Formula 17, in which $l(y_i, \hat{y}_i^{(t-1)})$ represents a constant value, $g_i$ represents the first derivative of the loss function, and $h_i$ represents the second derivative. The complexity of the tree is characterized by two components: the number of nodes in the tree and the L2 norm, as illustrated in Formula 18. Finally, by summing the loss functions of all classifiers and incorporating the regularization term, the final loss function is obtained, as depicted in Formula 19.

$$\Omega(f_t) = \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{T} w_j^2 \tag{18}$$

$$L = \sum_{j=1}^{T} \left[ \left( \sum_{i \in I_j} g_i \right) w_j + \frac{1}{2} \left( \sum\nolimits_{i \in I_j} h_i + \lambda \right) w_j^2 \right] + \gamma T \tag{19}$$

The design of XGBoost demonstrates powerful performance in solving practical problems, especially when dealing with high-dimensional data and missing values. It is worth noting that although XGBoost performs well in many tasks, it is not omnipotent. In certain scenarios, models based on attention mechanisms may have advantages. Therefore, when choosing a model, we have comprehensively considered the specific problem and data characteristics.

The attention mechanism has recently gained popularity as an algorithm for learning information from feature crosses. It is necessary to effectively learn the cross information between multiple features in the comfort prediction of global path planning. For instance, the cross-information between traffic flow density and weather provides a more insightful characterization of the situation compared to solely considering traffic flow density or weather individually. Therefore, we attempt to use self-attention to enrich each feature by utilizing the

cross information between this feature and other features.

$$f, F_{(Weather, Length, \ldots)} = \left(V_f^Q, V_f^K, V_f^V\right) \tag{20}$$

$$V_f^Q = W_f^Q f \tag{21}$$

$$V_f^K = W_f^K f \tag{22}$$

$$V_f^V = W_f^V f \tag{23}$$

$$\text{Attention}\left(f, F_{(Weather, Length, \ldots)}\right) = \frac{\text{Softmax}\left(V_f^Q V_F^{K^T}\right)}{\sqrt{d_K}} \tag{24}$$

$$O_f = \sum_F V_f^V \text{Attention}\left(f, F_{(Weather, Length, \ldots)}\right) \tag{25}$$

Self-attention, each feature $f, F_{(Weather, Length, \ldots)}$ undergoes a linear transformation to obtain three hidden vectors specific to itself, as shown in Formula 20. These are called the query vector $V_f^Q$, the key vector $V_f^K$, and the value vector $V_f^V$, as shown in Formulas 21, 22, and 23. Where $f$ represents the feature, and $W_f^Q$, $W_f^K$, and $W_f^V$ are the corresponding parameter weight matrices. Then, as shown in Formula 24, the query vector $V_f^Q$ of a feature is dot-multiplied with the transposed key matrix $V_F^{K^T}$ (where $V_F^K$ represents the matrix composed of all $V_f^K$). This results in the attention score vector of the feature for all features. The Softmax function is a normalization function used to eliminate differences in magnitudes. $\sqrt{d_K}$ is a scaling factor, where $d_K$ represents the vector dimension of $V_f^K$, used to scale the dot product result to prevent excessively large values from reducing the effectiveness of Softmax normalization. Finally, as shown in Formula 25, the value vector $V_f^V$ of all features is dot-multiplied with the attention score $Attention(f, F_{(Weather, Length, \ldots)})$ to obtain the value of the feature influenced by all other features $O_f$.

On the other hand, the influence between different features can be interpreted differently in various domains. For example, in the domain of road traffic conditions, the interaction between traffic density and weather can reflect road congestion; in the domain of safe driving, the interaction between these two features reflects the likelihood of accidents. Therefore, based on the attention mechanism, we use the multi-head self-attention mechanism to simulate different domains. The description of multi-head attention is outlined below[47]:

$$\text{where head}_{f, F_{(Weather, Length, \ldots)}} = \text{Attention}\left(f, F_{(Weather, Length, \ldots)}\right) \tag{26}$$

$$\text{MultiHead}\left(f, F_{(Weather, Length, \ldots)}\right) = W^O \text{Concat}\left(\text{head}_{f, F_{(Weather, Length, \ldots)}}\right) \tag{27}$$

The interpretation of different features $f$ in different domains requires representation as multiple heads, where self-attention is computed separately for each head. This is reflected in Formula 21, 22, and 23 by the subscript $f$ and in Formula 26 by $f, F_{(Weather, Length, \ldots)}$, resulting in the self-attention representation for each $head_{f, F_{(Weather, Length, \ldots)}}$, which can be comprehensively represented by Formula 27. On this basis, a concatenation function Concat($\text{head}_{f, F_{(Weather, Length, \ldots)}}$) is added to form a vector. After concatenating the vectors, a linear transformation is performed to obtain the final output of multi-head self-attention MultiHead($f, F_{(Weather, Length, \ldots)}$) using the weight matrix $W^O$. The multi-head self-attention mechanism captures cross-information from different dimensions, making it more effective than self-attention. As shown in Supplementary Fig. 4, this model mainly consists of three parts: embedding, attention model, and linear layer. First, the features are mapped to a high-dimensional vector space, then the multi-head attention model is used to capture cross-information, and finally, the vectors are concatenated, and a linear layer is used to complete the regression. In this work, the number of heads is 16, and the length of the hidden representation is 128.

Because of the maturity, classicism, and wide application of the BP neural network, it is widely used as the baseline for experimental models in the industry. In this article, we also consider the universality of the experiment and use BP neural network as the experimental reference baseline for XGBoost and multi-head Self-Attention.

### Path-planning algorithm

Any path planning algorithm can incorporate comfort considerations based on comfort prediction models, such as the A* algorithm[48]:

$$f(n) = g(n) + h(n) \tag{28}$$

where $f(n)$ represents the total estimated cost of the cheapest solution through node n. The function $g(n)$ denotes the cost from the start node to node n, effectively measuring the path cost incurred so far. The heuristic function $h(n)$ estimates the cost from node n to the goal node, providing a forward-looking assessment of the remaining distance or effort required to reach the goal. The balance between $g(n)$ and $h(n)$ is critical to the performance of the A* algorithm. By incorporating both the actual cost to reach the current node and an optimistic estimate of the cost to reach the goal, the A* algorithm ensures that it explores paths that appear promising based on past information while also considering potential future costs. This dual consideration helps to efficiently guide the search towards the goal, ideally finding the optimal path with minimal computational overhead.

This paper proposes a path-planning algorithm based on the A* algorithm. This method comprehensively considers path length, travel time, path congestion, and path comfort, represented as follows:

$$C(n) = \alpha L(n) + \beta T(n) + \gamma I(n) + \delta J(n) \tag{29}$$

$$L(n) = \sum_{m=1}^{M} l_m \tag{30}$$

$$T(n) = \sum_{q=1}^{Q} t_q \tag{31}$$

$$I(n) = \frac{\sum_{o=1}^{O} i_o}{O} \tag{32}$$

$$J(n) = \frac{\sum_{p=1}^{P} j_p}{P} \tag{33}$$

$C(n)$ is the cost function, $L(n)$ represents the total distance of the path. $T(n)$ denotes the time required to travel on the path. $I(n)$ signifies the level of path congestion. $J(n)$ represents the evaluation of path comfort. $\alpha, \beta, \gamma, \delta$ are the weights of the above four functions. $l_m$ is the distance of a segment of a path, $t_q$ is the time taken for a segment of a path, $i_o$ is the degree of congestion of a segment of a path, and $j_p$ is the comfort level of a segment of a path. Although these four functions are somewhat interrelated, such as path length affecting travel time, we still calculate them separately because they retain their independence.

In subsequent experiments involving real vehicles, normalize $L(n)$, $T(n)$, $I(n)$ and $J(n)$ to the range [0, 1] using the following formula. Then,

assign weights to the four functions.

$$X_{normalized}(n) = \frac{X(n) - X_{\min}}{X_{\max} - X_{\min}} \tag{34}$$

$X_{\min}$ is the minimum value of the function, $X_{\max}$ is the maximum value of the function, $X(n)$ is the current value of the function, and $X_{normalized}$ is the normalized value.

The path planning method in this study includes four modes: shortest path, shortest time, congestion avoidance, and most comfortable, as shown in Fig. 6. In order to rationally assign weights, the following considerations are proposed. In the shortest path mode, priority should be given to the minimization of the total path length, and therefore a higher weight should be assigned to $L(n)$. In the fastest time mode, priority should be given to the minimization of the total journey time while ensuring that the path length weights are maximized, and therefore, higher weights should be assigned to $T(n)$. Similarly, in the congestion avoidance mode, higher weights should be assigned to $I(n)$ in order to alleviate traffic congestion. Finally, the Most Comfortable Route model emphasizes the factors affecting passenger comfort and, therefore, deserves to be assigned a higher weight to $J(n)$. The weights for these models are shown in Supplementary Table 4.

Not just restricted to A* algorithms, any path planning algorithm can incorporate comfort considerations based on ADCP models.

### Real car testing

The purpose of this experiment is to further verify the effectiveness of applying the ADCP method to autonomous driving path planning and to explore whether the ADCP method can be applied to traditional driver-driven vehicles. The equipment and personnel required for the experiment are shown in Fig. 8, and important equipment such as acceleration sensors have been labeled.

Five combinations of starting and ending points were selected for testing. Each group underwent testing at a consistent time every weekday, utilizing the same autonomous driving solution loaded onto the autonomous vehicle. Data, including jerk and acceleration, was collected, and metrics such as weighted average and maximum values were computed. Concurrently, tests were conducted on the vehicle without loading the model, using the standard global path planning, for comparative analysis. Meanwhile, for the same destination, the ADCP model was used for path planning and tested using traditional manual driving mode. The schematic diagram illustrating this experiment is shown in Supplementary Fig. 5.

A total of 33 volunteers completed the scoring tests. We used a quota sampling technique to match the survey sample with the population (age and gender quotas) to enhance the representativeness of the survey results. Among them, 30 participants, representing the mainstream driving and riding age group (males = 18; females = 12; average age: 22), were recruited through public announcements. Additionally, three volunteers from the middle-aged and older group (males = 2; females = 1; average age: 50) were invited to participate in the experiment after a screening process. During the recruitment phase, a brief preliminary questionnaire about exclusion criteria was sent to volunteers who expressed interest in participating. This questionnaire included the following questions:

(i) The average number of car rides per week; (ii) Whether you have any health issues that affect your riding comfort (e.g., motion sickness, back problems, chronic pain, etc.); (iii) Have you ever experienced any unusual discomfort while driving or riding in a car? (iv) Are you willing to participate in a long-duration car-riding experiment that includes breaks? (up to 2 h in total).

This information was obtained through direct questioning, and we established criteria for screening participants for each question. Specifically, participants were initially excluded from the study if their average number of vehicle rides per week was below 5 or if they responded negatively to any of the other questions.

We provided the volunteers with written informed consent forms and copies of the questionnaires. They did not receive any monetary or credit compensation for their participation, other than the cost of transportation to the experimental site. Since the established experiments on the comfort of autonomous driving is rare, we referenced ISO 2631-1 to a certain extent. This project has been approved by the Ethics committee under the approval number THU01-20240121.

After obtaining informed consent, a single-blind approach was used to prevent certain information (such as the difference between traditional driving and autonomous driving) from biasing participants due to prior knowledge or expectations, which could affect the experiment's results. Participants were divided into ten groups and invited to the experimental site. In each session, participants were unaware of whether they were riding in an autonomous vehicle or whether the vehicle had undergone comfort optimization using a model, aiming to obtain more reasonable scoring results. Each participant group was assigned an alphanumeric code for subsequent survey statistics. The survey was conducted via a questionnaire, including demographic questions (gender, age) and the final evaluation scores. Additionally, to ensure participants were focused during the survey, a control question was added (requiring participants to mark a specific answer option). Participants who did not successfully complete the entire survey were excluded from the study and subsequent analyses, resulting in a total sample of 33 participants.

Based on the total sample, to ensure the fairness and reliability of the experimental results, some experiments with significant deviations were repeated with replacement groups, ultimately showing good internal consistency. In the experimental design, all non-experimental variables that might influence the results were controlled. For example, the models of all test vehicles were kept consistent to eliminate these factors' influence on the outcomes.

To minimize bias in the analysis process and ensure the reliability of the results, after final data processing, participants' scores were randomly ordered using a computer-generated random number table. This process effectively avoided bias from researchers' expectations during data analysis. Additionally, a double-blind state was maintained during data processing and analysis, ensuring that those analyzing the data were unaware of the specific participants or experimental groups corresponding to each score, thus guaranteeing objectivity in data processing. A rigorous standardized data processing procedure was adopted, including data cleaning, organization, and analysis, with different teams or individuals reviewing the work to minimize human error and bias. The appropriate statistical method was chosen to analyze the data, controlling for potential confounding variables to further ensure the accuracy and reliability of the results.

### Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

## Data availability

The datasets generated in this study are available via Figshare under accession code https://doi.org/10.6084/m9.figshare.24915945[49]. The ADCP datasets that are continuously updated and continuously collected are available at https://github.com/shujukaiyuanzhuanyong/adcp-, which can be used for further studies and to validate the method. Source data are provided with the paper.

## Code availability

The code for the autonomous driving comfort prediction model is available as Code Ocean (https://doi.org/10.24433/CO.4148871.v1)[50]. Source data are provided with this paper.

## References


1. Yurtsever, E., Lambert, J., Carballo, A. & Takeda, K. A survey of autonomous driving: common practices and emerging technologies. *IEEE Access* **8**, 58443–58469 (2020).
2. Gasparetto, A., Boscariol, P., Lanzutti, A. & Vidoni, R. in *Motion and Operation Planning of Robotic Systems: Background and Practical Approaches* (eds. Carbone, G. & Gomez-Bravo, F.) (Springer International Publishing, 2015).
3. Sánchez-Ibáñez, J. R., Pérez-del-Pulgar, C. J. & García-Cerezo, A. Path planning for autonomous mobile robots: a review. *Sensors* **21**, 7898 (2021).
4. Zhang, H., Lin, W. & Chen, A. Path panning for the mobile robot: a review. *Symmetry* **10**, 450 (2018).
5. Muhammad, K., Ullah, A., Lloret, J., Ser, J. D. & De Albuquerque, V. H. C. Deep learning for safe autonomous driving: current challenges and future directions. *IEEE Trans. Intell. Transp. Syst.* **22**, 4316–4336 (2021).
6. Htike, Z., Papaioannou, G., Velenis, E. & Longo, S. Motion planning of self-driving vehicles for motion sickness minimisation. In *2020 European Control Conference (ECC)* 1719–1724 (IEEE, 2020).
7. Koulocheris, D., Papaioannou, G. & Chrysos, E. A comparison of optimal semi-active suspension systems regarding vehicle ride comfort. *IOP Conf. Ser. Mater. Sci. Eng.* **252**, 012022 (2017).
8. Siddiqi, M. R., Milani, S., Jazar, R. N. & Marzbani, H. Ergonomic path planning for autonomous vehicles-An investigation on the effect of transition curves on motion sickness. *IEEE Trans. Intell. Transport. Syst.* **23**, 7258–7269 (2022).
9. Artuñedo, A., Godoy, J. & Villagra, J. Smooth path planning for urban autonomous driving using OpenStreetMaps. In *2017 IEEE Intelligent Vehicles Symposium (IV)* 837–842. https://doi.org/10.1109/IVS.2017.7995820. (IEEE, 2017).
10. Wang, Y. et al. Trajectory planning and safety assessment of autonomous vehicles based on motion prediction and model predictive control. *IEEE Trans. Vehicular Technol.* **68**, 8546–8556 (2019).
11. Luciani, S., Bonfitto, A., Amati, N. & Tonoli, A. Comfort-oriented design of model predictive control in assisted and autonomous driving. In *ASME 2020 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference* (American Society of Mechanical Engineers Digital Collection, 2020).
12. Zhu, M. et al. Safe, efficient, and comfortable velocity control based on reinforcement learning for autonomous driving. *Transport. Res. Part C Emerg. Technol.* **117**, 102662 (2020).
13. Jin, X., Yan, Z., Yin, G., Li, S. & Wei, C. An adaptive motion planning technique for on-road autonomous driving. *IEEE Access* **9**, 2655–2664 (2021).
14. Lu, B. et al. Adaptive potential field-based path planning for complex autonomous driving scenarios. *IEEE Access* **8**, 225294–225305 (2020).
15. Becerra, I. et al. Human perception-optimized planning for comfortable VR-based telepresence. *IEEE Robot. Autom. Lett.* **5**, 6489–6496 (2020).
16. Chen, J., Zhao, P., Mei, T. & Liang, H. Lane change path planning based on piecewise Bezier curve for autonomous vehicle. In *Proc. 2013 IEEE International Conference on Vehicular Electronics and Safety* 17–22 (IEE, 2013).
17. Hiemstra-van Mastrigt, S., Groenesteijn, L., Vink, P. & Kuijt-Evers, L. F. M. Predicting passenger seat comfort and discomfort on the basis of human, context and seat characteristics: a literature review. *Ergonomics* **60**, 889–911 (2017).
18. Lerspalungsanti, S., Albers, A., Ott, S. & Düser, T. Human ride comfort prediction of drive train using modeling method based on artificial neural networks. *Int. J. Automot. Technol.* **16**, 153–166 (2015).
19. Chen, Y.-C. et al. Spatial and temporal EEG dynamics of motion sickness. *Neuroimage* **49**, 2862–2870 (2010).
20. Bang, J.-S., Won, D.-O., Kam, T.-E. & Lee, S.-W. Motion sickness prediction based on dry EEG in real driving environment. *IEEE Trans. Intell. Transport. Syst.* **24**, 5442–5455 (2023).
21. Htike, Z., Papaioannou, G., Siampis, E., Velenis, E. & Longo, S. Fundamentals of motion planning for mitigating motion sickness in automated vehicles. *IEEE Trans. Vehicular Technol.* **71**, 2375–2384 (2022).
22. Htike, Z., Papaioannou, G., Siampis, E., Velenis, E. & Longo, S. Minimisation of motion sickness in autonomous vehicles. In *2020 IEEE Intelligent Vehicles Symposium (IV)* 1135–1140 (IEEE, 2020).
23. Salter, S., Diels, C., Herriotts, P., Kanarachos, S. & Thake, D. Model to predict motion sickness within autonomous vehicles. *Proc. Inst. Mech. Eng. D J. Automob. Eng.* **234**, 1330–1345 (2020).
24. Bisoffi, A., Biral, F., Da Lio, M. & Zaccarian, L. Longitudinal jerk estimation of driver intentions for advanced driver assistance systems. *IEEE/ASME Trans. Mechatron.* **22**, 1531–1541 (2017).
25. Zhan, X., Hasan, S., Ukkusuri, S. V. & Kamga, C. Urban link travel time estimation using large-scale taxi data with partial information. *Transport. Res. C Emerg. Technol.* **33**, 37–49 (2013).
26. Lebre, M., Mouel, F. L. & Menard, E. On the importance of real data for microscopic urban vehicular mobility trace. In *2015 14th International Conference on ITS Telecommunications (ITST)* 22–26 (IEEE, 2015).
27. Ohlhof, T. et al. Generation and update of Vmap data using satellite and airborne imagery. In *International Archives of Photogrammetry and Remote Sensing* XXXIII (2000).
28. Brooks, R. The new and improved base framework for National Atlas Data. https://ostrnrcan-dostrncan.canada.ca/handle/1845/256554 (1999).
29. Gamba, P. & Herold, M. *Global Mapping of Human Settlement: Experiences, Datasets, and Prospects* (CRC Press, 2009).
30. Meijer, J. R., Huijbregts, M. A. J., Schotten, K. C. G. J. & Schipper, A. M. Global patterns of current and future road infrastructure. *Environ. Res. Lett.* **13**, 064006 (2018).
31. U.S. Department Of Transportation Federal Highway Administration. Next generation simulation (NGSIM) vehicle trajectories and supporting data. https://doi.org/10.21949/1504477 (2017).
32. Castellanos, J. C. & Fruett, F. Embedded system to evaluate the passenger comfort in public transportation based on dynamical vehicle behavior with user's feedback. *Measurement* **47**, 442–451 (2014).
33. De Winkel, K. N., Irmak, T., Happee, R. & Shyrokau, B. Standards for passenger comfort in automated vehicles: acceleration and jerk. *Appl. Ergonomics* **106**, 103881 (2023).
34. Huang, Q. & Wang, H. Fundamental study of jerk: evaluation of shift quality and ride comfort. https://www.sae.org/publications/technical-papers/content/2004-01-2065/ (2004).
35. Ge, A. Theory and design of vehicle automatic transmission. (1993).
36. Shouren, Z. Automatic gearbox. (1984).
37. International Organization for Standardization. ISO 2631-1:1997, Mechanical vibration and shock—evaluation of human exposure to whole body vibration—part 1: general requirements (1997).
38. Murphey, Y. L., Milton, R. & Kiliaris, L. Driver's style classification using jerk analysis. In *2009 IEEE Workshop on Computational Intelligence in Vehicles and Vehicular Systems* 23–28 (IEEE, 2009).
39. Zhou, J. et al. Autonomous driving trajectory optimization with dual-loop iterative anchoring path smoothing and piecewise-jerk speed optimization. *IEEE Robot. Autom. Lett.* **6**, 439–446 (2021).
40. Abiodun, O. I. et al. State-of-the-art in artificial neural network applications: a survey. *Heliyon* **4**, e00938 (2018).
41. Mäkinen, E. A Survey on Binary Tree Codings. *The Computer Journal* **34**, 438–443 (1991).
42. Li, J., Cheng, J., Shi, J. & Huang, F. Brief introduction of back propagation (BP) neural network algorithm and its improvement. In

*Adv. Computer Science and Information Engineering* (eds. Jin, D. & Lin, S.) 553–558 (Springer, 2012).
43. Chen, T. & Guestrin, C. XGBoost: a scalable tree boosting system. In *Proc. 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 785–794 (Association for Computing Machinery, 2016).
44. Nielsen, D. Tree Boosting With XGBoost (2016).
45. Ramraj, S., Uzir, N., Sunil, R. & Banerjee, S. Experimenting XGBoost algorithm for prediction and classification of different datasets. *Int. J. Control Theory Appl.* **9**, 651–662 (2016).
46. Mitchell, R. & Frank, E. Accelerating the XGBoost algorithm using GPU computing. *PeerJ Comput. Sci.* **3**, e127 (2017).
47. Vaswani, A. et al. In *Advances in Neural Information Processing Systems* vol. 30 (Curran Associates, Inc., 2017).
48. Tang, G., Tang, C., Claramunt, C., Hu, X. & Zhou, P. Geometric A-star algorithm: an improved A-star algorithm for AGV path planning in a port environment. *IEEE Access* **9**, 59196–59210 (2021).
49. Zhengxian, C. et al. Predicting driving comfort in autonomous vehicles using road information and multi-head attention models. figshare https://doi.org/10.6084/m9.figshare.24915945 (2025).
50. Zhengxian, C. et al. Predicting driving comfort in autonomous vehicles using road information and multi-head attention models. Code Ocean. https://doi.org/10.24433/CO.4148871.v1 (2025).

## Acknowledgements

This work is supported by the National Key Research and Development Program of China: 2022YFB2503003 (Chaosheng Huang), and the National Natural Science Foundation of China Project: 52172388 (Chaosheng Huang).

## Author contributions

Z.C. and C.H. led the research program and proposed the initial concept of autonomous driving comfort prediction. Z.C. and Y.L. developed the algorithms and wrote the paper. Z.C. and H.H. collected and constructed the datasets. W.N., B.X., Z.L., and W.Y. drew pictures and analyzed the results. C.H. and J.L. provided solid suggestions for the program. Z.C., H.W., Y.S., and Y.L. performed the experiments. C.H. takes responsibility for the organization of the overall paper and approved the submission.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at https://doi.org/10.1038/s41467-025-57845-z.

**Correspondence** and requests for materials should be addressed to Chaosheng Huang.

**Peer review information** *Nature Communications* thanks Israel Becerra and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.