

A transferable vehicle energy consumption evaluation framework for quantifying vehicle electrification energy benefits

Author names: *Zhenyu jia^a, Zeping Cao^a, Jiawei Yin^a, Lin Wu^a, Ning Wei^b, Yanjie Zhang^c, Zhiwen Jiang^a, Qijun Zhang^{a,*}, Hongjun Mao^{a,*}*

Author Address: a Tianjin Key Laboratory of Urban Transport Emission Research & State Environmental Protection Key Laboratory of Urban Ambient Air Particulate Matter Pollution Prevention and Control, College of Environmental Science and Engineering, Nankai University, Tianjin, 300071, China

b Jinchuan Group Information and Automation Engineering Co. Ltd., Jinchang, 737100, China

c Tianjin Youmei Environment Technology, Ltd., Tianjin, 300380, China

Corresponding Author's e-mail address: zhangqijun@nankai.edu.cn (Q.Zhang),
hongjunm@nankai.edu.cn (H.Mao).

ABSTRACT. Vehicle electrification is an important approach for energy saving and emission reduction. However, the advantages of electric vehicles in terms of energy consumption are difficult to quantify accurately. Based on the 3-D markov chain and machine learning methods, this study proposes a framework to evaluate energy consumption benefits and validates it with two typical cities, Tianjin and Xining. The results show that electric vehicles can achieve energy consumption benefits from 16.4 % to 77.8 %. While electric vehicles have significant energy consumption

benefits, there are regional differences. Compared to internal combustion engine vehicles, battery electric vehicles showing greater benefits in Tianjin and plug-in hybrid electric vehicles in Xining. The model interpretation indicates that the difference stems from Tianjin's more frequent and intense deceleration and Xining's more varied terrain. A standardized energy consumption evaluation framework that integrates regional and vehicle type differences is essential to the electric vehicles promotion.

KEYWORDS: Energy consumption, Vehicle electrification, Driving cycle database, Machine learning, Markov chain.

Introduction

Promoting electric vehicles is an important approach to reduce air pollution and greenhouse gas emissions from motor vehicles, and to reduce energy consumption¹. Over the past decade, the world's light-duty electric vehicles (LDEVs), including battery electric vehicles (BEVs) and plug-in hybrid electric vehicles (PHEVs), increased from 0.4 million units in 2013 to 16.4 million units in 2022². It is expected that by 2030, the global electric vehicle population could potentially increase to tenfold its current quantity³. In the context of such widespread rollout, it is crucial to evaluate the real-world benefits of replacing conventional fuel vehicles with electric vehicles. Research on this topic primarily focuses on the reduction of pollutants ⁴⁻⁶, greenhouse gas emissions ^{7,8}, and associated health benefits ^{9,10}. A critical and foundational factor that limits the rationality and accuracy of these studies is the

absence of a standard frame for assessing vehicle energy consumption. Therefore, a standardized energy consumption evaluation framework that integrates regional and vehicle type differences is essential to help the public and policymakers recognize the potential energy benefits of EVs.

Standard driving cycles (DCs) represent a unified method for providing comparability in energy consumption across various vehicle types worldwide. Commonly used DCs include NEDC (New European Driving Cycle), WLTC (Worldwide Harmonized Light Vehicles Test Cycle), and CLTC(China light-duty vehicle test cycle)¹¹. These cycles transform complex real-world traffic conditions into a simplified speed-time curve using techniques like slicing, weighting, and concatenation^{12,13}, thus providing a uniform benchmark for evaluating energy consumption across different vehicle types.

In a broad context, standard DCs play an outstanding role. However, when the focus shifts to specific regions or cities, these cycles often exhibit considerable deviations. Transient DCs like WLTC, despite including authentic driving data, still struggle to comprehensively encompass all operational characteristics of vehicle operation due to the diversity in road conditions, terrain, and traffic situations across different regions¹⁴.

The study by Fontaras, G., et al., identified that carbon emissions measured using the WLTC exhibit a deviation of 15-22% compared to real road experiments¹⁵. For electric vehicles, Wang, H et al. observed that different standard DCs can result in energy consumption deviations of approximately 19%-30%¹⁶. Additionally, due to the presence of energy recovery systems in electric vehicles and their relatively heavier weight compared to fuel vehicles, they are more sensitive to changes in gradient than

fuel vehicles^{17,18}. However, the generalized use of velocity-time curves fails to account for the differential impact of terrain on various vehicle types, leading to increased inaccuracies in comparing energy consumption across different vehicle types.

Due to the inability of standard DCs to accurately reflect the energy consumption characteristics of different regions, researchers have conducted numerous studies aimed at determining the energy consumption performance of vehicles in specific areas. A prevalent approach entails conducting statistical analysis on actual energy consumption data gathered from real-world scenarios. This approach is straightforward and effective, allowing for the direct identification of factors influencing energy consumption. Studies based on this method have been applied to investigate the impact of temperature on energy consumption^{19,20}, as well as the relationship between driving distance and energy efficiency²¹. Nevertheless, the complexity of real-world traffic scenarios indicates that energy consumption in different regions is significantly influenced by variables such as terrain and traffic conditions²². This complexity requires thorough experimental efforts to establish credible patterns of energy consumption, which is undoubtedly time-consuming and laborious. With the rise of various new vehicle models in the current context of vehicle electrification, the task of conducting sufficient experiments to accurately evaluate real-world conditions becomes an almost insurmountable challenge²³. An alternative to the aforementioned method is the driving cycle method. Similar to the concept of standard DCs, this approach transforms a region's complex traffic

conditions into fixed-duration local DCs through slicing, weighting, and other methods. Upon obtaining the DCs, researchers can ascertain the energy consumption characteristics of different regions using methods such as applying chassis dynamometer testing or utilizing energy consumption models. Currently, cutting real driving data and utilizing Markov chains are the two main methods for constructing DCs. The cut real driving data method encompasses micro-trip method²⁴, trip segment method and driving fragmentation method^{25,26}. Regardless of the method of cutting used, the segments obtained frequently surpass the required quantity for constructing a DC. It is necessary to conduct a cluster analysis and select segments so that a DC could be established^{27,28}. However, the clustering method is constrained by its propensity to yield locally optimal solutions and the process of selecting segments risks omitting crucial real motion data²⁹. This can result in DCs that do not sufficiently reflect real-world conditions. In addition to the cut real driving data method, Markov chain-based methods can also be used to construct DCs. During the vehicle's traveling process, the probability distribution of its next state is solely determined by its current state. This aligns with the Markov chain's "memorylessness assumption", wherein each subsequent state is independent of the past states and relies only on the present state. State transition probability matrix constructed from the vehicle's current and next motion states capture changes in velocity and acceleration. This enables a holistic consideration of the driving characteristics for the entire range of collected sample conditions, thereby enhancing the efficiency of utilizing the original data^{30,31}. Additionally, the Markov chain method more accurately embodies

the inherent randomness of speed compared to the cut real driving data method³².

Currently, a variety of studies are concentrated exclusively on assessing the energy consumption of vehicles in single, distinct regions. Research by various scholars has revealed significant energy consumption differences across regions. For example, in cities such as Beijing, Xi'an, Tianjin, and Dublin, the kinematic characteristics of vehicles show considerable variation. This is reflected in their differing electric energy consumption rates, which are 0.1354Wh/(kg·km), 0.1267Wh/(kg·km), 0.1067Wh/(kg·km), and 0.1404Wh/(kg·km) respectively³³⁻³⁶. However, due to the lack of uniformity in DC construction methods in these studies, such energy differences are influenced by the varying methods used. Wang T et. al.'s research demonstrates that using the same data to construct DCs with various clustering techniques can lead to notable differences in energy consumption, with outcomes of 18.52kWh/100km and 17.88kWh/100km, respectively³⁷. Furthermore, the methods researchers use to establish the mapping relationship between DCs and energy consumption often vary, making it even more challenging to pinpoint the sources of vehicle energy consumption differences. Therefore, based on existing testing protocols and relevant studies, it remains challenging to accurately determine the energy consumption benefits when promoting electric vehicles in a specific location or city. To address these issues, this research proposes a framework designed to standardize the evaluation of energy consumption across different regions. The frame utilizes the Markov chain method with real driving data to generate a specific number of local DCs, forming a driving cycle (DC) database, which extracts the key driving

and elevation features from local driving data. This approach maintains the randomness of individual DCs within the DC database, while simultaneously mitigating the potential for significant deviations in randomly generated cycles. This approach enhances the stability of the energy consumption values represented by the database. Moreover, to bypass the exhaustive and time-consuming multiple chassis dynamometer tests and to improve the precision of energy consumption estimates³⁸, the study implements a high-precision energy consumption prediction model based on advanced machine learning techniques, replacing traditional prediction models, thereby providing more accurate and localized energy consumption forecasts. Finally, the framework quantifies the energy consumption differences between vehicle types across regions by integrating the energy consumption prediction model with the DC database. By applying this process to ICEVs, BEVs, and PHEVs, the energy consumption reduction benefits of promoting these vehicles can be scientifically compared. To demonstrate the usability of this framework, this study was corroborated using Tianjin and Xining as representative city examples. Tianjin, a typical plain city, has an average elevation of 3.5m, with plains covering 93% of the city and a car ownership rate of 296 vehicles per thousand people³⁹. Xining, a typical plateau city, has an average elevation of 2261 meters, characterized by rugged terrain, narrower and longer roads⁴⁰, and a car ownership rate of 362 vehicles per thousand people⁴¹. The development and validation of this framework provide a unified method to quantify the energy consumption differences between vehicle types across regions. This aids policymakers in comprehending the advantages of energy consumption

when promoting electric vehicles in various regions, thus facilitating the development of specialized electric vehicle promotion strategies tailored to specific areas.

Methods.

Data Collection. In order to verify the applicability and accuracy of the framework, the study collected real driving data from Tianjin and Xining, two cities with distinct geographical and traffic characteristics. Supplementary Figure 1 illustrates the specific geographic locations of Xining and Tianjin. The data collected in this study consists of two parts: real-world driving data and energy consumption data. The driving data were collected from the Global Positioning System (GPS) equipped in each vehicle, with a sampling rate of 1Hz. In total, data was collected from 35 vehicles, including 20 from Tianjin and 15 from Xining, amounting to approximately 200 hours and covering a distance of over 5637 kilometers. The collected data primarily includes information like latitude, longitude, elevation, and vehicle speed. Supplementary Figures 2 and 3 illustrate the specific driving routes in Xining and Tianjin, respectively.

Regarding energy consumption data, they were gathered using the VAGCOM Diagnostic System (VCDS) at a sampling rate of 1Hz. The data include kinematic information such as speed and acceleration, as well as energy consumption indicators like instantaneous fuel consumption, instantaneous electric current, and battery voltage. They were sourced from a total of six vehicles, three each from Tianjin and Xining. In each city, the vehicles consisted of one BEV, one PHEV, and one ICEV. The driving routes are illustrated in Supplementary Figures 4 and 5. Despite

variations in certain specifications, the vehicles are equipped with the same type of power system, ensuring consistency in energy consumption patterns. The detailed parameters of the vehicles have been listed in Table 1.

Build regional driving cycle database. Due to the significant differences in terrain and driving characteristics between Xining and Tianjin, the study has chosen speed(v), acceleration (a), and elevation(h) as the input variables for the Markov chain. To mitigate potential bias in energy consumption estimation caused by the randomness of individual DCs, this paper constructed two DC databases for Tianjin and Xining. These databases effectively utilized the randomness of the Markov chain, significantly improving the utilization of driving data. While maintaining the randomness of individual conditions, they also ensured the stability and reliability of the final energy consumption results. The process for generating the DC database is as follows: 1) Discretize v, a, and h with resolutions of 0.1 m/s, 0.2 m/s², and 1 m, respectively. 2) Aggregate v, a, and h into a v-a-h network. For each specific v-a-h combination, calculate the frequency of transitioning to the next state. Use the maximum likelihood estimation method to estimate probabilities with frequencies, generating state transition matrix for each state. 3) Select a state with the longest vehicle idling time as the initial state ($v_0=0$ km/h, $a_0=0$ m/s², $h_0 = 2222$ m for Xining; $v_0=0$ km/h, $a_0=0$ m/s², $h_0 = -2$ m for Tianjin, where v_0 , a_0 , h_0 represent the initial speed, acceleration, and elevation) and randomly choose the next state based on the state transition matrix. Continue this cycle, and at 900 s, check the maximum and minimum values of acceleration and slope. If $a_{\max} \leq 4.6$ m/s², $a_{\min} \geq -8$ m/s², $\text{slope}_{\max} <$

10 %, and $\text{slope}_{\min} > -10\%$, continue the calculation process until a 1800s DC is generated. Otherwise, stop the cycle to conserve computational resources. 4) If the following conditions are met: $a_{\max} \leq 4.6 \text{ m/s}^2$, $a_{\min} \geq -8 \text{ m/s}^2$, $\text{slope}_{\max} < 10\%$, $\text{slope}_{\min} > -10\%$, and the final speed is 0, then add the DC to the database. Repeat steps 2-4 until the database contains a specified number of DCs. Figure 1 details this process. It's worth mentioning that due to the low precision of GPS data, instantaneous slopes are not directly calculated. Instead, a linear regression is performed on a 160 m running segment, and the regression slope is used as the slope, discarding cases where the slope is greater than 10% or less than -10 %⁴².

The number of DCs in the DC database is determined based on the accumulative absolute error between the Speed-Acceleration-Height Frequency Distribution (SAHFD) of the DCs and real driving data (DiffSum). SAHFD reflects the driving and terrain characteristics of real driving data or DCs. The calculation method of SAHFD is as follows. Firstly, discretize v, a, and h with resolutions of 0.1 m/s, 0.2 m/s², and 1 m respectively. Afterward, aggregate v-a-h and calculate the frequency corresponding to different states. The resulting three-dimensional matrix of frequency distribution is SAHFD. The difference between real driving data and DC database, termed DiffSum, is used to describe the disparity between the cycle and real driving data. The calculation formula is given by equation (1):

$$\sum_{i=1}^{n_{\text{height}}} \sum_{j=1}^{n_{\text{col}}} \sum_{k=1}^{n_{\text{row}}} |f_{\text{DC database}_{i,j,k}} - f_{\text{Driving data}_{i,j,k}}| \#(1)$$

Where $f_{\text{DC database}_{i,j,k}}$ represents the frequency corresponding to the state of the i-th layer, j-th column, and k-th row in the entire DC database, $f_{\text{Driving data}_{i,j,k}}$ represents the

frequency corresponding to the state of the i-th layer, j-th column, and k-th row in the real driving data.

Increasing the number of DCs in the DC database can lead to a decrease in DiffSum. In other words, it makes the entire database closer to real driving data. However, this decreasing trend is not endless. When the number of DCs in the database reaches a certain point, further increasing the count leads to only a marginal decrease in DiffSum. The minimum number of DC database required to make DiffSum converge is selected as the number of DCs database. These DCs are then added to the database in the order of generation.

Energy Consumption Model Construction. To enhance the precision of the model as much as possible, this article employs a data-driven machine learning approach to construct an energy consumption model. XGBoost, a gradient boosting tree algorithm proposed by Chen T et al.⁴³, excels at adaptively learning non-linear relationships in data. It also employs regularization to avoid overfitting. The model inputs include 13 different features, such as vehicle speed (v), acceleration (a), Vehicle Specific Power (VSP), Relative Positive Acceleration ¹⁰, vehicle jerk (J), average speed over the previous 5 seconds (avg.v), speed variance over the previous 5 seconds (var.v), average acceleration over the previous 5 seconds (avg.a), acceleration variance over the previous 5 seconds (var.a), elevation (H), the sum of elevation over the previous 5 seconds (sum.h), slope (s), and change of slope(Δs).

VSP (kw/ton), introduced in 1998 by Jimenez et al.⁴⁴. considers variations in kinetic and potential energy, along with the energy expended against friction and air

resistance during the operation of a vehicle. Although the VSP formula includes parameters calibrated based on the authors' experiments, XGBOOST can still identify key features related to the target variable in the data, enabling accurate predictions under different driving conditions. The VSP formula is given by equation (2):

$$VSP = v[1.1a + 9.81[a \tan(\sin\theta)] + 0.132] + 0.000302v^3 \#(2)$$

where VSP is the Vehicle Specific Power, v is instantaneous driving speed (m/s), a is instantaneous acceleration (m/s^2), and θ is the angle between the road surface and the horizontal plane.

RPA, representing the vehicle's driving dynamics (m/s^2), is a time-averaged measure of the specific power, normalized by average driving speed. It indicates the intensity of driving. The formula for RPA is given by equation (3):

$$RPA = \frac{\int_{T-5}^T (v_i \times a_i^+) \cdot dt}{x} \#(3)$$

where T is a specific moment ($T \geq 5$), v_i is the speed at moment i (m/s), and a_i^+ is the positive acceleration at moment i (m/s^2). When $T \leq 5$, both speed and acceleration are essentially zero, so the RPA for the initial 5 seconds of each run is set to zero.

J , the vehicle jerk, reflects the bumpiness during driving and can be used to measure aggressive versus gentle driving. Aggressive driving may lead to higher energy consumption. The formula for J is given by equation (4):

$$J = \frac{da}{dt} = \frac{d^2v}{dt^2} \#(4)$$

where v is the vehicle's instantaneous speed, and a is its acceleration.

Slope is the road gradient (%). Due to the inaccuracy of GPS devices in measuring elevation, the segments are first binned into 160m distances. Then, linear regression is

applied on the elevation and cumulative distance within each segment, and the regression slope is used as the gradient. The formula is as (5):

$$H = \text{slope} X + b \#(5)$$

where slope is the estimated slope, X is the cumulative distance vector starting from 0, with a maximum not exceeding 160 meters, and b is the intercept of the linear regression.

Other parameters such as avg.v, var.v, avg.a, var.a, sum.h, Δs, etc., can be found in Supplementary Table 1.

In addition to the model input parameters, additional parameters are required for the construction of energy consumption data, which are recorded by VCDS. For BEV, this includes battery voltage and intermediate circuit current. For ICEV, it's fuel consumption. And for PHEV, it's battery voltage, intermediate circuit current, and fuel consumption. The formulas are given by equations (6), (7), and (8):

$$P_{ICEV} = \frac{F_{Ci}}{F_E} \#(6)$$

$$P_{BEV} = C_f U_{bi} I_{bi} (\Delta t) \#(7)$$

$$P_{PHEV} = \frac{F_{Ci}}{F_E} + C_f U_{bi} I_{bi} (\Delta t) \#(8)$$

where P_{ICEV} (kWh/s), P_{BEV} (kWh/s), P_{PHEV} (kWh/s) are the instantaneous energy consumption of ICEV, BEV, PHEV, respectively. U_{bi} is the instantaneous battery voltage (V/s), I_{bi} is the instantaneous intermediate circuit current (A/s), Δt (s) is the time interval, and C_f is the unit conversion factor from J/s to kWh/s, which is 1/3600000 in this study. F_E is the fuel consumption coefficient, which for the 92-gasoline used in this experiment is 0.1161 (L/(kWh)).

Upon formulating the feature and target values, the energy consumption datasets for different vehicle types (BEV, ICEV, PHEV) in Tianjin and Xining are combined into three new datasets. Although the vehicles in each type are of different models and have slightly different energy consumption, they use the same powertrain. Considering this, before merging datasets, a special correction method was employed to process the energy consumption data of different vehicles. This method is based on the ratio of test values obtained under the WLTC test. For example, for ICEVs, the WLTC energy consumption test conducted in Xining revealed a value of 55.1 kWh/km, while in Tianjin, the corresponding result was 53.8 kWh/km, resulting in a ratio of 0.976. Consequently, the study adjusted the energy consumption data for Xining by multiplying it with 0.976. For PHEVs, the Xining data was multiplied by 0.988. For BEVs, the Tianjin data was multiplied by 0.979. This correction ensured data consistency and enhanced the accuracy of our analysis.

Three separate machine learning models were established for the three different types of vehicles. Each dataset was randomly divided into 70% training set and 30% test set. During the model training process, a 10-fold cross-validation method was used to evaluate the model's performance. This method effectively utilizes limited data and enhances the reliability of the model's generalization ability. Moreover, to optimize the model's performance, a cross-grid search was conducted to find the best combination of model parameters within a predefined range. The specific range of hyperparameters is shown in Supplementary Table 2.

Model Evaluation and Coupled Result Analysis. To assess the predictive

accuracy of the XGBOOST models, two metrics, Root Mean Square Error (RMSE) and Coefficient of Determination (R^2), were employed for evaluating the three machine learning models. Their formulas are given by equations (9) and (10):

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{n=1}^N (y_n - \hat{y}_n)^2} \#(9)$$

$$R^2 = 1 - \frac{\sum_{n=1}^N (y_n - \hat{y}_n)^2}{\sum_{n=1}^N (y_n - \bar{y})^2} \#(10)$$

where N is the total number of samples being evaluated, y_n and \hat{y}_n represent the actual and predicted values for the n_{th} sample, respectively. To demystify the prediction behavior of XGBOOST during the modeling process and break the "black box" nature of machine learning models, the Shapley Additive exPlanations (SHAP) method, based on Shapley values, is utilized. SHAP employs the concept of Shapley value from game theory to explain machine learning models by determining the overall contribution of each feature to the model's prediction. Additionally, marginal contribution plots based on Shapley values can also delineate contributions of different features to energy consumption.

The methodology used for processing energy consumption data is also applied to handle the DC database, specifically in the calculation of the 13 feature values. Subsequently, the processed DC databases are coupled with different machine learning models to predict vehicle energy consumption. To better understand the influence of different features on the model outputs and thereby discern the differences in DC databases that lead to discrepancies in energy consumption, the SHAP method is also applied to interpret the outputs of the three distinct models.

Overall, the framework can be summarized into four main parts: acquiring real data, constructing DC databases, building energy consumption models, and coupling energy consumption models with DC databases. Figure 2 illustrates the specific process of this framework.

Results.

Construction of the Driving Cycle Databases. Based on the 3-D MARKOV method, DC databases for Tianjin and Xining were established. The DC database in Tianjin contains 60 DCs, and the DC database in Xining contains 115 DCs. Figure 3 illustrates partial DCs in the databases of Tianjin and Xining, along with the descent process of the databases' DiffSum as the number of DCs in the database increases. Compared to a single DC, a DC database with elevation data better reflects regional differences in driving characteristics. Nine different metrics were employed to quantify differences between the Tianjin and Xining DC databases, including average speed, average acceleration, average deceleration, average slope, accelerated time ratio, deceleration time ratio, constant speed time ratio, idle time ratio and standardized elevation (Figure 3). The calculation formulas are provided in Supplementary Table 3. To facilitate comparison of average elevations between Tianjin and Xining, elevation values for both cities were scaled to a 0-1 range using a min-max scaler.

Compared to Xining, Tianjin displays significantly higher average speed, smaller average deceleration, longer average deceleration time ratio, and shorter average idle and constant speed time ratio. Furthermore, while average slope is similar in both

cities, Xining has a significantly higher slope variance, indicating more significant terrain variability. Additionally, all nine metrics fluctuated within a certain range with few outliers. Box plots demonstrated that the Markov chain method retains realistic driving characteristics while introducing randomness.

Statistical methods further confirmed the significance of these differences. Firstly, the distribution of 18 groups of data across nine metrics for Tianjin and Xining underwent a Shapiro-Wilk normality test. Results indicated that Tianjin's average acceleration, standardized elevation, and Xining's average slope and idle time ratio did not follow a Gaussian distribution. Then, a Levene's test for homogeneity of variances was conducted between the two cities for the same metrics, revealing differences in variance for idle time ratio. Two-sample t-tests were conducted on groups that satisfied the conditions of homogeneity of variance and normal distribution, revealing significant differences among the groups. For the remaining groups, the Wilcoxon Rank-Sum test was employed, revealing a significant difference in idle time ratio between the two cities. Supplementary Table 4 provides a detailed overview of the results and their significance.

Construction and Interpretation of the Energy Consumption Model. Using XGBoost, energy consumption models for ICEV, BEV, and PHEV were constructed. The R^2 values for these models on the test set were 0.83, 0.92, and 0.74, respectively, while the RMSE values were 0.0017, 0.0018, and 0.0036. High R^2 and low RMSE on the test set indicate good generalization ability of the models. Figure 5 shows the models' performance on training and test sets, with similar trends on both datasets

suggesting no overfitting. The PHEV model had the lowest R^2 , possibly due to its complex powertrain structure. Yet, a coefficient of 0.74 is still higher than most conventional models, providing robust estimates for future predictions.

For machine learning models, particularly 'black box' models, overreliance on their outputs can lead to critical errors. These models may sometimes erroneously capture patterns inconsistent with reality, yielding seemingly accurate results. In order to ensure accuracy in subsequent coupling with the DC database and to better understand the internal mechanisms of the models, Shapley Additive exPlanations (SHAP) were introduced. Figure 6 displays the average absolute Shapley values based on the training set, indicating the importance of features in the model's output. To further understand different features' impact on energy consumption predictions, Shap marginal contribution plots were drawn. Figure 7 shows the marginal contribution of VSP, with other features' contributions detailed in Supplementary Figure 7-9. On these plots, the x-axis represents the actual feature values of samples, and the y-axis represents the contribution of these feature values to the model output at that point. The color mapping reflects the model's predicted values for corresponding samples, with redder colors indicating higher energy consumption and bluer colors indicating lower consumption.

For ICEVs, VSP and acceleration (a) are key features. As a and VSP gradually increase from 0, their contribution to ICEV energy consumption noticeably increases, with corresponding real energy consumption also rising. However, as a and VSP decrease from 0, their impact on predicted energy consumption remains largely

unchanged. In fact, when a and VSP are positive, they reflect the power demand during ICEV operation, with higher values indicating increased engine output power and energy consumption. During deceleration, the braking system operates independently of the engine, thus having minimal impact on energy consumption. Additionally, the impact of vehicle speed on energy consumption in the characteristics exhibits a trend of initially increasing and then decreasing, consistent with the pattern of the engine having an optimal efficiency range⁴⁵.

For BEVs, VSP still holds the most crucial position, but the importance of other features changes significantly. Features like avg.v and speed see considerable increases in their contributions., while a and J see significant decreases. Moreover, features related to terrain, such as sum.h, H, and slope, gain considerable importance (indicated by a decrease in average absolute contribution). Specifically, as VSP decreases from zero, its contribution also decreases. The contribution of speed correlates positively with its value, while the contribution of avg.v inversely correlates with speed — higher avg.v values contribute less. This phenomenon is likely related to BEVs' operational mode. Firstly, BEVs can recover energy during deceleration or downhill driving, thus responding better to changes in energy consumption and terrain across the entire speed range. In contrast, ICEVs primarily exhibit noticeable energy consumption changes during acceleration. Hence, in BEVs, speed and avg.v have a greater impact on energy consumption compared to ICEVs. Secondly, BEVs and ICEVs have different energy consumption strategies. Most of BEVs' energy consumption is used to maintain speed, with less energy allocated for

acceleration, making them less sensitive to changes in a and j . Unlike ICEVs, BEVs have a larger body weight and lower energy consumption, with changes in gravitational potential energy more directly reflected in energy consumption. Additionally, energy savings from downhill driving further enhance the impact of terrain on energy consumption. These factors are reflected in the increased importance of terrain-related features (sum.h, H, and slope).

For PHEVs, the importance of features other than VSP is relatively similar. Additionally, the importance of sum.h, H, and slope remains substantial. Unlike ICEVs and BEVs, the powertrain of PHEVs includes both an engine and an electric motor, and the energy consumption patterns of these two systems are different. During operation, changes in PHEV power demand usually coincide with changes in powertrain conditions, making it more challenging for the model to identify contributions from the vehicle's motion state. Consequently, the importance of features that directly describe changes in the vehicle's motion state, such as acceleration a and J , is significantly reduced, displaying a more uniform importance. Conversely, features that directly reflect the vehicle's energy or changes directly related to it, such as VSP, sum.h, H, and slope, are more important. Moreover, PHEVs also have some characteristics of BEVs, such as heavier body weight and energy recovery capabilities, further enhancing the importance of sum.h, H, and slope.

Model coupling results and interpretation of model outputs. By integrating the DC databases from both cities with the energy models for ICEVs, BEVs, and PHEVs, box plots were plotted to illustrate the variations in energy consumption

across different cities and vehicle types (Figure 8). The precise findings were as follows: for Tianjin, ICEV consumed 0.554 kWh/km, BEV 0.123 kWh/km, and PHEV 0.463 kWh/km; for Xining, ICEV consumed 0.534 kWh/km, BEV 0.138 kWh/km, and PHEV 0.401 kWh/km (represented by the medians of the box plots).

Supplementary Table 5 presents the 95% confidence intervals for these estimates for both Tianjin and Xining. It was evident that using the same vehicle type in different cities resulted in significant energy consumption differences. Compared to Xining, the energy consumption of Tianjin ICEV was higher by 3.7%, PHEV by 11.3%, and BEV lower by 10.9%. In Tianjin, PHEV energy consumption was 83.6% of ICEV, and BEV consumption was 26.5% of PHEV and 22.2% of ICEV. In Xining, PHEV consumption was 76.4% of ICEV, and BEV consumption was 33.8% of PHEV and 25.8% of ICEV. In summary, promoting PHEVs and BEVs demonstrates energy consumption advantages in both cities. In Tianjin, the energy consumption benefits for promoting BEVs surpass those in Xining; meanwhile, in Xining, promoting PHEVs shows higher energy consumption benefits compared to Tianjin.

To identify the sources of energy consumption differences among different vehicle types in each database, a typical DC was selected from each of the Tianjin and Xining databases. Energy consumption curves for the three vehicle types and the changes in a , v , J , and slope over time were plotted (Figure 9). Generally, the vehicle types exhibited similar characteristics in terms of energy consumption peaks and lows. High speeds and accelerations led to energy consumption peaks. Conversely, during deceleration, ICEVs showed lower energy consumption, while BEVs and PHEVs

demonstrated negative energy consumption, indicating energy recovery. Specifically, ICEVs exhibited higher peak values and did earlier compared to the other two types. Both BEVs and PHEVs could recover energy, but BEVs recovered more. The lower energy conversion efficiency of internal combustion engines primarily causes higher instantaneous energy consumption. Additionally, during acceleration, a rapid increase in torque is often needed within a short time. Electric motors can provide maximum torque instantly upon starting, whereas the torque increase in internal combustion engines is linear. To rapidly increase torque, internal combustion engines inject more fuel to increase speed, resulting in higher and earlier peaks. Compared to BEVs, PHEVs' less efficient drivetrain results in relatively higher energy consumption, and their smaller electric motors mean less energy recovery; this leads to greater instantaneous and total energy consumption.

To identify the specific features responsible for the differences in model outputs between the two cities, the Shapley method was applied to analyze the output of the three models in both cities, with Supplementary Figure 10 showing the average absolute contributions. The energy consumption of Tianjin's ICEVs is close to Xining's. Two factors synergistically contributed to this outcome. On the one hand, there is a certain disparity in the positive VSP between Tianjin and Xining, with mean values of 5.20 kW/ton and 4.21 kW/ton, respectively. On the other hand, the different terrains of the two cities also have varying impacts on the fuel consumption of ICEVs. Compared to Tianjin, Xining has higher contributions in terms of H, sum.h, and slope, which to some extent, increase the fuel consumption of internal combustion vehicles

in Xining. Overall, the influence of terrain is slightly less than that of VSP, resulting in slightly higher energy consumption in Tianjin compared to Xining. Similarly, for PHEVs, variables describing terrain features showed higher importance in the output for Xining than for Tianjin, yet this difference was less than the impact caused by VSP, avg.a, and speed. The final energy consumption results showed Tianjin has higher consumption than Xining. However, for BEVs, energy consumption in Xining was actually higher than in Tianjin. Although Tianjin's higher average speed had a negative effect on BEV energy consumption, Tianjin's more frequent and severe deceleration behaviors were more conducive to energy recovery in BEVs, resulting in lower overall energy consumption. Additionally, Xining's greater slopes had a more significant impact on the heavier BEVs, increasing their energy consumption and further widening the energy consumption gap between the two cities.

Discussion.

In this section, we provide a review of the methods used in this study. The Markov Chain method was selected to extract driving and elevation features for different regions. This method is computationally efficient, simple yet effective, and can handle three-dimensional input data. In this study, besides speed and acceleration, elevation data is also incorporated. This makes the method particularly suitable for characterizing driving and terrain features across different regions, and for comparing terrain differences. Furthermore, unlike traditional two-dimensional speed-time driving cycle curves, the driving cycle database in this study is a three-dimensional curve that includes elevation, speed, and time. Therefore, a longer duration is required

to adequately capture the driving characteristics of a region, and the complexity of information varies by region, necessitating different sample sizes.

XGBoost was chosen for training the machine learning models. This method offers low training costs and has a strong research foundation, particularly in the context of PHEV energy consumption modeling. Additionally, the SHAP model was used to explain the behavior of the machine learning models within the proposed framework. This approach is computationally efficient, cost-effective, and allows for a macro-level understanding of the impact of a specific variable on model performance. It also supports various models, meaning that in the future, smaller computational models could be used in large-scale implementations while still coupling with the current framework.

The study innovatively proposes an energy consumption evaluation framework that can unify the energy consumption evaluation scales between different vehicle types, thereby quantifying the feasibility of promoting electric vehicles in different regions.

The framework standardizes the energy consumption evaluation scales across different vehicle types, addressing the current inconsistencies in vehicle evaluation standards and scales across regions. Its excellent generalizability allows for seamless adaptation and expansion to various regions, enabling precise energy consumption assessments for vehicles.

Using the established regional DC databases and instantaneous energy consumption models, the levels of EVs energy consumption in Tianjin and Xining were quantified. Compared to Xining, the energy consumption of Tianjin ICEV was higher by 3.7%,

PHEV by 11.3%, and BEV lower by 10.9%. In Tianjin, PHEV energy consumption was 83.6% of ICEV, and BEV consumption was 26.5% of PHEV and 22.2% of ICEV. In Xining, PHEV consumption was 76.4% of ICEV, and BEV consumption was 33.8% of PHEV and 25.8% of ICEV. Nowadays, electrification has gradually become one of the main policies for vehicle energy saving and emission reduction. Our framework is universal and can be directly used to evaluate the energy consumption benefits when promoting electric vehicles in new regions. By collecting relevant data and using the approach based on Markov chain databases and high-precision prediction models, robust regional energy consumption benefits can be provided, thereby offering a scientific basis for policymakers when formulating related policies. Our research shows that the electrification of transportation plays an important role in reducing urban air pollution and alleviating road traffic energy demand. Promoting different vehicle types in different regions will also yield different energy benefits. However, we recommend that the method used in this study should only be applied in cases where the extraction of elevation information is necessary. If elevation is not a relevant factor, the 3D Markov Chain method employed in this study may not be required, and alternative modeling methods could be used to better suit different scenarios. Besides, the current framework still has certain issues that need to be improved in future research: 1) The variety of models chosen within each vehicle type is limited, and it does not encompass vehicles with diverse engine models. The current study only selected a few representative vehicles and cannot cover all vehicle types and engine models available in the market. Different vehicle models and engine

types may have significantly different energy consumption characteristics, and the existing model cannot fully represent this diversity. Using a generalized energy consumption model may smooth out the differences between different models. Future research should consider including a wider range of vehicle models and engine data to explore the performance of more models in different regions. Additionally, how to address the impact of different engine models on the model output results (the fusion of datasets with different patterns may lead to data mixing, thereby affecting the accuracy of the model output results) requires further study. 2) The number of DCs in the DC database needs mathematical proof. In this study, the convergence of DiffSum was indeed observed; however, the exact number of DC entries in the operating condition database required for convergence, as well as the manner of convergence, needs to be rigorously proven mathematically. The lack of a clear mathematical proof may lead to a shallow understanding of the model's convergence behavior, potentially affecting its stability and applicability in real-world scenarios. If the number of DC entries in the operating condition database varies significantly, it may impact the reliability and consistency of the model results. Future research should delve deeper into the theoretical conditions for DiffSum convergence, particularly the relationship between the number of DCs and convergence, and provide a mathematical proof. Furthermore, experiments or simulations could be considered to explore the impact of varying the number of DCs on model convergence and result stability. How to better save computational resources while ensuring the robustness of results is an important issue to be resolved in the future.

Supplementary information.

The online version contains supplementary material available in file Supplement File.

Data availability

The data generated and/or analyzed during the current study are not publicly available for legal/ethical reasons but are available (as an aggregate data) from the corresponding author upon reasonable request.

Code availability

The code used in the analysis was developed in Python version 3.8.2 and is not publicly available for legal/ethical reasons. However, it can be made available upon reasonable request from the corresponding author.

Acknowledgements

This work was supported by the National Key Research and Development Program of China (2024YFE0112100, 2023YFC3705405), the National Natural Science Foundation of China (42107114, 42177084) and the Fundamental Research Funds for the Central Universities of China (63241318, 63241322).

Author Contribution

Z. J. and Z. C. contributed equally to this work. Z. J. and Z. C. conceived and designed the study in consultation with J. Y.; Z. J. and Z. C. collected the data, implemented the model, and created the visualizations. L. W. and Y. Z. processed the data and performed the formal analysis. N. W. contributed to the methodology. Z. J. developed the software. Q. Z. contributed to the review and editing of the manuscript.

H. M. acquired the funding. Z. J., Z. C. wrote the original manuscript with contributions from all co-authors. All authors reviewed the results and approved the final version of the manuscript.

Competing interests

The authors declare no competing financial interest

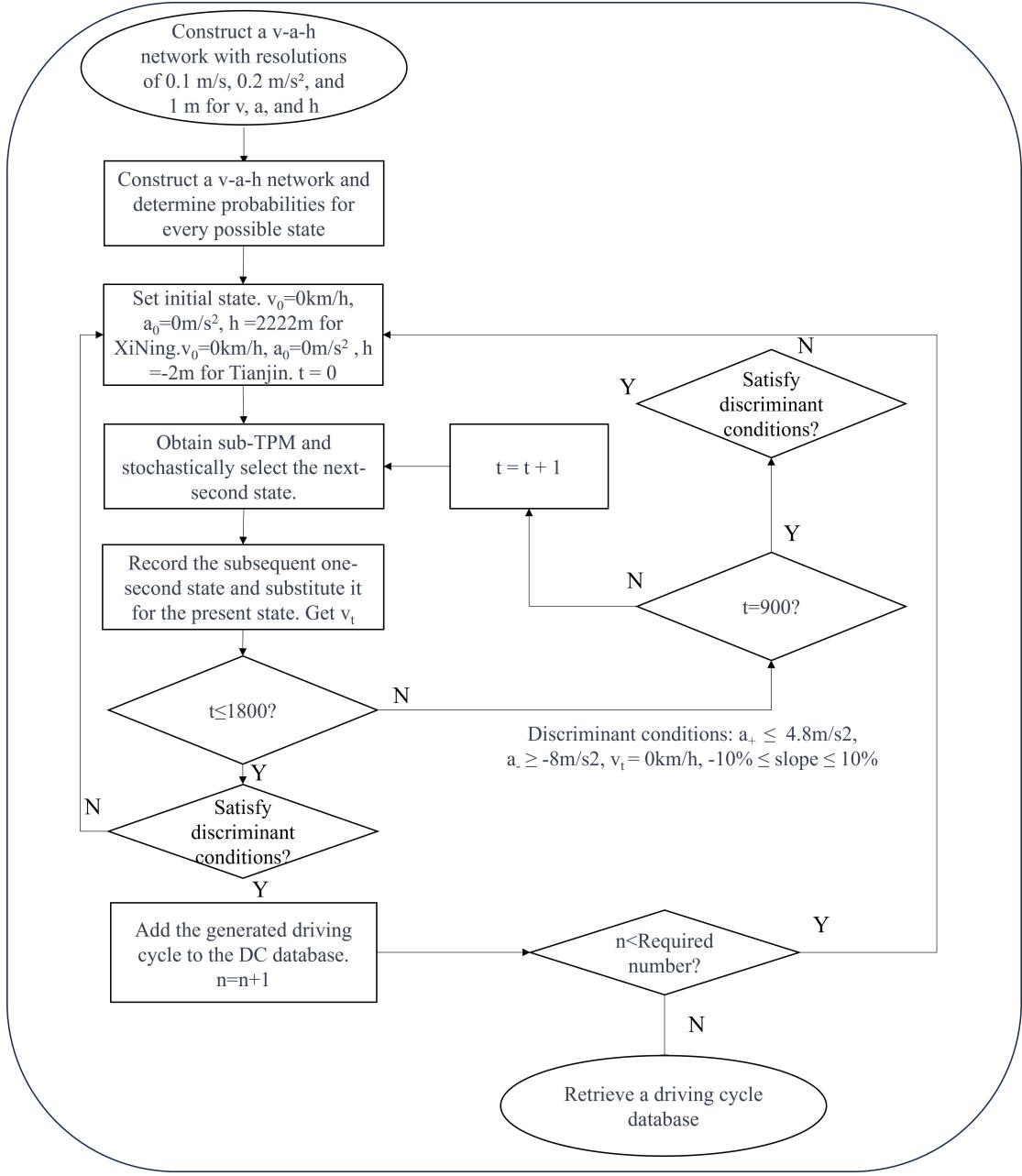
REFERENCES

- 1 Jia, Z. *et al.* Regional vehicle energy consumption evaluation framework to quantify the benefits of vehicle electrification in plateau city: A case study of Xining, China. *Applied Energy* **377**, 124626, doi:<https://doi.org/10.1016/j.apenergy.2024.124626> (2025).
- 2 IEA. *Global electric car stock*, <<https://www.iea.org/data-and-statistics/charts/global-electric-car-stock-2010-2022>> (2010-2022).
- 3 IEA. *World Energy Outlook 2022*, <<https://www.iea.org/reports/world-energy-outlook-2022>> (2022).
- 4 Requia, W. J., Mohamed, M., Higgins, C. D., Arain, A. & Ferguson, M. How clean are electric vehicles? Evidence-based review of the effects of electric mobility on air pollutants, greenhouse gas emissions and human health. *Atmospheric Environment* **185**, 64-77 (2018).
- 5 Hu, X., Chen, N., Wu, N. & Yin, B. The potential impacts of electric vehicles on urban air quality in Shanghai City. *Sustainability* **13**, 496 (2021).
- 6 Yu, A., Wei, Y., Chen, W., Peng, N. & Peng, L. Life cycle environmental impacts and carbon emissions: A case study of electric and gasoline vehicles in China. *Transportation Research Part D: Transport and Environment* **65**, 409-420, doi:<https://doi.org/10.1016/j.trd.2018.09.009> (2018).
- 7 Guo, X., Sun, Y. & Ren, D. Life cycle carbon emission and cost-effectiveness analysis of electric vehicles in China. *Energy for Sustainable Development* **72**, 1-10 (2023).
- 8 Buberger, J. *et al.* Total CO₂-equivalent life-cycle emissions from commercially available passenger cars. *Renewable and Sustainable Energy Reviews* **159**, 112158, doi:<https://doi.org/10.1016/j.rser.2022.112158> (2022).
- 9 Buekers, J., Van Holderbeke, M., Bierkens, J. & Int Panis, L. Health and environmental benefits related to electric vehicle introduction in EU countries. *Transportation Research Part D: Transport and Environment* **33**, 26-38, doi:<https://doi.org/10.1016/j.trd.2014.09.002> (2014).
- 10 Tobollik, M. *et al.* Health impact assessment of transport policies in Rotterdam: Decrease of total traffic and increase of electric car use. *Environmental Research* **146**, 350-358, doi:<https://doi.org/10.1016/j.envres.2016.01.014> (2016).
- 11 Liu, Y. *et al.* Development of China light-duty vehicle test cycle. *International Journal of Automotive Technology* **21**, 1233-1246 (2020).
- 12 Yu, L., Zhi-xin, W. & Meng-liang, L. Development of China Light-duty Passenger Car Test Cycle. *Journal of Transportation Systems Engineering and Information Technology* **20**, 233 (2020).

- 13 Tutuiaru, M. *et al.* Development of the World-wide harmonized Light duty Test Cycle (WLTC) and a possible pathway for its introduction in the European legislation. *Transportation research part D: transport and environment* **40**, 61-75 (2015).
- 14 Nyberg, P., Frisk, E. & Nielsen, L. Driving cycle equivalence and transformation. *IEEE Transactions on Vehicular Technology* **66**, 1963-1974 (2016).
- 15 Mafi, S. *et al.* Developing local driving cycle for accurate vehicular CO₂ monitoring: A case study of Tehran. *Journal of Cleaner Production* **336**, 130176 (2022).
- 16 Wang, H., Zhang, X. & Ouyang, M. Energy consumption of electric vehicles based on real-world driving patterns: A case study of Beijing. *Applied energy* **157**, 710-719 (2015).
- 17 Travasset-Baro, O., Rosas-Casals, M. & Jover, E. Transport energy consumption in mountainous roads. A comparative case study for internal combustion engines and electric vehicles in Andorra. *Transportation Research Part D: Transport and Environment* **34**, 16-26 (2015).
- 18 Al-Wreikat, Y., Serrano, C. & Sodré, J. R. Driving behaviour and trip condition effects on the energy consumption of an electric vehicle under real-world driving. *Applied Energy* **297**, 117096 (2021).
- 19 Yang, D., Liu, H., Li, M. & Xu, H. Data-driven analysis of battery electric vehicle energy consumption under real-world temperature conditions. *Journal of Energy Storage* **72**, 108590 (2023).
- 20 Hao, X., Wang, H., Lin, Z. & Ouyang, M. Seasonal effects on electric vehicle energy consumption and driving range: A case study on personal, taxi, and ridesharing vehicles. *Journal of Cleaner Production* **249**, 119403 (2020).
- 21 Zou, Y., Wei, S., Sun, F., Hu, X. & Shiao, Y. Large-scale deployment of electric taxis in Beijing: A real-world analysis. *Energy* **100**, 25-39, doi:10.1016/j.energy.2016.01.062 (2016).
- 22 Wu, X., Freese, D., Cabrera, A. & Kitch, W. A. Electric vehicles' energy consumption measurement and estimation. *Transportation Research Part D: Transport and Environment* **34**, 52-67 (2015).
- 23 Jia, Z. *et al.* Energy saving and emission reduction effects from the application of green light optimized speed advisory on plug-in hybrid vehicle. *Journal of Cleaner Production* **412**, 137452 (2023).
- 24 Lin, J. & Niemeier, D. A. An exploratory analysis comparing a stochastic driving cycle to California's regulatory cycle. *Atmospheric Environment* **36**, 5759-5770 (2002).
- 25 Kamble, S. H., Mathew, T. V. & Sharma, G. K. Development of real-world driving cycle: Case study of Pune, India. *Transportation Research Part D: Transport and Environment* **14**, 132-140 (2009).
- 26 Mayakuntla, S. K. & Verma, A. A novel methodology for construction of driving cycles for Indian cities. *Transportation research part D: transport and environment* **65**, 725-735 (2018).
- 27 Yuhui, P., Yuan, Z. & Huibao, Y. Development of a representative driving cycle for urban buses based on the K-means cluster method. *Cluster Computing* **22**, 6871-6880 (2019).
- 28 Tao, S., Ding, K., Li, Z. & Zhang, H. Development of a representative driving cycle for evaluating exhaust emission and fuel consumption for Chinese switcher locomotives. *Applied Energy* **322**, 119499, doi:<https://doi.org/10.1016/j.apenergy.2022.119499> (2022).
- 29 Sisodia, D., Singh, L., Sisodia, S. & Saxena, K. Clustering techniques: a brief survey of different clustering algorithms. *International Journal of Latest Trends in Engineering and Technology*

- (IJLTET) **1**, 82-87 (2012).
- 30 Peng, J., Pan, D. & He, H. in *2015 International Conference on Intelligent Systems Research and Mechatronics Engineering*. 1491-1497 (Atlantis Press).
- 31 Qin, S., Qiu, D. & Zhou, J. Driving cycle construction and accuracy analysis based on combined clustering technique. *Automotive Engineering* **34**, 164-169 (2012).
- 32 Wei, N. et al. Standard environmental evaluation framework reveals environmental benefits of green light optimized speed advisory: A case study on plug-in hybrid electric vehicles. *Journal of Cleaner Production* **404**, 136937 (2023).
- 33 Gong, H. et al. Generation of a driving cycle for battery electric vehicles: A case study of Beijing. *Energy* **150**, 901-912 (2018).
- 34 Zhao, X., Ye, Y., Ma, J., Shi, P. & Chen, H. Construction of electric vehicle driving cycle for studying electric vehicle energy consumption and equivalent emissions. *Environmental Science and Pollution Research* **27**, 37395-37409 (2020).
- 35 Jing, Z., Wang, G., Zhang, S. & Qiu, C. Building Tianjin driving cycle based on linear discriminant analysis. *Transportation Research Part D: Transport and Environment* **53**, 78-87 (2017).
- 36 Brady, J. & O'Mahony, M. Development of a driving cycle to evaluate the energy economy of electric vehicles in urban areas. *Applied Energy* **177**, 165-178, doi:10.1016/j.apenergy.2016.05.094 (2016).
- 37 Wang, T., Jing, Z., Zhang, S. & Qiu, C. Utilizing principal component analysis and hierarchical clustering to develop driving cycles: a case study in Zhenjiang. *Sustainability* **15**, 4845 (2023).
- 38 Ho, C. S. et al. Optical properties of vehicular brown carbon emissions: Road tunnel and chassis dynamometer tests. *Environmental Pollution* **320**, 121037 (2023).
- 39 Statistics, T. M. B. o. *Statistical Bulletin on the National Economic and Social Development of Tianjin Municipality in 2020*, https://stats.tj.gov.cn/tjsj_52032/tjgb/202103/t20210317_5386752.html (2021).
- 40 Prieto-Curiel, R., Patino, J. E. & Anderson, B. Scaling of the morphology of African cities. *Proceedings of the National Academy of Sciences* **120**, e2214254120 (2023).
- 41 Province, T. p. s. G. o. Q. *Xining City's motor vehicle ownership is approaching the one-million mark, with nearly one out of every three individuals possessing a motor vehicle and a corresponding driver's license.*, <http://www.qinghai.gov.cn/zwgk/system/2023/10/12/030027217.shtml> (2023).
- 42 Boroujeni Behdad, Y., Frey, H. C. & Sandhu Gurdas, S. Road Grade Measurement Using In-Vehicle, Stand-Alone GPS with Barometric Altimeter. *Journal of Transportation Engineering* **139**, 605-611, doi:10.1061/(ASCE)TE.1943-5436.0000545 (2013).
- 43 Chen, T. & Guestrin, C. in *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*. 785-794.
- 44 Jimenez-Palacios, J. L. Understanding and quantifying motor vehicle emissions with vehicle specific power and TILDAS remote sensing. *Massachusetts Institute of Technology* (1998).
- 45 Bao, F. B. Research on calculation method of fuel consumption rate base on engine universal characteristics. *Advanced Materials Research* **239**, 1490-1494 (2011).

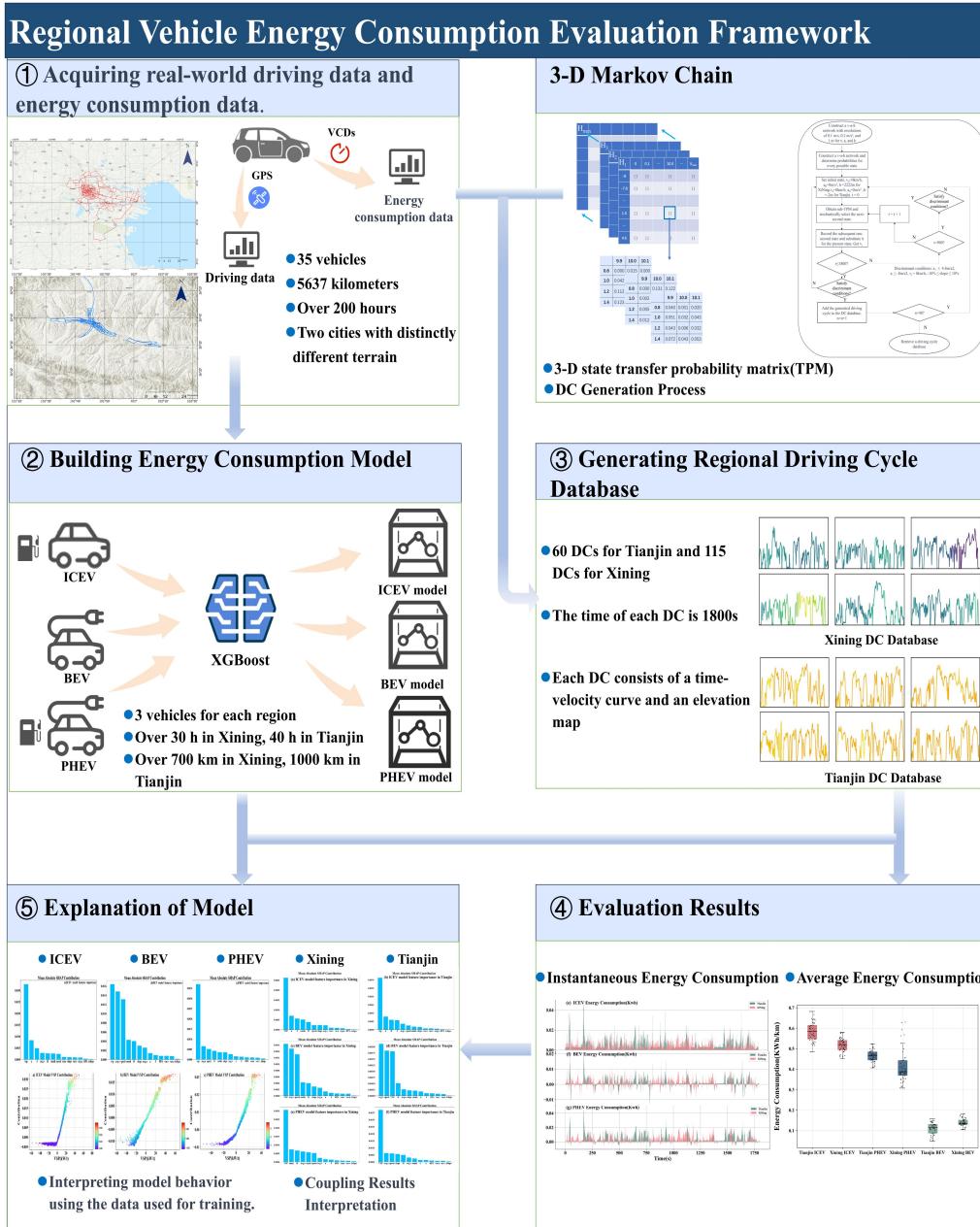
Figure 1. The process of building a DC Database.



The figure illustrates the complete process of constructing a driving cycle database based on a velocity-acceleration-elevation (v-a-h) network. First, the v-a-h network is established, and transition probabilities for all possible states are determined. An initial state is set, and subsequent states are stochastically selected using a sub-transition probability matrix (sub-TPM). The generated driving cycle is recorded and evaluated against predefined discriminant conditions. If the conditions are met, the cycle is added to the database. This process repeats until the required number of cycles is reached or 1800 seconds are exceeded. A final check adds the last generated cycle if needed.

number of driving cycles is obtained, after which the final driving cycle database is retrieved.

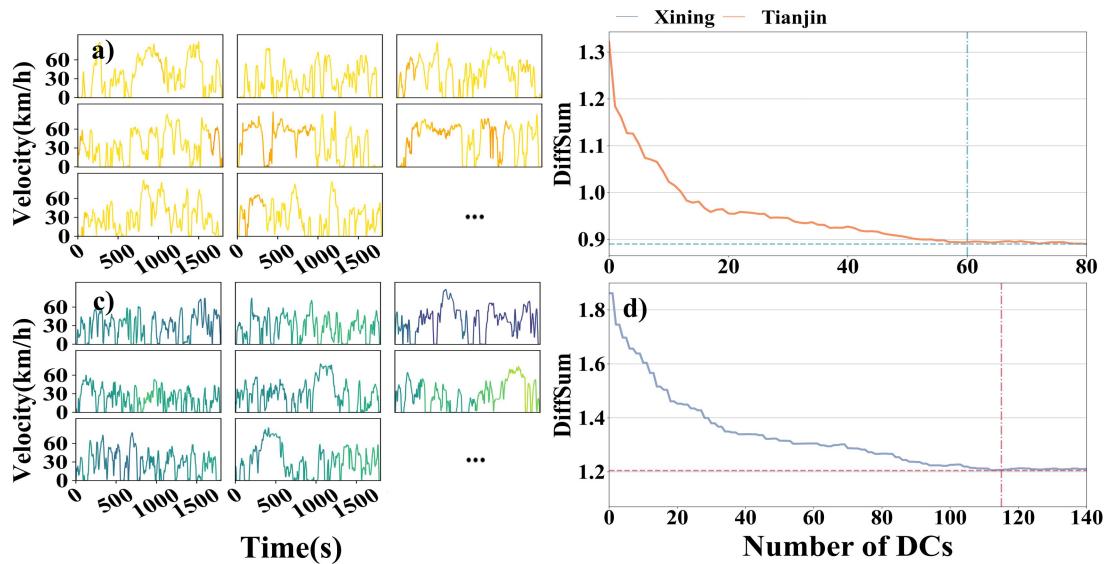
Figure 2. Regional vehicle energy consumption evaluation framework.



The figure illustrates the complete process of the energy consumption framework. First, energy consumption data and driving data are collected. The driving data is fed into a three-dimensional Markov chain to establish a regional DC database. The energy consumption data is used to

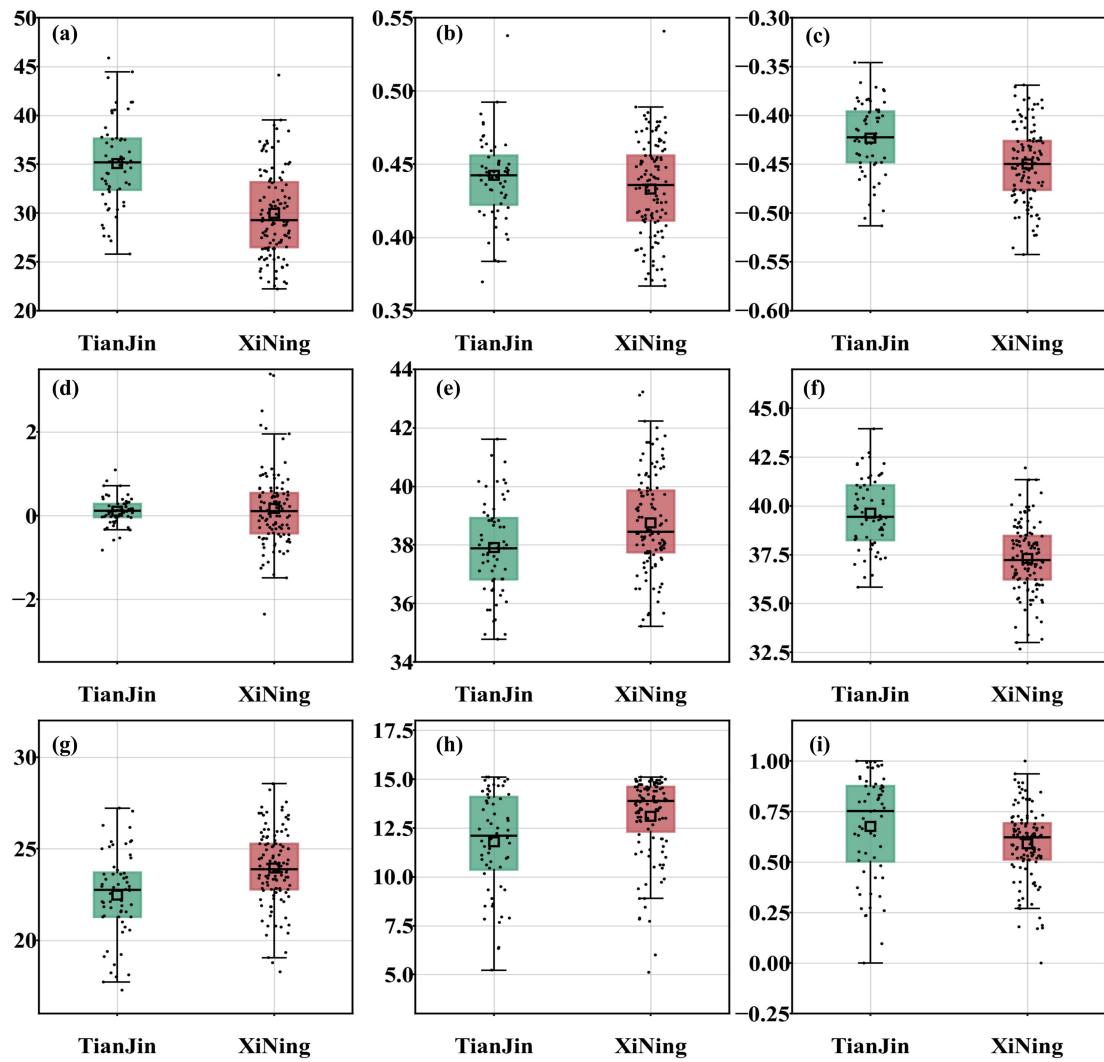
develop a machine learning model, which is interpreted using the SHAP model interpretation framework. Finally, the DC database and the machine learning model are coupled, and the SHAP framework is applied again to interpret the coupling results.

Figure 3. DC databases for Xining and Tianjin, and the variation of DiffSum with the number of DCs.



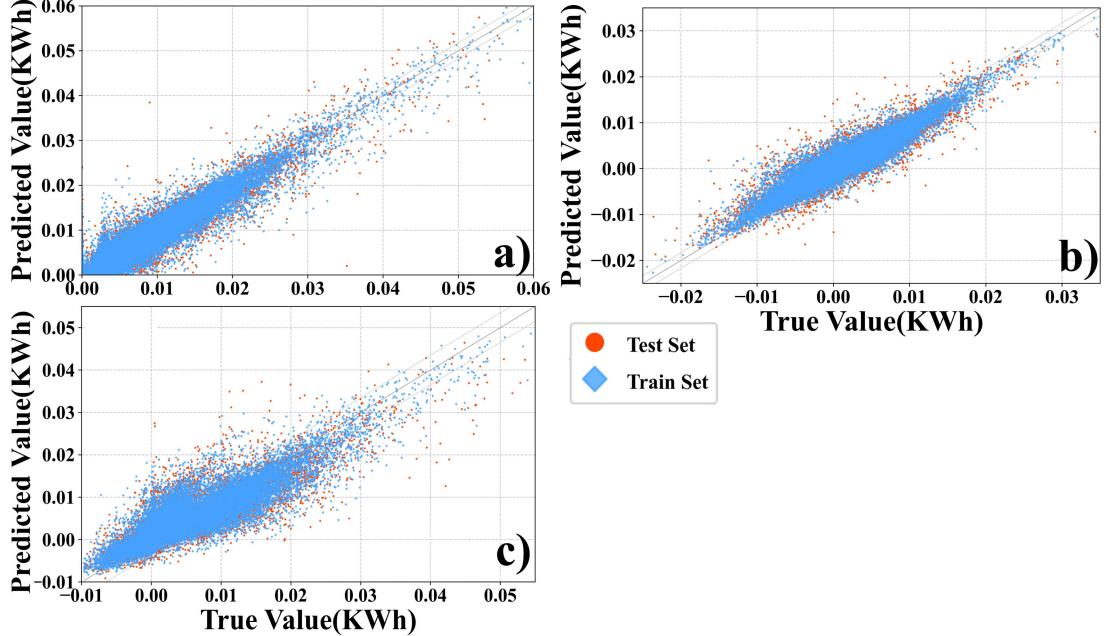
(a) Tianjin's DC database. (b) The trend of DiffSum variation in Tianjin. (c) Xining's DC database. (d) The trend of DiffSum variation in Xining. Tianjin's database converges in 60. Xining's database converges in 115. In each DC, the horizontal axis represents time, totaling 1800 s. The vertical axis represents speed (km/h), and the color mapping represents elevation (m). Detailed DC databases for each region are depicted in Supplementary Figure 6.

Figure 4. The distribution of 9 metrics in Figure 4, including both average and ratio indicators.



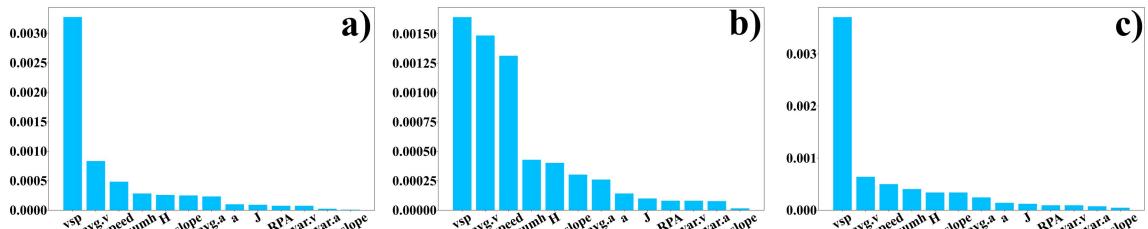
(a) Average Speed (km/h). (b) Average Acceleration (m/s²). (c) Average Deceleration (m/s²). (d) Average Slope (%). (e) Accelerated Time Ratio (%). (f) Deceleration Time Ratio (%). (g) Constant Speed Time Ratio (%). (h) Idle Time Ratio (%). (i) Standardized Elevation. Green represents Tianjin, red represents Xining, with the vertical axis of the squares indicating the corresponding means.

Figure 5. The energy consumption predicted values and true values of machine learning models for three types of vehicles.



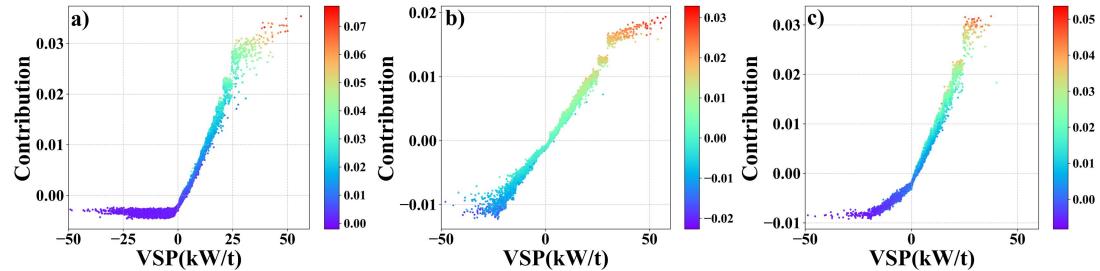
(a) The model performance for ICEV. ICEV model's R^2 is 0.92, and the RMSE is 0.0018. (b) The model performance for BEV. BEV model's R^2 is 0.83, and the RMSE is 0.0017. (c) The model performance for PHEV. PHEV model's R^2 is 0.74, and the RMSE is 0.0036. The solid line represents a slope of 1, while the dashed lines represent $\pm \text{RMSE}$. Red represents the points in the test set, while blue represents the points in the training set.

Figure 6. The average feature importance of the three models.



(a) ICEV's mean absolute SHAP contribution. (b) BEV's mean absolute SHAP contribution. (c) PHEV's mean absolute SHAP contribution.

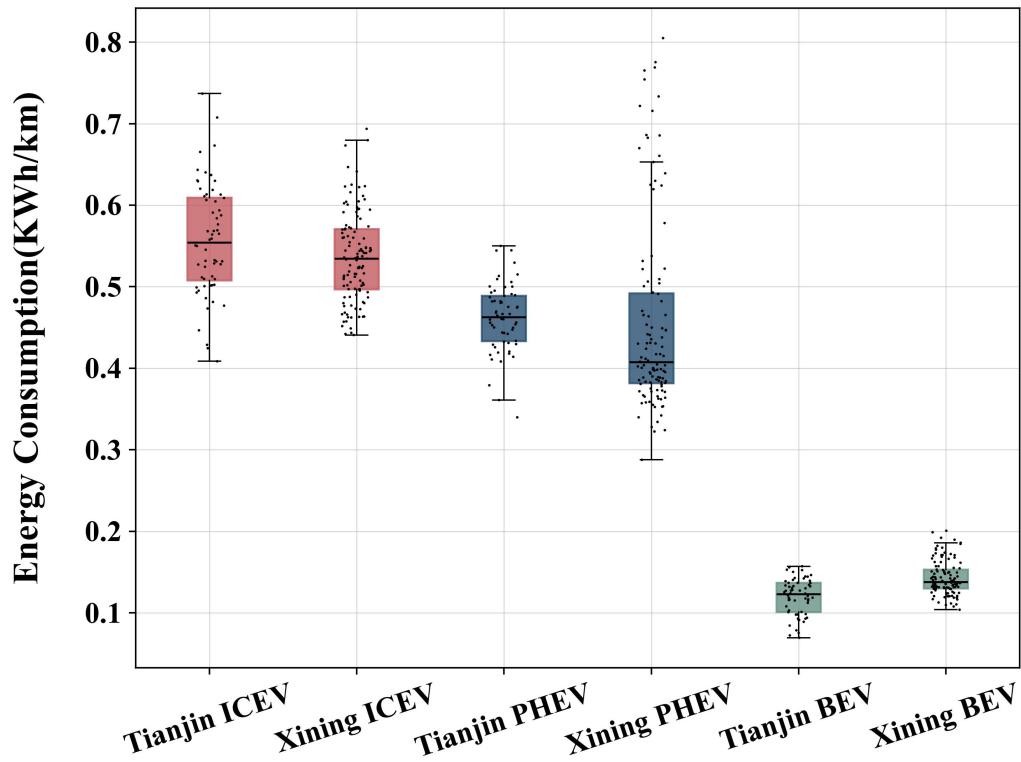
Figure 7. Marginal Contribution of VSP for three types of vehicles' model.



(a) The marginal contribution of VSP for ICEV. (b) The marginal contribution of VSP for BEV. (c)

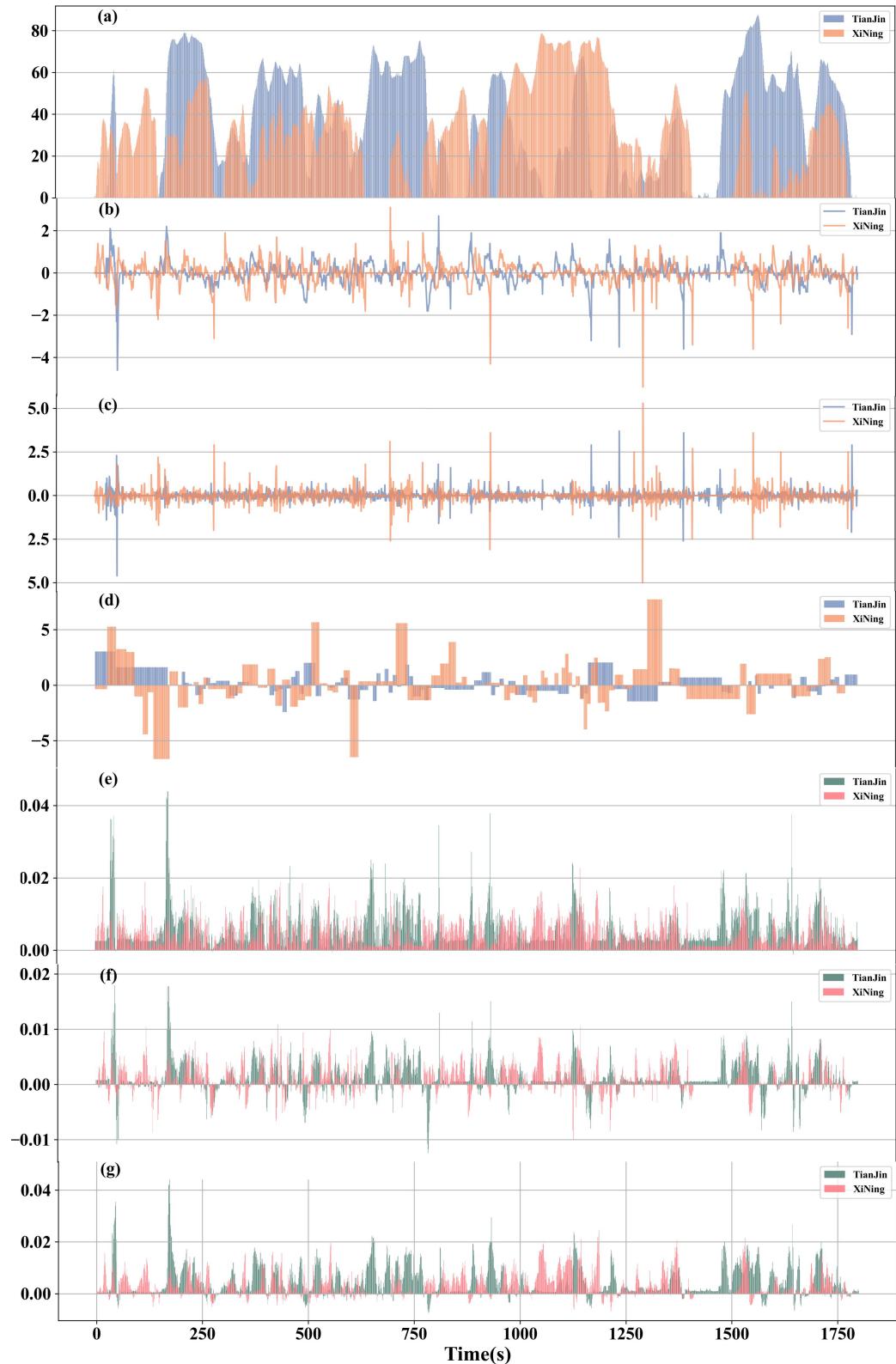
The marginal contribution of VSP for PHEV.

Figure 8. The overall energy consumption distribution of different vehicle types under the DC databases of Tianjin and Xining.



The black scatter points represent the energy consumption predictions for each DC, while the box plot represents the overall distribution of these predictions.

Figure 9. One DC's kinematic characteristics and instantaneous energy consumption.



(a) Velocity. (b) Acceleration. (c) Jerk. (d) Slope. (e) ICEV Energy Consumption. (f) Energy

Consumption. (g) Energy Consumption.

Table 1. Detailed vehicle and energy consumption information

	Volkswagen Tango 2019	Volkswagen Magotan 280	Volkswagen ID.4 X 2022	Volkswagen ID.4 cross 2022	Volkswagen Passat PHEV 2020	Volkswagen Magotan GTE 2022
City	Xining	Tianjin	Xining	Tianjin	Xining	Tianjin
Fuel Type	Gasoline 1.4T	Gasoline 1.4T	Electricity	Electricity	Hybrid 1.4T	Hybrid 1.4T
Engine	4-cylinder GDI	4-cylinder GDI	—	—	4-cylinder GDI	4-cylinder GDI
			83.4 kWh	84.8 kWh	13 kWh	13 kWh
Battery	—	—	Ternary Li- ion Battery	Ternary Li- ion Battery	Ternary Li- ion Battery	Ternary Li- ion Battery
Curb weight(kg)	1404	1410	2120	2130	1750	1750
Test Cycle	WLTC	WLTC	CLTC	CLTC	WLTC	WLTC
Fuel consumption per 100km(L)	6.4	6.25	—	—	2.23	2.29
Electricity consumption per 100 km(kWh)	—	—	14	14.3	16	15.1
Total energy consumption per 100 km(kWh)	55.1	53.8	14	14.3	35.2	34.8