



# A Preference-aware Meta-optimization Framework for Personalized Vehicle Energy Consumption Estimation

Siqi Lai

The Hong Kong University of Science  
and Technology (Guangzhou)  
siqlai@hkust-gz.edu.cn

Weijia Zhang

The Hong Kong University of Science  
and Technology (Guangzhou)  
wzhang411@connect.hkust-  
gz.edu.cn

Hao Liu\*

The Hong Kong University of Science  
and Technology (Guangzhou)  
The Hong Kong University of Science  
and Technology  
liuh@ust.hk

## ABSTRACT

Vehicle Energy Consumption (VEC) estimation aims to predict the total energy required for a given trip before it starts, which is of great importance to trip planning and transportation sustainability. Existing approaches mainly focus on extracting statistically significant factors from typical trips to improve the VEC estimation. However, the energy consumption of each vehicle may diverge widely due to the personalized driving behavior under varying travel contexts. To this end, this paper proposes a preference-aware meta-optimization framework (META-Pec) for personalized vehicle energy consumption estimation. Specifically, we first propose a spatiotemporal behavior learning module to capture the latent driver preference hidden in historical trips. Moreover, based on the memorization of driver preference, we devise a selection-based driving behavior prediction module to infer driver-specific driving patterns on a given route, which provides additional basis and supervision signals for VEC estimation. Besides, a driver-specific meta-optimization scheme is proposed to enable fast model adaptation by learning and sharing transferable knowledge globally. Extensive experiments on two real-world datasets show the superiority of our proposed framework against ten numerical and data-driven machine learning baselines. The source code is available at <https://github.com/usail-hkust/Meta-Pec>.

## CCS CONCEPTS

- Information systems → Spatial-temporal systems.

## KEYWORDS

energy consumption estimation, spatiotemporal prediction, driving preference learning, meta learning

## ACM Reference Format:

Siqi Lai, Weijia Zhang, and Hao Liu\*. 2023. A Preference-aware Meta-optimization Framework for Personalized Vehicle Energy Consumption Estimation. In *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD '23)*, August 6–10, 2023, Long Beach, CA, USA. ACM, New York, NY, USA, 11 pages. <https://doi.org/10.1145/3580305.3599767>

\* Corresponding author.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

KDD '23, August 6–10, 2023, Long Beach, CA, USA.

© 2023 Copyright held by the owner/author(s). Publication rights licensed to ACM. ACM ISBN 979-8-4007-0103-0/23/08...\$15.00  
<https://doi.org/10.1145/3580305.3599767>

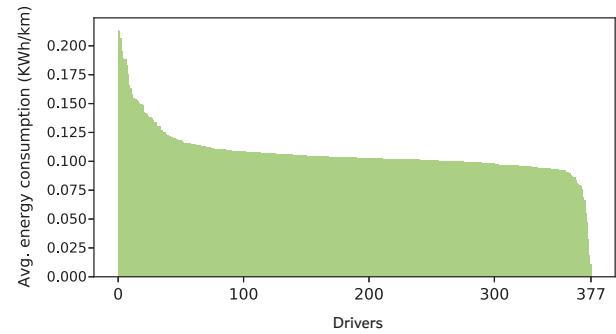


Figure 1: The driver-varying energy consumption distribution of a real-world dataset in Shenzhen, all 377 vehicles are of the same type.

## 1 INTRODUCTION

The rapid transportation network expansion and traffic demand growth raise public concerns about the efficiency, sustainability, and resilience of the urban transportation system. From 2015 to 2030, as reported by the United Nations<sup>1</sup>, the number of vehicles on the road is approximately double, and global traffic is likely to increase by 50%. Therefore, the accurate estimation of Vehicle Energy Consumption (VEC) is of great importance to the decision-making of urban governance and travel planning [24].

Prior studies on VEC estimation can be roughly divided into two categories: the *Numerical methods* [6, 8, 20] and the *Data-driven methods* [10, 14, 17, 18, 22]. Specifically, the numerical methods mainly focus on identifying factors that have the most significant influence on VEC. For example, Y. Al-Wreikat *et al.* [1] propose to partition drivers into multiple classes and quantify the average VECs. The data-driven methods, on the other hand, aim to automatically extract and utilize relevant knowledge by leveraging machine learning tools, e.g., linear regression (LR) [7], decision tree (DT) [25], support vector machine (SVM) [21], etc. Inspired by the recent advances of deep learning, deep neural networks such as Long-Short-Term Memory (LSTM) [4] and Transformer [31] have also been adopted to analyze road-level energy consumption. However, existing approaches make predictions based on hand-crafted statistical features, which overlook the personalized nature of varying driving behaviors under different travel contexts.

In this work, we investigate the personalized vehicle energy consumption estimation problem with a consideration of fine-grained

<sup>1</sup><https://www.un.org/sustainabledevelopment/progress-report/>

individual driving preferences hidden in historical trajectories. However, three major challenges arise towards this goal. First, the driving preference describes the driver’s long-term intention (e.g., overtaking, braking, changing lanes under different travel contexts), which is critical to the overall energy consumption. As depicted in Figure 1, the average energy consumption of different drivers may vary significantly, indicating the necessity of incorporating driver preference for VEC estimation. However, the existing approach [31] represents latent driver preference via handcrafted features, which lead to lossy driving preference preservation. Thus, how to capture the latent driving preference in an effective way is the first challenge. Second, the behavior of a driver under different travel contexts may also vary. Estimating the intentional driving behavior of a given route can provide additional signals to guide the learning direction of the prediction model. However, the historical trajectory is noisy and may in large-scale. How to quantify the personalized driving behavior on a target route based on historical data in a cost-effective manner is another challenge. Third, although a unified model can provide predictions for all drivers, the estimation for drivers with only a few trajectories may be biased and error-prone. The last challenge is how to share transferable knowledge between drivers so that to benefit the long-tail prediction.

To this end, in this paper, we propose a preference-aware meta-optimization framework, META-PEC, to deliver more effective personalized vehicle energy consumption estimation. Specifically, we first propose a driving preference learning module to capture the latent spatiotemporal preference of each driver hidden in high-dimensional historical trajectories in an end-to-end manner. Moreover, we devise a selection-based driving behavior prediction module to estimate the future behaviors of a driver on a given route. In particular, the predicted behaviors provide additional supervision signals for model learning by incorporating the information from similar historical trips. Furthermore, we propose a driver-specific meta-optimization scheme to allow fast model adaption to data-insufficient drivers, where the transferable knowledge is encoded in a global parameter initialization.

In summary, the major contributions of this paper are as follows. (1) We investigate the personalized vehicle energy consumption estimation problem, which is beneficial to various downstream applications, such as fuel-efficient trip planning and sustainable urban transportation system policy-making. (2) We propose a preference-aware meta-optimization framework to incorporate the latent driving behavior knowledge hidden in past trajectories. To our knowledge, this is the first work that utilizes meta-learning to tackle the cold-start problem in the VEC estimation task. (3) Extensive experiments on two real-world datasets demonstrate the superiority of META-PEC compared with ten numerical and data-driven state-of-the-art approaches.

## 2 PRELIMINARIES

This section introduces some important definitions and the formal problem statement.

**Road network.** The road network consists of a set of road segments  $E = \{e_1, e_2, \dots, e_N\}$  and intersection joints, where  $N$  is the number of road segments in the city. We use  $\mathbf{x}_i^e$  to denote the road segment  $i$ ’s features (e.g., length, number of lanes).

**Table 1: Statistics of datasets.**

Category	VED	ETTD
# of trips	25,661	18,546
# of drivers	348	377
Time span	11/1/2017-11/7/2018	10/22/2014
# of total distance	130,864 km	87,251 km
Avg. route length	4.81 km	5.10 km
Avg. speed	38.59 km/h	13.54 km/h

**Route.** A route  $R = [e_1, e_2, \dots, e_n]$  is a road segment sequence that a vehicle will traverse, where  $n$  is the number of road segments. We denote  $R_c$  as the route of the target trip and  $R_i$  ( $i \in \mathbb{N}^+$ ) as the route of a historical trip.

**Trajectory.** A trajectory  $T_i = [(p_j, t_j, \mathbf{x}_j^l, y_j)]_{j=1}^m$  is a sequence of sample points logged by GPS devices, where  $p_j$  is the distance the driver has traveled from the origin to the current location ( $p_1 = 0$ ),  $t_j$  is the time elapsed since the trip started ( $t_1 = 0$ ),  $\mathbf{x}_j^l$  are features describing the current state of the vehicle (e.g., speed, acceleration),  $y_j$  is the energy consumption from the origin to the current point  $j$  ( $y_1 = 0$ ), and  $m$  is the number of sample points. We take  $y = y_m$  as the ground truth label of the total energy consumption of the trip.

**Historical trips.** A driver  $u$ ’s historical trips is defined as  $H^u = \{(R_i, T_i)\}_{i=1}^M$ , where  $R_i$  is the route,  $T_i$  is the corresponding trajectory, and  $M$  is the total number of historical trips.

**Problem definition.** We define personalized vehicle energy consumption estimation as a supervised-learning task. Given a target route  $R_c$  and a driver  $u$ , we aim to predict the energy consumption of the trip based on  $u$ ’s historical trips  $H^u$ ,

$$f_\theta : (R_c, H^u) \rightarrow y, \quad (1)$$

where  $f_\theta$  is the model parameterized by  $\theta$  that we aim to learn,  $y$  is the ground truth energy consumption.

## 3 DATA DESCRIPTION AND ANALYSIS

In this section, we describe the datasets used for META-PEC with a primary data analysis.

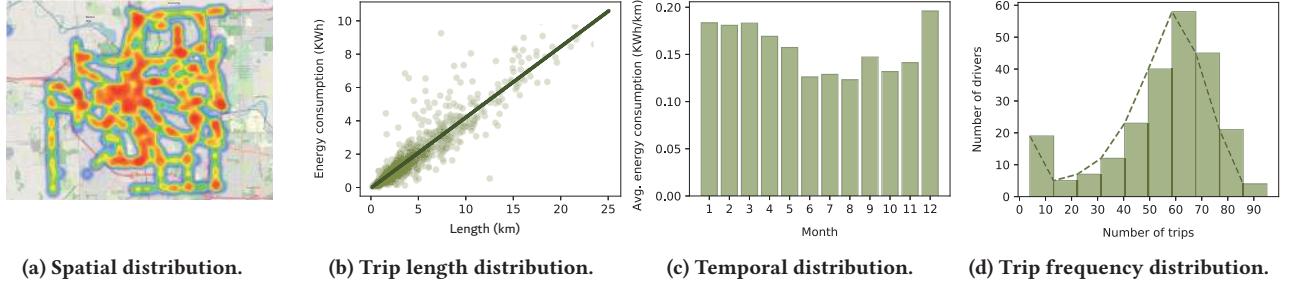
### 3.1 Data Description

We study our problem on two real-world datasets. The first is a large-scale energy usage dataset of diverse vehicles in Ann Arbor, Michigan, USA, known as the vehicle energy dataset (VED<sup>2</sup>). Another is an electric taxi trajectory dataset (ETTD<sup>3</sup>), collected from Shenzhen, Guangdong, China. The statistics of two datasets are summarized in Table 1.

The VED dataset [19] consists of vehicles’ trajectories and the corresponding dynamic factors (i.e., energy, speed, etc.) collected by the Second On-Board Diagnostics (OBD-II) logger. The dataset is ranged from Nov 2017 to Nov 2018, covering various driving scenarios and weather conditions. In total, VED contains trajectories of approximately 130,864 kilometers. The fleet comprises 348 vehicles (230 Internal Combustion Engine Vehicles (ICEVs), 91 Hybrid Electric Vehicles (HEVs), 24 Plug-in Hybrid Electric Vehicles (HEVs),

<sup>2</sup><https://github.com/gsoh/VED>

<sup>3</sup><http://guangwang.me/#/data>



**Figure 2: Distribution of the VED dataset:** (a) the spatial distribution of vehicle energy consumption on each road segment. (b) the trip distance-energy distribution. (c) the temporal distribution of energy consumption in a year. (d) the driver-specific trip frequency distribution.

and 3 Electric Vehicles (EVs). We regard the fuel consumption of ICEVs and HEVs, and the electricity cost of PHEVs and EVs as the ground truth energy consumption [14].

The ETTD dataset [28] contains trajectories of 377 electric taxis with 1,155,654 GPS records and speed profiles collected on Oct 22, 2014. The driving conditions range from approximately all scenarios city-wide. The total distance is about 87,251 kilometers. We consider the required mechanical energy at the wheel as the ground truth energy consumption [32] in the ETTD dataset.

### 3.2 Data Preprocessing

**3.2.1 Data Preparation.** We preprocess each dataset as follows. For each dataset, we split the raw trajectory into multiple trips if the driver has stopped for more than five minutes. The route of each trip is extracted by a map-matching algorithm [35]. We calculate the energy consumption on each GPS sample point by following [14]. Since the ETTD dataset only provides the speed and coordinate information, the VEC is calculated based on an estimation of the required mechanical energy at the wheel by following [32]. We aggregate the energy consumption of each GPS point as the ground truth of each trip.

**3.2.2 Data Anonymization.** The original datasets in our study did not include any identifiable driver information such as names, phone numbers, or any other personal details. To further protect the sensitive information of each driver, we mask each driver with an anonymized identifier.

### 3.3 Data Analysis

To help understand the VEC distribution, we conduct primary data analysis on the VED dataset. Overall, the energy consumption of each trip is influenced by various factors. First, Figure 2a plots the spatial distribution of energy consumption in Ann Arbor, where warmer color represents a higher energy consumption. We observe the VEC varies in different road segments of the city, and vehicles cost more energy in the downtown area of the city than in the suburbs. Figure 2b shows the positive correlation between trip length and energy consumption. Meanwhile, Figure 2c presents the varying energy consumption over the year, indicating a negative correlation between the energy consumption and temperature. Finally, the trip frequency of each driver is reported in Figure 2d. We observe a two-peak distribution where the second-highest peak

in the long-tail represents the large portion of drivers with only a few historical trips. Such observation inspires us to develop a meta-optimization scheme to alleviate the cold-start problem.

## 4 THE PROPOSED METHOD

The overall structure of our proposed META-Pec framework is illustrated in Figure 3, which mainly consists of four components. (1) *Driving preference learning*: it performs spatiotemporal behavior learning to extract the driver’s driving preferences from historical trips. (2) *Selection-based driving behavior prediction*: it forecasts the driver’s future driving behaviors on each road by jointly considering road conditions and driving preferences, providing the extra basis and supervision signals for VEC estimation. (3) *VEC estimation*: it predicts the VEC of the target trip. (4) *Driver-specific meta-optimization*: it learns the global parameter initialization and enables the model to fast adapt to each driver. We next present each component in detail.

### 4.1 Feature Construction

This section introduces the features we utilize based on the datasets mentioned above. Every feature vector will be projected into a  $d$  dimensional representation via an embedding layer before feeding into the META-Pec model.

**4.1.1 The trip feature.** Trip features provide prior knowledge about the trip, including month, departure time, and route length.

**4.1.2 The statistical driving behavior feature.** The statistical driving behavior features provide macroscopic personalized information. We extract them from the driver’s historical trips, including average speed, acceleration, VEC per hour, and VEC per kilometer. They are concatenated with the trip features and denoted as  $x^s$ . We also extract these features on each road of the target trip as the ground-truth labels for the selection-based driving behavior prediction module. The features on the road  $e_i$  are denoted as  $y^e_i$ .

**4.1.3 Vehicle state feature.** Vehicle state features describe the vehicle’s current condition in the trajectory, including local time, speed, acceleration, VEC per hour, and VEC per kilometer.

**4.1.4 Road feature.** Road features describe the characteristics of the target road, which include the road type, the one-way indicator,

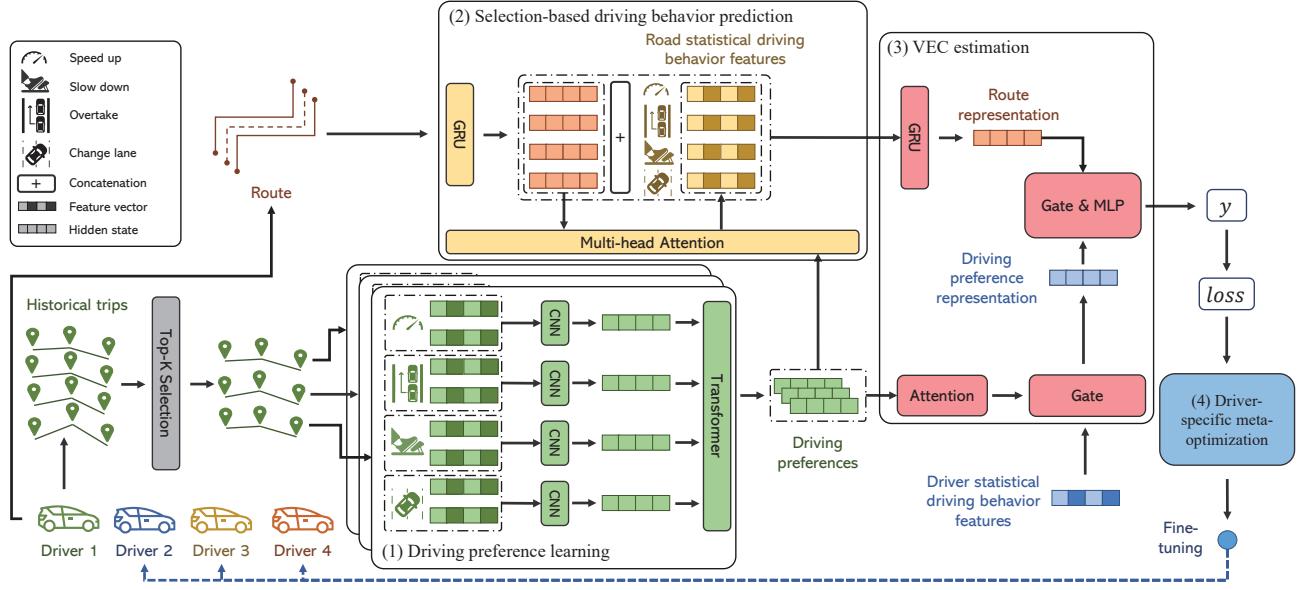


Figure 3: The framework overview of META-PEC.

the number of lanes, the allowed maximum speed, and the length of the road.

## 4.2 Driving Preference Learning

Instead of measuring the driver's personalized information only by handcrafting statistical features, we propose to exploit fine-grained trajectory-level driving preferences hidden in historical trips. As demonstrated in Section 3.3, VEC varies under different locations, time periods, and driver behaviors. We develop a spatiotemporal driving preference learning module to incorporate driving preference in an end-to-end way.

**4.2.1 Distance-time encoding.** We first encode the location and time information in each trajectory in a normalized scale to ease spatiotemporal learning. Inspired by the success of SEFT [13] and Transformer [26] in preserving the location relation via position encoding, we introduce distance-time encoding. Specifically, we use sine and cosine functions of different frequencies to convert the 1-dimensional data axis into a multi-dimensional input, which is defined as follows:

$$\text{Enc}_{2k}^*(a) = \sin\left(\frac{a}{\alpha^{2k/d}}\right), \quad (2)$$

$$\text{Enc}_{2k+1}^*(a) = \cos\left(\frac{a}{\alpha^{2k/d}}\right), \quad (3)$$

where  $\alpha$  is the maximum value that is expected in the data,  $k$  denotes the dimension, and  $*$  can be an arbitrary encoding target. For distance encoding  $\text{Enc}^{\text{dist}}(\cdot)$ ,  $a$  stands for the distance the driver has traveled from the origin to the current location. For time-encoding  $\text{Enc}^{\text{time}}(\cdot)$ ,  $a$  represents the time elapsed since the trip started. Then, we attach two types of encodings into the vector of the vehicle state features:

$$\mathbf{x}_i^l = \mathbf{x}_i^l + \text{Enc}^{\text{dist}}(p_i) + \text{Enc}^{\text{time}}(t_i). \quad (4)$$

**4.2.2 Driving behavior learning.** Then, we extract driving behaviors hidden in historical trajectories. We consider the driver's behaviors as a sequence of vehicle states, where each state  $\mathbf{x}_i^l$  at a time stamp describes an instant condition of the vehicle (e.g., current speed, acceleration, VEC per km, etc.). We derive the latent driver behavior representation by analogous the complicated behavior as a semantic object in an image [9].

Formally, we first split the vehicle state sequence of the historical trajectory  $\mathbf{X}^l = [\mathbf{x}_1^l, \mathbf{x}_2^l, \dots, \mathbf{x}_m^l]$  into multiple segments, then a convolutional neural network (CNN) is applied to embedding the driving behavior,

$$\mathbf{Z}[i] = \text{SELU}(\mathbf{X}^l[q \cdot (i-1) + 1 : q \cdot i] \odot \mathbf{F} + \mathbf{b}), \quad (5)$$

where  $q < m$  is the number of states contained in a segment,  $\mathbf{Z}[i]$  is the learned representation of the extracted behavior at step  $i$ ,  $\mathbf{Z}$  is the sequence of driving behaviors in each step,  $\mathbf{F} \in \mathbb{R}^{q \times d}$  is the filter,  $\odot$  denotes the convolution operation,  $\text{SELU}$  is chosen as the activation function, and  $\mathbf{b} \in \mathbb{R}^d$  is a learnable parameter. We further employ a Transformer encoder to derive the unified driver preference representation  $\mathbf{z}$ ,

$$\mathbf{z} = \text{TransformerEncoder}(\mathbf{Z}). \quad (6)$$

## 4.3 Selection-based Driving Behavior Prediction

Based on the driving preferences extracted above, we further infer the fine-grained trajectory, i.e., the detailed driving behavior such as acceleration and speed on the target trip, to provide the additional basis and supervision signals for the VEC estimator. Instead of predicting driving behaviors solely based on the road segment sequence, we propose to selectively reference similar routes of the target one to align the input (i.e., the feature vector) at training and

inference time. The selective approach also incorporates personalized information without introducing the significant computational overhead and extra noise.

**4.3.1 Top- $K$  historical trip selection.** We propose to select top- $K$  historical trips to ease the driving behavior prediction. Specifically, we construct two lightweight score functions to measure the distance between two trips. The first distance function is based on the overlap of road segments,

$$\text{score}_{c,i}^{\text{route}} = |R_c \cap R_i|, \quad (7)$$

where  $\cap$  is the *intersection* operation. Meanwhile, we calculate the temporal similarity,

$$\text{score}_{c,i}^{\text{time}} = |s_c - s_i|, \quad (8)$$

where  $s_c$  and  $s_i$  are the departure time of the target and the historical trip, respectively. Finally, we calculate the overall score of the historical trip,

$$\text{norm}(\text{score}_{c,i}^*) = \frac{\text{score}_{c,i}^* - \text{Min}(\text{score}_c^*)}{\text{Max}(\text{score}_c^*) - \text{Min}(\text{score}_c^*)}, \quad (9)$$

$$\text{score}_{c,i} = \text{norm}(\text{score}_{c,i}^{\text{route}}) - \text{norm}(\text{score}_{c,i}^{\text{time}}), \quad (10)$$

where  $\text{score}_c^*$  denotes the similarity scores of the target trip compared with all historical trips of the driver. We select  $K$  trips with the highest scores to extract the driver's driving preference. Note the reference trips may be less than  $K$  for long-tail drivers with only a few historical trips. Please refer to Appendix A.1 for the complete top- $K$  historical trip selection algorithm.

**4.3.2 Driving behavior prediction.** For driving behavior prediction, we first employ GRU [5], a variant of the recurrent neural network, to encode contextual information of the road segment sequence in the target route:

$$\mathbf{h}_i^e = \text{GRU}_1(\mathbf{h}_{i-1}^e, \mathbf{x}_i^e; \Omega_1), \quad (11)$$

where  $\Omega_1$  is parameters of GRU. With selected historical trips, then a multi-head attention module followed by a fully connected layer (FC) is applied to predict driving behaviors on each road segment:

$$\beta_{ij} = \text{softmax}\left(\frac{(\mathbf{W}_Q \mathbf{h}_i^e)^\top \mathbf{W}_K \mathbf{z}_j}{\sqrt{d}}\right), \quad (12)$$

$$\text{head} = \sum_{j=1}^K \beta_{ij} (\mathbf{W}_V \mathbf{z}_j), \quad (13)$$

$$\hat{\mathbf{y}}_i^e = \text{FC}(\text{head}_1 \parallel \dots \parallel \text{head}_h), \quad (14)$$

where  $\hat{\mathbf{y}}_i^e$  is the predicted statistical driving behavior features on the road  $e_i$ ,  $\mathbf{z}_j$  is the driving preference shown in the historical trip  $j$ ,  $\mathbf{W}_Q, \mathbf{W}_K, \mathbf{W}_V \in \mathbb{R}^{d \times d}$  are learnable parameters,  $h$  is the number of heads, and  $\parallel$  denotes the concatenation operation.

#### 4.4 VEC Estimation

Based on the five categories of features we obtained above, *i.e.*, the trip features, the statistical driving behavior features, the road features, the driving preferences extracted in section 4.2, and the target trip's driving behaviors predicted in section 4.3, we estimate the VEC of the target trip.

Firstly, we fuse all the driver's historical driving preference representations by an attention function as follows:

$$\mu_i = \text{softmax}(\mathbf{W}_z(\mathbf{x}^s * \mathbf{z}_i) + \mathbf{b}), \quad (15)$$

$$\mathbf{h}^z = \sum_{i=1}^K \mu_i \mathbf{z}_i, \quad (16)$$

where  $\mathbf{h}^z$  is the representation of driving preference extracted from top- $K$  historical trips,  $\mathbf{W}_z \in \mathbb{R}^{d \times d}$  and  $\mathbf{b} \in \mathbb{R}^d$  are learnable parameters, and  $*$  denotes element-wise multiplication.

For the road segment sequence, we concatenate each road segment's embedding vector with its corresponding predicted statistical driving behavior features as  $\hat{\mathbf{x}}_i^e = [\mathbf{x}_i^e \parallel \hat{\mathbf{y}}_i^e]$ . Then, we adopt another GRU to encode the updated contextual relations among roads:

$$\hat{\mathbf{h}}_i^e = \text{GRU}_2(\hat{\mathbf{h}}_{i-1}^e, \hat{\mathbf{x}}_i^e; \Omega_2). \quad (17)$$

We take the last hidden state  $\hat{\mathbf{h}}_n^e$  from (17) as the representation of the target route.

Afterward, we propose a gating mechanism to fuse the representation of driving preference, the statistical driving behavior features, and the target route,

$$\mathbf{h} = (\mathbf{h}^z \star \mathbf{x}^s) \star \hat{\mathbf{h}}_n^e, \quad (18)$$

where  $\mathbf{h}$  is the representation of the target route, and  $\star$  denotes the parameterized gating mechanism.

Finally, we estimate the VEC of the target route by feeding  $\mathbf{h}$  into a multi-layer perception (MLP), then multiply it with the vehicle type embedding  $\mathbf{W}_{tp} \in \mathbb{R}^d$ :

$$\mathbf{h} = \text{MLP}(\mathbf{h}), \quad (19)$$

$$\hat{y} = \mathbf{W}_{tp}^\top \mathbf{h}, \quad (20)$$

where  $tp \in \{\text{ICEV}, \text{HEV}, \text{PHEV}, \text{EV}\}$  indicates the type of the vehicle.

#### 4.5 Driver-specific Meta-optimization

A well-trained global model may fail to estimate the VEC for long-tail drivers with insufficient historical trips. Inspired by the success of Model-Agnostic Meta-Learning (MAML) [11] in handling few-shot problems, we propose to learn a meta-optimized universal parameter initialization that can fast adapt to all drivers. Then we fine-tune private models for each driver based on the model initialized by meta-training.

Formally, we apply MAE as our loss function for both driving behavior prediction and VEC estimation.

$$\mathcal{L}_{beh}(\hat{\mathbf{y}}^e) = \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^f |\mathbf{y}_i^e[j] - \hat{\mathbf{y}}_i^e[j]|, \quad (21)$$

$$\mathcal{L}_{EC}(\hat{y}) = |\mathbf{y} - \hat{y}|, \quad (22)$$

$$\mathcal{L} = \mathcal{L}_{beh}(\hat{\mathbf{y}}^e) + \mathcal{L}_{EC}(\hat{y}), \quad (23)$$

where  $j$  and  $f$  denote the index and the number of statistical driving behavior features, respectively.

Then, we regard VEC estimation for an individual driver as an independent task. The drivers' datasets can be denoted as  $\{(D_s^u, D_q^u)\}_{u=1}^U$ ,

where  $U$  is the number of drivers,  $D_s^u$ ,  $D_q^u$  denotes the support set, and the query set of driver  $u$ , respectively.

In  $i$ -th epoch, we run every dataset on MAML to learn the globally adaptive model parameters. In every meta-training step, we perform fast adaptation on the support set  $D_s^u$ , calculate the loss  $\mathcal{L}_{D_s^u}^i(f_\theta)$ , and update parameters as follow:

$$\theta' \leftarrow \theta - \eta \nabla_\theta \mathcal{L}_{D_s^u}^i(f_\theta), \quad (24)$$

where  $\theta'$  are updated parameters,  $\eta$  is the inner loop learning rate, and  $\nabla_\theta \mathcal{L}_{D_s^u}^i(f_\theta)$  denotes the gradient of loss on the support set. Then, we evaluate the model by running the query set  $D_q^u$  and calculate the loss  $\mathcal{L}_{D_q^u}^i(f_{\theta'})$ . At the end of each epoch, we apply bi-level optimization to update model parameters as:

$$\theta \leftarrow \theta - \gamma \nabla_\theta \sum_{u=1}^U \mathcal{L}_{D_q^u}^i(f_{\theta'}), \quad (25)$$

where  $\gamma$  is the outer loop learning rate. We run several meta-training epochs until it performs well on the validation dataset.

Afterward, we merge the support and query set as  $D^u$ , then fine-tune the initialized parameters on driver  $u$ . In the end, we would obtain  $U$  sets of parameters, and each of them can be well adapted to its corresponding driver. Please refer to Appendix A.2 for the complete meta-optimization algorithm.

## 5 EXPERIMENTS

### 5.1 Experimental Setup

In this section, we introduce the metrics of our experiments, the baseline models we compared, and the implementation details. Moreover, please refer to A.4 for the prototype system design.

**5.1.1 Metrics.** We adopt three widely used evaluation metrics in regression tasks: Mean Square Error (MSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE).

**5.1.2 Implementation Details.** The embedding size of features is set as 20. The top- $K$  value is 5. We take the filter size  $q = 4$  for CNN and the number of heads  $h = 4$  for the Transformer encoder and the multi-head attention module. The hidden size of GRU and MLP modules is set as 20 and 40, respectively. We utilize Adam optimizer with the inner loop learning rate  $6 \times 10^{-4}$  and  $3 \times 10^{-4}$ , the outer loop step size  $6 \times 10^{-3}$  and  $3 \times 10^{-3}$ , and the fine-tuning learning rate  $6 \times 10^{-4}$  and  $10^{-4}$  for the VED and ETTD datasets. The L2 penalty is set to  $10^{-5}$ . For each dataset, we randomly select each driver's 10% and 20% of records as the validation and test sets, respectively. The rest is left for training. The support set is split from the training set of the driver with a ratio of 10%, and the rest is left as the query set.

**5.1.3 Baselines.** We compare META-PEC<sup>4</sup> with numerical and data-driven methods. The following ten baselines are compared.

**Average** [7] estimates the energy consumption by multiplying the VEC per kilometer by the length of the trip.

**MLR** is known as the multiple linear regression model. [7] utilizes

<sup>4</sup>Source code is available at <https://github.com/usail-hkust/Meta-Pec>

**Table 2: Overall performance of META-PEC and all baselines on two datasets. The best and second-best results are highlighted in boldface and underlined, respectively.**

Model	VED			ETTD		
	MSE	MAE	MAPE	MSE	MAE	MAPE
Average	0.4142	0.1569	62.52%	0.8245	0.5279	43.27%
MLR	0.0551	0.0529	48.12%	0.7627	0.5014	38.95%
XGBoost	0.0446	0.0416	28.05%	0.7531	<u>0.4989</u>	42.31%
DNN	0.0397	0.0427	26.55%	0.7956	0.5187	38.38%
LSTM	0.0393	0.0421	27.56%	0.7646	0.5045	39.36%
Transformer	0.0382	0.0398	24.50%	0.7669	0.5068	39.30%
LDFeRR	0.0470	0.0425	28.16%	0.7483	0.5085	38.76%
Enc-Dec	<u>0.0378</u>	0.0398	25.55%	0.7496	0.5071	37.81%
PLd-FeRR	0.0405	<u>0.0395</u>	24.37%	0.7519	0.504	<u>37.66%</u>
Meta-TTE	0.0458	0.0463	40.72%	<u>0.7416</u>	0.5139	40.12%
<b>META-PEC</b>	<b>0.0350</b>	<b>0.0349</b>	<b>22.82%</b>	<b>0.7306</b>	<b>0.4898</b>	<b>37.10%</b>

real-world measured driving data and domain knowledge to construct a particular MLR model for VEC estimation, and predicts the speed profile based on a NN model.

**XGBoost** is a well-known gradient boosting model. [3] utilizes XGBoost as the state-of-the-art model.

**DNN** [4, 7, 25] is a classic NN model and widely used by multiple researchers in VEC estimation tasks.

**LSTM** [4, 18] is a typical recurrent neural network widely utilized to handle the route sequence data and make road-level VEC estimations.

**Transformer** [26] is a well-known self-attention-based sequential model.

**LDFeRR** [17] utilizes an attention-based GRU to estimate the road-level VEC.

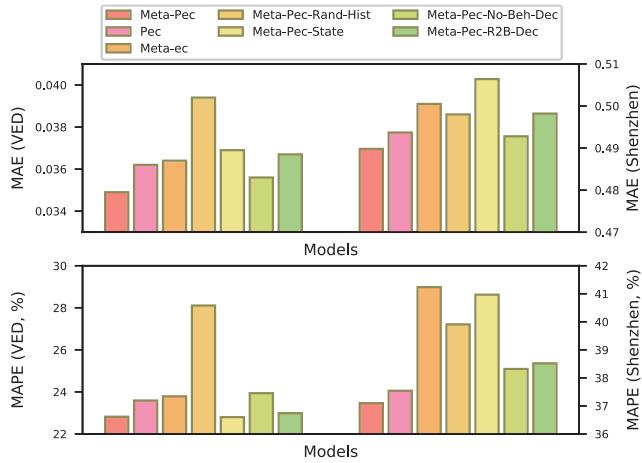
**Enc-Dec** [18] utilizes an encoder-decoder structure to estimate the road-level VEC.

**PLd-FeRR** [31] identifies the features indicating the driving preference, and a Transformer-based model is deployed for VEC estimation.

**Meta-TTE** [27] is a meta-learning-based model utilized for the estimated time of arrival (ETA) prediction tasks. We modify its output from travel time into the energy consumption of the target trip.

### 5.2 Overall Results

Table 2 shows the experimental results of the ten baselines and our proposed method on two real-world datasets. META-PEC outperforms all the baseline models and achieves state-of-the-art. In detail, almost all classic methods (Average, MLR, XGBoost) perform worse than deep learning models since they only consider the trip's and driver's statistical information. They estimate energy consumption from a macroscopic view. Well-designed VEC estimation-oriented models (LDFeRR, Enc-Dec, PLd-FeRR) have the best performance among most baseline models, indicating more information is captured by the model as we consider more driver-specific features. Meta-TTE is a model for ETA prediction, which performs worse on the VED since this dataset is constructed based on real vehicle energy usage. In contrast, VECs of the ETTD dataset are calculated

**Figure 4: Ablation tests of the model on two datasets.**

based on mechanical energy, indicating a gap existing between the ETA and VEC prediction tasks. Note Meta-TTE is designed to provide accurate travel time estimates even when there are changes in traffic conditions or road networks, while our model leverages meta-learning to acquire globally shared knowledge from various drivers and enable rapid adaptation to insufficient data scenarios.

Further looking into the results, META-PEC significantly surpasses all deep-learning-based models by (7.51%, 11.66%, 6.37%) on the VED dataset, indicating our proposed method can make more appropriate driver-specific VEC estimations. The reasons are three-fold. First, unlike other baseline models which only model personalization features by handcrafting statistical features, we extract driving preferences from the historical trajectories at the trajectory-level. Second, the driving behavior prediction module also helps in providing extra estimation basis and supervision signals. More in-detailed analysis of this module will be provided in Section 5.3.5. Finally, different from methods that learn a unified model for all users, we utilize a meta-optimization strategy to perform fast adaptation on long-tail drivers, which will be studied in Section 5.4.

### 5.3 Ablation Study

In order to verify the effectiveness of each module, we conduct ablation studies on six variants of our proposed META-PEC, including (1) PEC: the base model, which does not utilize the meta-optimization strategy, (2) META-EC: the model does not use any personalization module (*i.e.*, the driving preference learning and driving behavior prediction module), (3) META-PEC-RAND-HIST: the model randomly selects  $K$  historical trips to extract the driver’s preference, (4) META-PEC-STATE: the model learns the driving preference by modeling vehicle state sequence rather than driver behaviors, (5) META-PEC-No-BEH-DEC: the model that does not predict driving behaviors on each target road, and (6) META-PEC-R2B-DEC: the model that predicts driving behaviors only based on the road features rather than jointly considering driving preferences and the road conditions. The comparison results among all variants are shown in Figure 4.

**Table 3: Performances on long-tail drivers**

Model	VED			ETTD		
	MSE	MAE	MAPE	MSE	MAE	MAPE
PEC	0.0452	0.0686	22.85%	2.9598	1.0203	70.41%
META-PEC	<b>0.0316</b>	<b>0.0587</b>	<b>21.86%</b>	<b>2.8950</b>	<b>0.8848</b>	<b>50.25%</b>

**5.3.1 The effectiveness of the meta-optimization module (PEC).** Although PEC and META-PEC have similar performance, META-PEC is better (3.74% at most on the VED dataset) due to its meta-optimization module, indicating the model has learned the globally shared knowledge and is able to adapt quickly to each driver’s particular preference and provide accurate VEC estimations for long-tail drivers.

**5.3.2 The effectiveness of personalization modules (META-EC).** We designed two modules for personalized VEC estimation, including the driving preference learning module and the selection-based driving behavior prediction. After we exclude these two components and only leave statistical information as personalized features, it performs badly on the two datasets (drops 11.16% at most on the ETTD dataset), indicating the handcrafting statistical data is coarse-grained and helps little to the accurate personalized estimations.

**5.3.3 The effectiveness of the top- $K$  historical trip selection (META-PEC-RAND-HIST).** After we replace the top- $K$  historical trip selection strategy by random picking, this variant model performs much worse on two datasets (drops 23.19% at most on the VED dataset), representing the irrelevant trips provide little information about how the driver will drive on the target route.

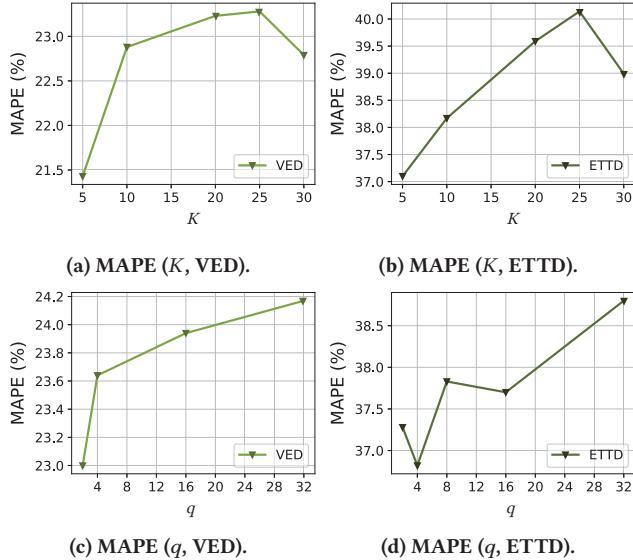
**5.3.4 The effectiveness of modeling driver behaviors (META-PEC-STATE).** META-PEC-STATE does not deploy CNN to extract behaviors but only uses a transformer module to model the vehicle state sequence. It performs terribly on the two datasets (drops 10.43% at most on the ETTD dataset), indicating it is reasonable to encode driver behaviors rather than directly model vehicle states on the trajectory.

**5.3.5 The effectiveness of driving behavior prediction (META-PEC-No-BEH-DEC).** The driving behavior prediction module can offer more VEC estimation basis for VEC estimations and extra supervision signals as a joint learning module. As we exclude the driving behavior prediction module, the performance drops on the two datasets (4.91% at most on the VED dataset).

**5.3.6 The effectiveness of the Behavior to Behavior (B2B) prediction (META-PEC-R2B).** Previous studies like [18] predict road-level driver energy consumption only based on road features, which require the model to transfer the road domain into the energy usage behavior domain (R2B). It is a more complex manner and does not consider any personalization information. After we switch B2B into R2B prediction, the performance drops on the two datasets (5.17% at most on the ETTD dataset).

### 5.4 Effectiveness on Long-tail Drivers

We consider drivers who have less than ten training samples as long-tail drivers, which account for 14.65% and 8.12% in the VED and ETTD datasets, respectively. To verify the effectiveness of



**Figure 5: Parameter sensitivity tests of the model on two datasets.**

**Table 4: The importance of each top- $K$  historical trips. "# of same roads" indicates the number of road segments shared by the historical and the target route.**

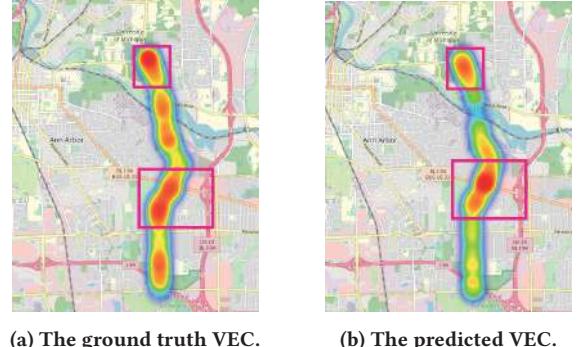
Rank	# of same roads	Importance
# 1	31	0.2689
# 2	29	0.2626
# 3	29	0.2637
# 4	25	0.1074
# 5	24	0.0974

our proposed meta-optimization module, we further compare the performance of META-PEC and PEC (*i.e.*, the model does not utilize the meta-optimization module) on long-tail drivers. Table 3 presents the results on two datasets with and without the meta-optimization module. We observe that META-PEC significantly surpasses PEC on long-tail drivers (30.08% and 28.64% at most on the VED and ETTD dataset, respectively), indicating our proposed model is more robust in handling drivers with limited instances and achieving personalized VEC estimation for various drivers.

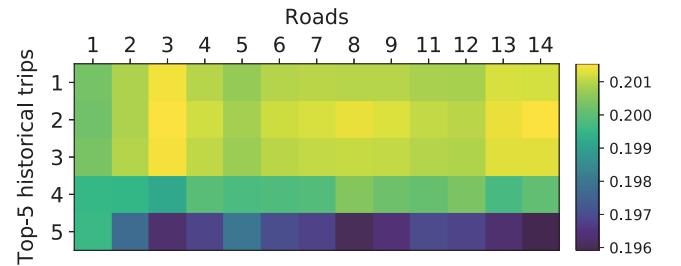
## 5.5 Parameter Sensitivity

We conduct experiments on the two datasets to study the impacts of the following hyper-parameters in META-PEC.  $K$  is the number of the most similar historical trips we extracted.  $q$  is the number of continuous vehicle states in a behavior segment. States contained in a segment are considered a driving behavior in the trip trajectory (*e.g.*, accelerating, braking, etc.).

Figure 5a and Figure 5b show the results with varying  $K$ . It is clear that the model's performance drops as we sample more historical trips, indicating dissimilar trips contain noisy information and are less helpful to the driver's driving preference. Keeping the most



**Figure 6: The spatial distribution of the ground truth and predicted VEC.**



**Figure 7: The learned multi-head attention weights.**

relevant trips can help the model focus on what is truly important. META-PEC achieves the best performance when setting  $K = 5$  on the two datasets.

Figure 5c and Figure 5d report the results of varying  $q$ . We observe that the performance drops when more states are included in a behavior segment, which is perhaps because the CNN needs to handle more behaviors simultaneously. As a result, the model performance decreased as the task became difficult. We set the  $q$  value as 2 and 4 on the VED and ETTD datasets, respectively.

## 5.6 Case Study

In this section, we conduct a case study to further validate the VEC estimation performance of our proposed META-Pec. We take an example of the VED dataset, which achieves absolute error (AE) at  $5 \times 10^{-4}$  and absolute percentage error (APE) at 0.56%.

**5.6.1 Spatial visualization.** We first analyze the spatial distribution of the estimated VEC. Figure 6a shows the route and the ground truth energy usage of the example route. Figure 6b shows the energy consumption on each road segment predicted by the selection-based driving behavior prediction module. As can be seen, our model properly estimated the high energy consumption area (in red rectangles) indicating META-PEC successfully captures the driver's preference and road conditions for accurate VEC estimations.

**5.6.2 Module-level analysis.** We report the learned importance of historical trips in Table 4. As can be seen, the importance of historical trips decreases gradually, indicating META-PEC owns the ability

to discriminate the most critical historical trips that can provide helpful information to the VEC estimation. Besides, we calculate the multi-head attention weights of each historical trip defined by Equation 12. As illustrated in Figure 7, the learned attention weights of top-3 trips are higher than the rest two, indicating that our driving behavior prediction module has also learned how to select the most useful information.

## 6 RELATED WORK

### 6.1 Vehicle Energy Consumption Estimation

Previous studies on vehicle energy consumption estimation can be mainly categorized into numerical approaches [1, 6, 8, 20] and data-driven approaches [4, 7, 10, 14, 17, 18, 22, 31]. Based on the vehicle dynamics equation as the underlying physical model, Cauwer *et al.* [6] proposed using multiple linear regression (MLR) models to identify correlations between the kinematic parameters of the vehicle and VEC. Similarly, Al-Wreikat *et al.* [1] proposed to evaluate the driving behavior, distance, temperature, traffic, and road grade effects on the VEC of an electric vehicle and the result shows the driver's behaviors have significant influences on VEC. Ojeda *et al.* [20] considered using a real physical vehicle model for the speed and fuel consumption prediction. Ding *et al.* [8] proposed to evaluate VEC by measuring the power generated by fuel combustion. Recently, machine learning based methods have been studied for vehicle energy consumption estimation. Cauwer *et al.* [7] developed a data-driven method utilizing real-world measured driving data and domain knowledge to construct MLR models for VEC estimation. Chen *et al.* [4] proposed to use of long short-term memory (LSTM) and artificial neural network (ANN) models to estimate the energy consumption of electric buses. Liu *et al.* [17] utilized an attention-based GRU to estimate the road-level VEC. DeepFEC [10] proposed a deep-learning-based model to forecast energy consumption on every road in a city based on real traffic conditions. Hua *et al.* [14] developed a transfer learning model for electric vehicle energy consumption estimation based on insufficient electric vehicles and ragged driving trajectories. PLD-FeRR [31] employs the Transformer for VEC estimation, with a consideration of the driver's driving preference through handcrafted personalized features.

### 6.2 Spatial-temporal Data Mining

Our study is closely related to spatiotemporal data mining [12, 15, 16, 23, 30, 33, 34, 37–39]. With the propensity of GPS devices, spatial-temporal data mining has been extensively studied in various applications. To name a few, Estimated time of arrival (ETA) prediction is a classic task that aims to estimate the travel time with a given origin, destination, and departure time. WDR [30] proposed to combine Wide-Deep Learning with an LSTM module for ETA. It is worth to mention that WDR also considers some personalized features (*e.g.*, driver profile, rider profile, vehicle profile, *etc.*) to improve the performance. HetETA [12] proposes to transform the road map into a heterogeneous graph and introduce a vehicle-trajectories-based network to consider traffic behavior patterns jointly. Huang *et al.* [15] argue modeling traffic congestion is important for accurate ETA prediction, and they constructed a congestion-sensitive graph and a route-aware graph transformer to learn the long-distance congestion correlations. Driving is a

complex activity influenced by multiple factors. Analyzing driving behavior helps to assess driver performance, including safety and energy efficiency, leading to enhancements in transportation systems. Studies like [2] and [29] propose using discretized state transition graphs derived from trajectories to identify different driving behaviors. In contrast, [33] proposes to provide better driving behavior predictions by modeling the correlation between drivers' skills and interactions hidden in their social networks. Besides, Next Point-of-Interest (POI) aims to recommend the next POIs drivers are most likely to visit based on their historical trajectories. Rao *et al.* [23] proposed a Spatial-Temporal Knowledge Graph (STKG), which can directly learn transition patterns between POIs. Trajectory prediction is similar to Next POI recommendation, which aims to predict the driver's future visiting grid cell in the trajectory. Xu *et al.* [34] designed a cluster-based [36] network initialization method based on a meta-learning algorithm to obtain initial personalized parameters for each trajectory. With the development of electric vehicle technologies, spatiotemporal data mining has also been applied to electric vehicle tasks. In order to help drivers find proper spots for charging, Zhang *et al.* [38] proposed a framework called MASTER for charging station recommendation by considering each charging station as an individual agent and formulating the problem as a multi-objective multi-agent reinforcement learning task. To balance the use of charging stations, MAGC [37] proposes to provide dynamic pricing for each charging request and achieve effective use of stations by formulating this problem as a mixed competitive-cooperative multi-agent reinforcement learning task with multiple long-term commercial goals.

## 7 CONCLUSION

In this paper, we investigated the personalized vehicle energy consumption estimation problem by explicitly exploiting driving behaviors hidden in historical trajectories. Specifically, we proposed a preference-aware meta-optimization framework (META-PEC) which consists of three major modules. We first proposed a driving preference learning module to extract latent spatiotemporal preferences from historical trips. After that, we constructed a selection-based driving behavior prediction module to estimate the possible driving behavior on a given route with the consideration of the driver's past relevant trips. Furthermore, a driver-specific meta-optimization module is proposed to learn a shared global model parameter initialization that can be fast adapted to each long-tail driver with a few historical trips. Extensive experiments on two large real-world datasets demonstrated the effectiveness of META-PEC against ten baselines. In the future, we plan to deploy META-PEC to more cities so as to provide insightful information for various decision-making tasks such as individual trip planning and sustainable transportation system management.

## ACKNOWLEDGMENTS

This research was supported in part by the National Natural Science Foundation of China under Grant No.62102110, Guangzhou Science and Technology Plan Guangzhou-HKUST(GZ) Joint Project No. 2023A03J0144, and Foshan HKUST Projects (FSUST21-FYTRI01A, FSUST21-FYTRI02A).

## REFERENCES

- [1] Yazan Al-Wreikat, Clara Serrano, and José Ricardo Sodré. 2021. Driving behaviour and trip condition effects on the energy consumption of an electric vehicle under real-world driving. *Applied Energy* 297 (2021), 117096.
- [2] Chao Chen, Qiang Liu, Xingchen Wang, Chengwu Liao, and Daqing Zhang. 2022. semi-Traj2Graph identifying fine-Grained driving style with GPS trajectory data via multi-task learning. *IEEE Transactions on Big Data* 8, 6 (2022), 1550–1565.
- [3] Xiaowei Chen, Zengxiang Lei, and Satish V. Ukkusuri. 2022. Prediction of road-level energy consumption of battery electric vehicles. In *Proceedings of the 25th IEEE ITSC International Conference on Intelligent Transportation Systems, October 8–12, 2022*. IEEE, 2550–2555.
- [4] Yuchu Chen, Yuntao Zhang, and Ruixiao Sun. 2021. Data-driven estimation of energy consumption for electric bus under real-world driving conditions. *Transportation Research Part D: Transport and Environment* 98 (2021), 102969.
- [5] Junyoung Chung, Çağlar Gülcabay, KyungHyun Cho, and Yoshua Bengio. 2014. Empirical evaluation of gated recurrent neural networks on sequence modeling. *CoRR* abs/1412.3555 (2014). arXiv:1412.3555
- [6] Cedric De Cauwer, Joeri Van Mierlo, and Thierry Coosemans. 2015. Energy consumption prediction for electric vehicles based on real-world data. *Energies* 8, 8 (2015), 8573–8593.
- [7] Cedric De Cauwer, Wouter Verbeke, Thierry Coosemans, Saphir Faid, and Joeri Van Mierlo. 2017. A Data-driven method for energy consumption prediction and energy-efficient routing of electric vehicles in real-world conditions. *Energies* 10, 5 (2017).
- [8] Yan Ding, Chao Chen, Shu Zhang, Bin Guo, Zhiwen Yu, and Yasha Wang. 2017. GreenPlanner: Planning personalized fuel-efficient driving routes using multi-sourced urban data. In *Proceedings of the 2017 IEEE Percom International Conference on Pervasive Computing and Communications, March 13–17, 2017*. IEEE Computer Society, 207–216.
- [9] Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, Jakob Uszkoreit, and Neil Houlsby. 2021. An image is worth 16x16 words: Transformers for image recognition at scale. In *Proceedings of the 9th ICCV International Conference on Learning Representations, May 3–7, 2021*. OpenReview.net.
- [10] Sayda Elmi and Kian-Lee Tan. 2021. DeepFEC: Energy consumption prediction under real-world driving conditions for smart cities. In *Proceedings of the 2021 ACM WWW International World Wide Web Conference, April 19–23, 2021*. ACM / IW3C2, 1880–1890.
- [11] Chelsea Finn, Pieter Abbeel, and Sergey Levine. 2017. Model-agnostic meta-learning for fast adaptation of deep networks. In *Proceedings of the 34th ICML International Conference on Machine Learning, 6–11 August 2017 (Proceedings of Machine Learning Research, Vol. 70)*. PMLR, 1126–1135.
- [12] Huiting Hong, Yucheng Lin, Xiaoqing Yang, Zang Li, Kun Fu, Zheng Wang, Xiaohu Qie, and Jieping Ye. 2020. HetETA: Heterogeneous information network embedding for estimating time of arrival. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, August 23–27, 2020*. ACM, 2444–2454.
- [13] Max Horn, Michael Moor, Christian Bock, Bastian Rieck, and Karsten M. Borgwardt. 2020. Set functions for time series. In *Proceedings of the 37th ICML International Conference on Machine Learning, 13–18 July 2020 (Proceedings of Machine Learning Research, Vol. 119)*. PMLR, 4353–4363.
- [14] Yining Hua, Michele Sevegnani, Dewei Yi, Andrew Birnie, and Steve McAslan. 2022. Fine-grained RNN with transfer learning for energy consumption estimation on EVs. *IEEE Transactions on Industrial Informatics* 18, 11 (2022), 8182–8190.
- [15] Jizhou Huang, Zhengjie Huang, Xiaomin Fang, Shikun Feng, Xuyi Chen, Jiaxiang Liu, Haitao Yuan, and Haifeng Wang. 2022. DuETA: Traffic congestion propagation pattern modeling via efficient graph learning for ETA prediction at Baidu Maps. In *Proceedings of the 31st ACM CIKM International Conference on Information and Knowledge Management, October 17–21, 2022*. ACM, 3172–3181.
- [16] Fan Liu, Hao Liu, and Wenzhao Jiang. 2022. Practical Adversarial Attacks on Spatiotemporal Traffic Forecasting Models. In *Proceedings of the 2022 NeurIPS Annual Conference on Neural Information Processing Systems, Nov 28 - Dec 9, 2022*.
- [17] Min Liu, Zhaozhui Peng, Xiaohui Yu, Senzhang Wang, and Qiao Song. 2021. LD-FeRR: A fuel-efficient route recommendation approach for long-distance driving based on historical trajectories. In *Proceedings of the 2021 SIAM SDM International Conference on Data Mining, April 29 - May 1, 2021*. SIAM, 73–81.
- [18] Ayman Moawad, Krishna Murthy Gurumurthy, Omer Ömer Verbas, Zhijian Li, Ehsan Islam, Vincent Freyermuth, and Aymeric Rousseau. 2021. A deep learning approach for macroscopic energy consumption prediction with microscopic quality for electric vehicles. *CoRR* abs/2111.12861 (2021). arXiv:2111.12861
- [19] Geunseob Oh, David J. LeBlanc, and Huei Peng. 2022. Vehicle energy dataset (VED), a large-scale dataset for vehicle energy consumption research. *IEEE Transactions on Intelligent Transportation Systems* 23, 4 (2022), 3302–3312.
- [20] Luis Leon Ojeda, Alexandre Chasse, and Romain Goussault. 2017. Fuel consumption prediction for heavy-duty vehicles using digital maps. In *Proceedings of the 20th IEEE ICST International Conference on Intelligent Transportation Systems, October 16–19, 2017*. IEEE, 1–7.
- [21] Federico Perrotta, Tony Parry, and Luís C. Neves. 2017. Application of machine learning for fuel consumption modelling of trucks. In *Proceedings of the 2017 IEEE Big Data International Conference on Big Data, December 11–14, 2017*. IEEE Computer Society, 3810–3815.
- [22] Linas Petkevicius, Simonas Saltenis, Alminas Civilis, and Kristian Torp. 2021. Probabilistic deep learning for electric-vehicle energy-use prediction. In *Proceedings of the 17th ACM SSTD International Symposium on Spatial and Temporal Databases, August 23–25, 2021*. ACM, 85–95.
- [23] Xuan Rao, Lisi Chen, Yong Liu, Shuo Shang, Bin Yao, and Peng Han. 2022. Graph-flashback network for next location recommendation. In *Proceedings of the 28th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, August 14 - 18, 2022*. ACM, 1463–1471.
- [24] Nadine Rauth, Thomas Franke, and Josef F. Krems. 2015. Understanding the impact of electric vehicle driving experience on range anxiety. *Human Factors: The Journal of Human Factors and Ergonomic Society* 57, 1 (2015), 177–187.
- [25] Millend Roy, Akshay Uttama Nambi, Anupam Sobti, Tanuja Ganu, Shivkumar Kalyanaraman, Shankar Akella, Jaya Subha Devi, and S. A. Sundaresan. 2022. Reliable energy consumption modeling for an electric vehicle fleet. In *Proceedings of the 2022 ACM COMPASS SIGCAS/SIGCHI Conference on Computing and Sustainable Societies, 29 June 2022 - 1 July 2022*. ACM, 29–44.
- [26] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. In *Proceedings of the 2017 NeurIPS Annual Conference on Neural Information Processing Systems, December 4–9, 2017*. 5998–6008.
- [27] Chenxing Wang, Fang Zhao, Haichao Zhang, Haiyong Luo, Yanjun Qin, and Yuchen Fang. 2022. Fine-grained trajectory-based travel time estimation for multi-city scenarios based on deep meta-learning. *IEEE Transactions on Intelligent Transportation Systems* 23, 9 (2022), 15716–15728.
- [28] Guang Wang, Xiuyuan Chen, Fan Zhang, Yang Wang, and Desheng Zhang. 2019. Experience: Understanding long-term evolving patterns of shared electric vehicle networks. In *Proceedings of the 25th ACM MobiCom Annual International Conference on Mobile Computing and Networking, MobiCom 2019, October 21–25, 2019*. ACM, 9:1–9:12.
- [29] Pengyang Wang, Yanjie Fu, Jiawei Zhang, Pengfei Wang, Yu Zheng, and Charu C. Aggarwal. 2018. You are how you drive: Peer and temporal-aware representation learning for driving behavior analysis. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, August 19–23, 2018*. ACM, 2457–2466.
- [30] Zheng Wang, Kun Fu, and Jieping Ye. 2018. Learning to estimate the travel time. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, August 19–23, 2018*. ACM, 858–866.
- [31] Zhan Wang, Zhaozhui Peng, Senzhang Wang, and Qiao Song. 2022. Personalized long-distance fuel-efficient route recommendation through historical trajectories mining. In *Proceedings of the 15th ACM WSDM International Conference on Web Search and Data Mining, February 21 - 25, 2022*. ACM, 1072–1080.
- [32] Yang Xing, Chen Lv, Dongpu Cao, and Chao Lu. 2020. Energy oriented driving behavior analysis and personalized prediction of vehicle states with joint time series modeling. *Applied Energy* 261 (2020), 114471.
- [33] Tong Xu, Hengshu Zhu, Xiangyu Zhao, Qi Liu, Hao Zhong, Enhong Chen, and Hui Xiong. 2016. Taxi driving behavior analysis in latent vehicle-to-vehicle networks: A social influence perspective. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, August 13–17, 2016*. ACM, 1285–1294.
- [34] Yuan Xu, Jiajie Xu, Jing Zhao, Kai Zheng, An Liu, Lei Zhao, and Xiaofang Zhou. 2022. MetaPTP: An adaptive meta-optimized model for personalized spatial trajectory prediction. In *Proceedings of the 28th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, August 14 - 18, 2022*. ACM, 2151–2159.
- [35] Can Yang and Gyöző Gidófalvi. 2018. Fast map matching, an algorithm integrating hidden Markov model with precomputation. *International Journal of Geographical Information Science* 32, 3 (2018), 547–570.
- [36] Yang Yang, De-Chuan Zhan, Yi-Feng Wu, Zhi-Bin Liu, Hui Xiong, and Yuan Jiang. 2021. Semi-Supervised Multi-Modal Clustering and Classification with Incomplete Modalities. *IEEE Transactions on Knowledge and Data Engineering* 33, 2 (2021), 682–695.
- [37] Weijia Zhang, Hao Liu, Jindong Han, Yong Ge, and Hui Xiong. 2022. Multi-agent graph convolutional reinforcement learning for dynamic electric vehicle charging pricing. In *Proceedings of the 28th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, August 14 - 18, 2022*. ACM, 2471–2481.
- [38] Weijia Zhang, Hao Liu, Fan Wang, Tong Xu, Haoran Xin, Dejing Dou, and Hui Xiong. 2021. Intelligent electric vehicle charging recommendation based on multi-agent reinforcement learning. In *Proceedings of the 2021 ACM WWW International World Wide Web Conference, April 19–23, 2021*. ACM / IW3C2, 1856–1867.
- [39] Jingbo Zhou, Shan Gou, Renjun Hu, Dongxiang Zhang, Jin Xu, Airong Jiang, Ying Li, and Hui Xiong. 2019. A collaborative learning framework to tag refinement for points of interest. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD 2019, Anchorage, AK, USA, August 4–8, 2019*. 1752–1761.

## A APPENDIX

### A.1 Top- $K$ historical trip selection Algorithm

---

**Algorithm 1:** Top- $K$  historical trip selection

---

**input :** The route of the target trip  $R_c$  and its departure time  $s_c$ , and historical trips' routes  $\{R_i\}_{i=1}^M$  and departure times  $\{s_i\}_{i=1}^M$ .  
**output :** The target trip's most similar  $K$  trips  $H^K$ .

```

1 for  $i \leftarrow 1$  to  $M$  do
2   Calculate the route similarity by:
     $score_{c,i}^{route} = |R_c \cap R_i|$ ;
3   Calculate the temporal similarity by:
     $score_{c,i}^{time} = |s_c - s_i|$ ;
4 end
5 for  $i \leftarrow 1$  to  $M$  do
6   Normalize each similarity score as:
     $norm(score_{c,i}^*) = \frac{score_{c,i}^* - Min(score_c^*)}{Max(score_c^*) - Min(score_c^*)}$ ;
7   Calculate the final similarity score by:
     $score_{c,i} = norm(score_{c,i}^{route}) - norm(score_{c,i}^{time})$ ;
8 end
9 Select  $K$  historical trips with the highest scores as  $H^K$ .

```

---

### A.2 Meta-optimization Algorithm

---

**Algorithm 2:** Meta-optimization

---

**input :** All drivers' task datasets  $\{(D_s^u, D_q^u)\}_{u=1}^U$ , inner loop learning rate  $\eta$ , outer loop learning rate  $\gamma$ , and fine-tuning learning rate  $\omega$ .  
**output :** Model parameters  $\{\theta^u\}_{u=1}^U$  adapted to every driver.

```

1 Random initialize model parameters  $\theta$ ;
2 for  $i \leftarrow 1$  to  $N_{epoch}$  do
3   for  $u \leftarrow 1$  to  $U$  do
4     Calculate the gradient of loss on the support set as:
         $\nabla_\theta \mathcal{L}_{D_s^u}^i(f_\theta)$ ;
5     Update parameters with gradient descent as:
         $\theta' \leftarrow \theta - \eta \nabla_\theta \mathcal{L}_{D_s^u}^i(f_\theta)$ ;
6     Calculate the loss on the query set as:  $\mathcal{L}_{D_q^u}^i(f_{\theta'})$ ;
7   end
8   Update parameters with gradient descent as:
     $\theta \leftarrow \theta - \gamma \nabla_\theta \sum_{u=1}^U \mathcal{L}_{D_q^u}^i(f_{\theta'})$ ;
9 end
10 for  $u \leftarrow 1$  to  $U$  do
11   Merge the support and query set as  $D^u$ ;
12   Fine-tune the model by:  $\theta^u \leftarrow \theta - \omega \nabla_\theta \mathcal{L}_{D^u}(f_\theta)$ ;
13 end

```

---

### A.3 Additional Case Study

The ground truth and predicted VEC spatial distribution of additional cases are presented in Figure 8-9.

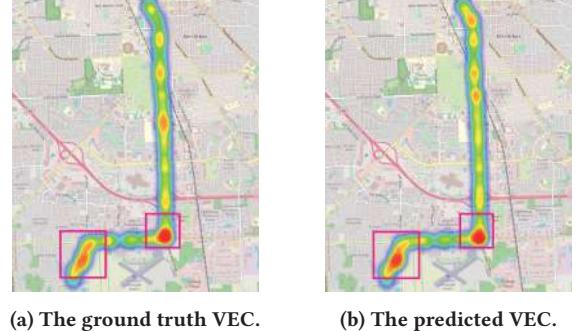


Figure 8: Additional case 1.

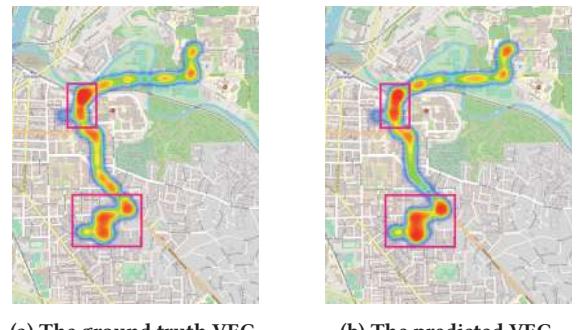


Figure 9: Additional case 2.

### A.4 Prototype System

We have implemented a demo system to provide personalized energy consumption estimation for given trips. Figure 10 shows the screenshot of our demo system. For each trip, the system displays the route, the origin and destination, and the estimated energy consumption of the trip, where warm colors indicate high energy consumption. Moreover, the system also provides a road view to visualize the energy consumption of each road segment based on the past trajectories traversed.

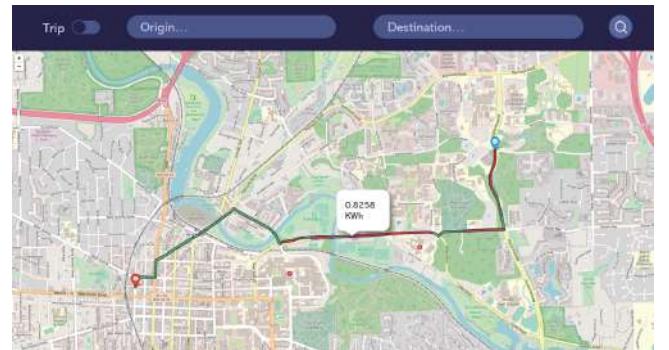


Figure 10: Prototype system.