Neural Computing and Applications (2024) 36:8565–8576
https://doi.org/10.1007/s00521-024-09512-5

ORIGINAL ARTICLE

A seq2seq learning method for microscopic emission estimation of on-road vehicles

Zhenyi Zhao[1] · Yang Cao[1] · Zhenyi Xu[1,2] · Yu Kang[1,2,3]

Received: 31 December 2022 / Accepted: 14 January 2024 / Published online: 26 February 2024

Abstract
Microscopic emission estimation based on driving states plays a crucial role in controlling the pollution of on-road vehicles. Existing research has evolved from fitting nonlinear models of driving cycles and emission factors to utilizing neural networks to exploit driving patterns and construct a nonlinear mapping between driving states and emission values. However, due to the small percentage of driving-cycle related to high emissions, it is still challenging to capture the vehicle emission peaks, which lead to the most noteworthy high-emission characteristics being regarded as abnormal disturbances instead. To address the issue, this paper proposes a peak-sensitive microscopic emission estimation framework characterized by sequence-to-sequence learning for on-road vehicles. Sequence-to-sequence learning emphasizes serial pattern mapping from driving sequences to emission sequences to provide more statistical constraints and reduce the estimation uncertainty brought by unstable driving behaviors. Specifically, the framework aims at capturing context features related to high emissions from sequences dynamically in an adaptive and self-learning way and is composed of a driving states embedding module and a dynamic aggregation module. Particularly, an incremental tracking loss (ITL) is proposed to adjust the incremental emissions at adjacent time steps by supervising the differences of the generated sequences, enabling the model to track sudden changes in emissions. Extensive experiments are conducted on the on-board diagnostics (OBD) dataset with 12628 sampling records collected from a heavy-duty diesel vehicle. The results show that the estimation accuracy of our proposed method is significantly better than state-of-the-art methods, and it can effectively capture high-emission peaks.

Keywords Microscopic emission model · Sequence to sequence · Dynamic feature aggregation · Incremental tracking loss

1 Introduction

The microscopic emission estimation task of vehicles usually establishes the relationship between driving states and emission, which is the basis of refinement management of vehicles' emissions in an intelligent transportation system [1, 2]. Macroscopic and mesoscopic emission models at other scales apply to the simulation of emissions in a country or local area by adopting typical driving feature parameters to characterize the generic emission features of all vehicles but ignore the dynamic emission process of a single vehicle. And microscopic emission models mainly focus on short-time or instantaneous emissions from one vehicle driving on the road under specific driving states such as speed, acceleration, and engine properties.

✉ Yang Cao
forrest@ustc.edu.cn

✉ Yu Kang
kangduyu@ustc.edu.cn

Zhenyi Zhao
zzy0025@mail.ustc.edu.cn

Zhenyi Xu
xuzhenyi@mail.ustc.edu.cn

[1] Department of Automation, University of Science and Technology of China, Hefei 230026, China

[2] Institute of Artificial Intelligence, Hefei Comprehensive National Science Center, Hefei 230088, China

[3] Institute of Advanced Technology, University of Science and Technology of China, Hefei 230088, China

Methods for microscopic emission estimation have evolved from the construction and localization of traditional vehicle emission models to fit the nonlinear relationship between vehicle driving states and emission values by neural networks. The microscopic emission models, such as IVE [3] and MOVES [4], divide driving states into several intervals and assign a unique emission value to each interval. In practical applications, the base emission values are further calibrated using parameters reflecting the instantaneous driving states of the vehicles. These models are constructed under test driving cycles in the laboratory environment. However, the driving states of the on-road vehicles proceeding along the highway are more complex and changeable. The calibration coefficients of such models have limited for analyzing complex driving states when calculating the instantaneous emissions.

With extensive use of on-board emission measure systems installed on vehicles to collect the driving parameters and exhaust emissions, some researchers adopt driving states and emissions from on-road vehicles to construct regression models, such as EMIT [5], CMEM [6], and VT-Micro[7]. These methods determine the formulas according to the physical process and then fit the corresponding coefficients. As a result, it is challenging for the models to adaptively reflect the complicated emission patterns of vehicles.

Recently, some studies have learned the nonlinear features hidden in emission patterns by adopting neural networks to estimate the instantaneous emission values. Le Cornec et al. [8] and Seo et al. [9] demonstrate that the multilayer perceptron (MLP) outperforms traditional and nonlinear regression models (NLR) for estimation. Deep learning methods also concern the influence of historical driving states on current emissions [10], especially recurrent neural network (RNN) and long short-term memory (LSTM), which preserve the long-term dependence of emission behavior well. Higher accuracy in estimating CO2 emissions is achieved by adopting LSTM to learn historical driving patterns and feedforward neural network (FNN) to extract current driving behaviors and external environmental features [11].

Relevant studies have achieved high accuracy of emissions estimation under stable driving states. However, the driving state of vehicles is erratic fluctuation under complex traffic conditions. As shown in Fig. 1, vehicles start and stop more than five times within an hour, and these unstable driving states are related to high emissions. Existing studies have a large error in emission estimation under the above driving conditions, and the reasons are as follows:

(1) Vehicles will start and stop frequently during road congestion or passing traffic lights, which results in sudden changes in driving states. The instantaneous emissions of vehicles are mostly related to historical driving state sequences with finite lengths. High emissions from an accelerating vehicle are unrelated to its constant speed a few minutes ago. Therefore, there is uncertainty to accurately determine the instantaneous emissions of vehicles at current moment using historical driving states.

(2) Although getting the most attention, high-emission stages only account for a small proportion of the overall driving trip for vehicles, as shown in Fig. 1. Existing networks tend to reduce the learning error of stable-state samples in optimization, hence can learn the correlation better when the samples are sufficient. However, high-emission behaviors with fewer samples are often ignored as outliers, making it difficult for the network to capture high-emission features under complex driving conditions.

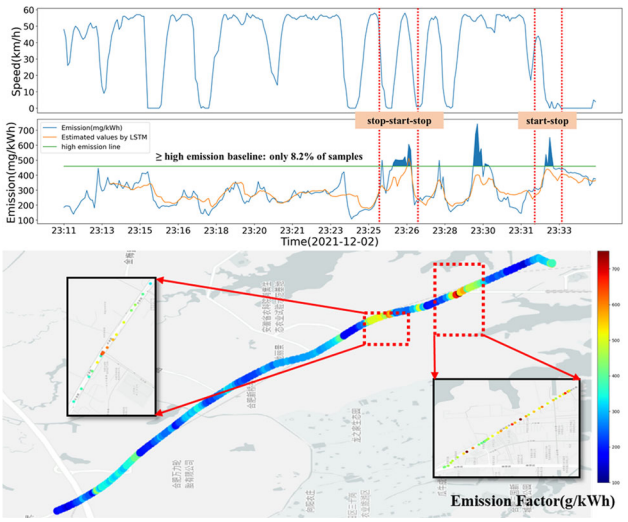

Fig. 1 Visualization of the trajectory, the corresponding driving speed, and instantaneous emission values collected by OBD from a diesel vehicle on 12/02/2021. High-emission states of vehicles account for less than 10% of the overall trajectory. These high-emission states are associated with frequent starts and stops of vehicles, which usually occur when passing traffic lights and road congestion. Existing studies have large errors in estimating peak values

In general, there are two challenges in the task of instantaneous emission estimation:

- How do identifying crucial temporal features related to instantaneous estimated emission in complex and varying conditions?
- How do enhancing the network's attention to high-emission features when the samples with high emission values are insufficient?

To address the limits, we propose a peak-sensitive seq2seq learning method for microscopic emission estimation of on-road vehicles. The main contributions are as follows:

(1) Instead of using the traditional sequence-to-state or state-to-state paradigm, we propose a sequence-to-sequence framework, which shifts the focus from the correlation between driving sequences and instantaneous emissions to the correlation between driving sequences and emission sequences. It provides more constraints about statistical property on the instantaneous emission estimation and reduces estimation uncertainty brought by unstable driving behaviors.
(2) We point out that the key factor toward precise estimation of high emission lies in identifying critical local features among the sequence. To achieve this goal, we design the network comprised of a driving states embedding module and dynamic aggregation module to learn local temporal driving features and dynamically aggregate the key features to get a causal analysis of driving patterns and emission patterns.
(3) Incremental Tracking Loss (ITL) is proposed to make the network more sensitive to high-emission features with insufficient learning examples. With ITL, the learning process is supervised not only by the overall distribution of the emission sequence but also by the emission increment of adjacent time steps namely difference values, enabling the network to track drastic emission changes.
(4) The proposed method effectively captures the high peaks of emissions. And better accuracy for emission estimation is achieved significantly than state-of-the-art methods on the OBD dataset collected from a heavy-duty diesel vehicle driving in Hefei, which contains 21 trajectories and 12628 samples.

2 Related work

The methods for microscopic emission estimation based on driving cycles are established by the portrayal of the instantaneous driving state of a vehicle [12]. The instantaneous emission values are calculated by coupling the driving states and emission factors through mathematical relationships or physical models using instantaneous speed and acceleration. For example, IVE [3], MOVES [4], PHEM [13], et al. first adopt the data in the laboratory environment to construct the basic emission factor and modify it by introducing parameters such as vehicle-specific power (VSP) or engine power describing the instantaneous driving state of the vehicle. CMEM [6], EMIT [5], VT-Micro [7], et al. determine the formula form through physical processes of vehicle's emission to fit the corresponding parameters with a large amount of data. To describe the complex emission pattern, the above models require a large number of experimental data and comprehensive driving parameters. Such traditional methods rely on too many driving parameters to modify the model to simulate driving states, which leads to high requirements for data attributes and is not conducive to model promotion.

In recent years, several studies have introduced deep learning methods to estimate the instantaneous emission values of vehicles [14–19]. Most of these methods utilize the actual driving states collected by global positioning system (GPS) [20, 21], portable emission measurement system (PEMS) [22] or OBD [23] to calculate the instantaneous emission of one on-road vehicle. Le Cornec et al. [8] and Seo et al. [9] employ backpropagation (BP) network to build the correlation between driving states and emissions and prove that BP network is superior to the traditional nonlinear regression model in emission estimation. In particular, the researchers find that the networks are more accurate in estimating emissions by introducing historical driving states. In earlier studies, the IVE model introduced engine stress (ES) to describe the level of historical VSP of vehicles, but the effect was limited. Motallebiaraghi et al. [17] utilizes BP network to map the one-dimensional driving state sequence to the emission factor and achieves outstanding accuracy. Further, researchers employ recurrent neural networks, such as RNN [24], LSTM [25], and GRU [15], which maintain long-term dependence on emission behavior well. Jia et al. [11] combine LSTM and FNN to encode the historical and current driving states and obtain excellent accuracy for emission estimation.

The algorithms based on neural networks improve the feature representation to compensate for the disadvantage of insufficient data attributes, which facilitates the model's generalizability. Meantime, these algorithms regard the small number of high-emission samples as noise, while the ignored high-emission samples are critical in the research. Therefore, how to give more attention to the emission features of the high-emission samples should be a problem can't be ignored. The algorithm proposed in the paper innovatively introduces the seq2seq framework and ITL loss function, which enhances feature representation of the model, thereby reducing the requirement for data quality and tracking the dynamic changes effectively under high emissions.

3 Preliminary

3.1 Dataset construction

The input attributes of the estimation model are driving states such as speed and engine conditions of vehicles. More specifically, the dataset tested in the research is collected from the on-board diagnostics (OBD) [26] installed on diesel vehicles. The driving variables represent as $\mathbf{s}_t = (t, v, a, ewt, tpd)$, where t is the sampling time, v and a donate the instantaneous speed (km/h) and acceleration (m/s^2) at t. ewt and tpd donate engine cooling water temperature ($^\circ C$) and throttle pedal degree ($\%$). In addition, the OBD also collects the quality of vehicles' emissions, and we transformed it into emission factor F_t (mg/kWh) of pollutant NO_x according to formulas in [27].

Figure 2 is a pairwise plot used to understand the relationship between all possible pairs of driving variables. Each subplot on the non-diagonal line is a scatter plot of the pairs of driving variables; the subplots on the diagonal line are histograms of the individual driving variables with statistics on the data distribution. In Fig. 2, the data points in the orange boxes are high-emission samples whose distribution has a fuzzy trend. For example, high emissions of vehicles usually occur when less throttle pedal degree tpd and high engine water temperature ewt. It is almost impossible for high emissions to ensue when the vehicle speed is too fast. There is a coupling relationship between driving variables and emission values.

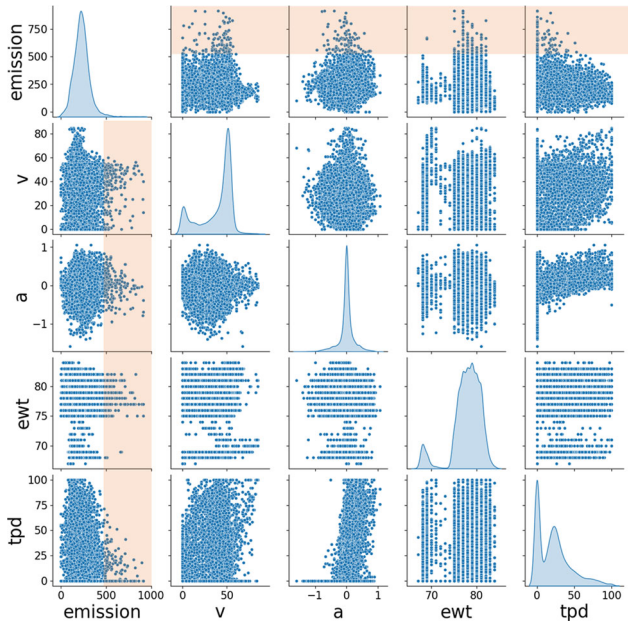

Fig. 2 Pairwise plot for emission values and driving variables: there is a correlation between all possible pairs of driving variables, especially the distribution of the high-emission samples has a vague trend

3.2 Problem description

The main goal of this paper is to estimate the instantaneous emission factor by the driving states of a vehicle. We consider the driving state at the estimated timestamp and the historical driving states with the appropriate timespan together. The input information hence could be represented as $\mathbf{X}_t = \{\mathbf{s}_{t-k}, \mathbf{s}_{t-k+1}, \ldots, \mathbf{s}_t\}$, which k is the length of historical driving state sequence, and the estimated result donates $\hat{F}_t = f(\mathbf{X}_t)$.

4 Methodology

This section introduces the proposed method for microscopic emission estimation of vehicles. As shown in Fig. 3, the model consists of two modules: driving states embedding module and dynamic aggregation module. Each module is described in detail below.

4.1 Driving states embedding module

As shown in Fig. 3, the driving states embedding module is achieved by a bidirectional GRU layer[28] which includes a forward GRU and a backward GRU. There are two gates in GRU cell: the update gate z_t defines the amount of previous memory saved to the current time step; the reset gate r_t determines how the new input information is combined with the previous memory. For the sequence of driving states $\mathbf{X}_t = \{\mathbf{s}_{t-k}, \mathbf{s}_{t-k+1}, \ldots, \mathbf{s}_t\}$ as input, the forward propagation is as follows:

$$\begin{aligned} z_i &= \sigma(W_z \cdot [h_{i-1}, \mathbf{s}_i]), \\ r_i &= \sigma(W_r \cdot [h_{i-1}, \mathbf{s}_i]), \\ \tilde{h_i} &= \tanh(W_{\tilde{h}} \cdot [r_t * h_{i-1}, \mathbf{s}_i]), \\ \overrightarrow{h}_i &= (1 - z_i) * \overrightarrow{h}_{i-1} + z_i * \tilde{h_i}. \end{aligned} \tag{1}$$

where $\overrightarrow{h}_i$ is output of the i-th GRU cell whose input is $\mathbf{s}_i$. $\overrightarrow{h}_{i-1}$ is output of the $\{i-1\}$-th GRU cell. $W_z, W_r, W_{\tilde{h}}$ donate the learnable weight matrices and $\sigma(\cdot)$ donates the activation function.

For the hidden state $\{\overrightarrow{h}_i | i = t-k, \ldots, t\}$ for the forward GRU outputs and the hidden state $\{\overleftarrow{h}_i, | i = t-k, \ldots, t\}$ for the backward GRU outputs, the output of the bidirectional GRU: $\mathbf{H} = [h_{t-k}^T, \ldots, h_{t-1}^T, h_t^T]^T \in \mathbb{R}^{k \times (2d_c)}$, where $h_i = [\overrightarrow{h}_i, \overleftarrow{h}_i] \in \mathbb{R}^{1 \times (2d_c)}$, $i = t-k, \ldots, t$ and d_c donates the number of encoder features. GRU generates the feature representation more comprehensively by mining the history propagation features of the sequence.

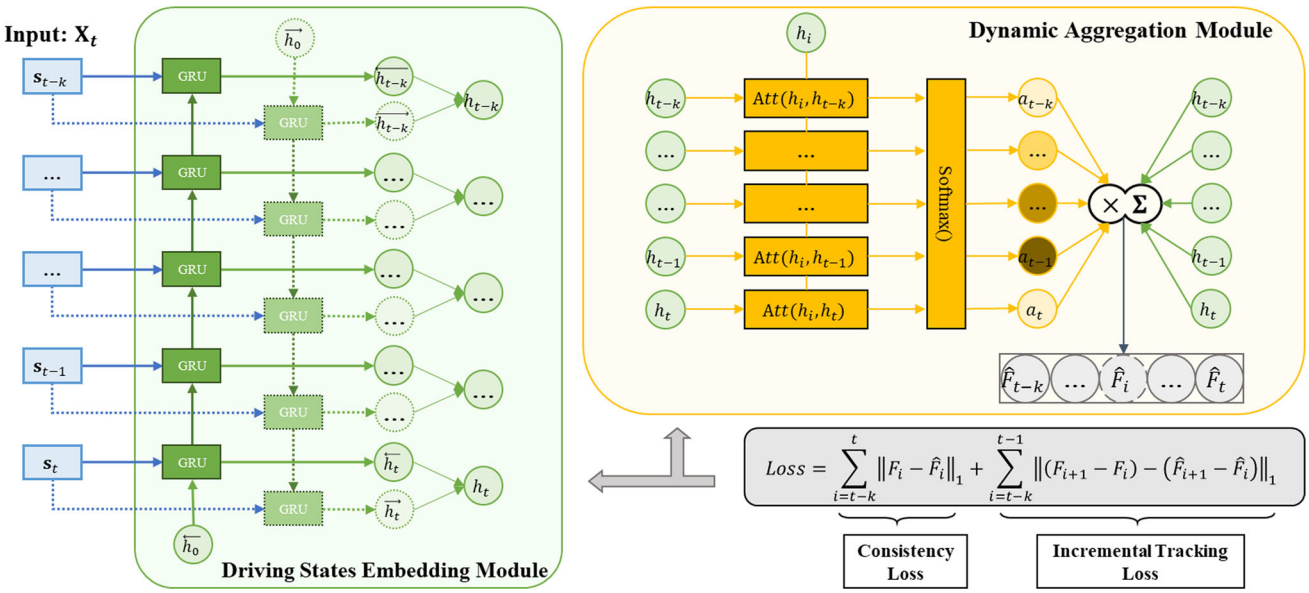

Fig. 3 Structure of proposed microscopic emission model. Driving states embedding module: Encoding vehicle driving state sequence to get the corresponding hidden states with GRU. Dynamic aggregation module: Calculating the attention weight for each hidden state in the encoding space to focus on the critical driving states and dynamic aggregate relevant features. Loss: Including the consistency loss of the generated emission and incremental tracking loss, supervising the model to generate the accurate sequence distribution

4.2 Dynamic aggregation module

The feature information in different time steps for driving states of vehicles is unequal. Dynamic aggregation module adopts attention mechanism [29] to assign more attention to the time step that contributes more to the estimated moment. As in Fig. 3, the attention layer calculates the attention weights

$$\boldsymbol{\alpha} = [\,\boldsymbol{\alpha}_t \quad \boldsymbol{\alpha}_{t-1} \quad \dots \quad \boldsymbol{\alpha}_{t-k}\,] = \begin{bmatrix} \alpha_{t,t} & \alpha_{t,t-1} & \dots & \alpha_{t,t-k} \\ \alpha_{t-1,t} & \alpha_{t-1,t-1} & \dots & \alpha_{t-1,t-k} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{t-k,t} & \alpha_{t-k,t-1} & \dots & \alpha_{t-k,t-k} \end{bmatrix}$$

of the hidden states $\mathbf{H}$ which is the output of GRU layer. For $\alpha_{i,j}$, $i,j = t-k,\dots,t$, the calculation formulas are as follow:

$$\alpha_{i,j} = softmax(\mathbf{score}(h_i, u_j)) = \frac{\exp(\mathbf{score}(h_i, u_j))}{\sum_{i=t-k}^{t} \exp(\mathbf{score}(h_i, u_j))} \tag{2}$$

where $u_j = \tanh(W_u h_j + b_u) \in \mathbb{R}^{1\times d_e}$ and score function $\mathbf{score}(h_i, h_j)$ is:

$$\mathbf{score}(h_i, u_j) = V^T \tanh(W_s h_i + U_s u_j) \tag{3}$$

where W_u, b_u, W_s, U_s and V are both learnable parameter vectors, d_e donates the number of decoder features. In this paper, $\boldsymbol{\alpha}_j$ represents the distribution of attention weight between the hidden states from $t-k$ to t and the hidden state at j. Therefore, the output of the attention layer is:

$$\mathbf{C} = [C_t, C_{t-1}, \dots, C_{t-k}] \in \mathbb{R}^{k\times 2d_c} \tag{4}$$

where $C_j = \sum_{i=t-k}^{t} \alpha_{i,j} h_i, j = t-k,\dots,t-1,t$. Finally, the emission factor from $(t-k)$-th to t-th time step, i.e., the output of decoder $\hat{\mathbf{F}} = [\hat{F}_{t-k}, \dots, \hat{F}_{t-1}, \hat{F}_t]$ is calculated as follow:

$$\hat{F}_i = \sigma(W_d[C_i, u_i] + b_d), i = t-k,\dots,t-1,t \tag{5}$$

where $W_d \in \mathbb{R}^{(2d_c+d_e)\times 1}$ and b_d are learnable parameters, and σ is sigmoid activate function.

4.3 Losses

In this paper, Φ donates all training parameters in the emission model. The objective function of the model has two components: first is the consistency loss

$$\mathcal{L}_s = \sum_{i=t-k}^{t} \|F_i - \hat{F}_i\|_1 \tag{6}$$

which decreases the mean absolute error between the generated emission factor sequence and ground-truth; another is the incremental tracking loss (ITL)

$$\mathcal{L}_d = \sum_{i=t-k}^{t-1} \|(F_{i+1} - F_i) - (\hat{F}_{i+1} - \hat{F}_i)\|_1 \quad (7)$$

which decreases the mean absolute error for the emission increment between adjacent time steps. The model with two loss functions of different magnitudes would face seesaw phenomena during optimization. To optimize the proposed model better, we adapt the strategy in [30] to balance the above two loss functions by introducing uncertainty weights. Therefore, the loss function is:

$$\mathcal{L}(\Phi) = \frac{1}{2\tau_s^2} \cdot \mathcal{L}_s + \frac{1}{2\tau_d^2} \cdot \mathcal{L}_d + \log \tau_s^2 \tau_d^2 \quad (8)$$

where τ_s and τ_d are optimized as parameters in training.

The number of features d_c and d_e of GRU is set to 512. The optimization algorithm for training is AdamW [31] method, and the learning rate is set to 0.001. When the loss of the model on the validation dataset is not further falling, the whole training process is terminated.

5 Experiments

5.1 Dataset description and preprocessing

The dataset evaluating the proposed method is collected from OBD mounted on a heavy-duty diesel vehicle in Hefei city, Anhui, China. Twenty-one trajectories of a vehicle from 2020/06/28 to 11/25 are selected, and there are 12628 samples whose sampling interval is 5 s. The average mileage per trajectory is approximately 12*km* and takes 56 min. Figure 4 shows some of the trajectories, most of which are distributed on the highways outside the city center.

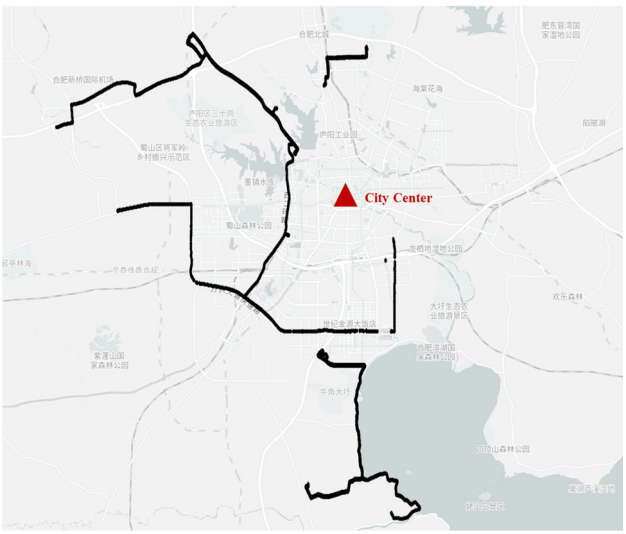

Fig. 4 The visualization of the trajectory distribution of the OBD dataset in Hefei. Heavy diesel vehicles mainly drive on the highways outside the city center

In data preprocessing, the traces whose valid values are less than 100 are deleted. And only the idle state of less than 180 s is retained at the beginning of each trajectory after removing invalid, abnormal, and null values. We do the maximum-minimum normalization for each attribute, respectively, before training [32, 33].

The effectiveness in estimating the high value of vehicles' emissions is also explored to confirm that our suggested model is peak-sensitive. According to the national standard "Limits and measurement methods for emissions from diesel fueled heavy-duty vehicles (CHINA VI) [34]" issued by Ministry of Ecology and Environment in 2018, the emission limit of NO_x from vehicles is 460 mg/kWh under the World Harmonized Transient Cycle (WHTC) [35]. So the samples with emission values higher than the limit are identified as high-emission samples.

Trajectories of dataset encompass different driving states like acceleration, deceleration, cruise, and idle along with multiple types of emission states, including high and low emissions. Table 1 illustrates the proportion of each state. Therefore, the data are reliable and representative, ensuring the validity verification of the proposed model.

In this study, the length of the sequence is set as 25, that is, k=24, and the batch size is 50. 80% of the samples in the dataset are divided into train data and 20% into test data. We perform k-fold cross-validation on training data. In each fold, 80% of the samples in the training data are used to train the model, and 20% of the samples are used for validation.

The proposed method is modeled using Pytorch [36], and all experiments are performed on a Linux cluster(CPU: Intel(R) Xeon(R) Silver 4216 CPUs, GPU: NVIDIA GeForce RTX 3090).

Table 1 The percentage of each driving state and emission state of vehicles in the dataset

Driving state			
Acceleration	Deceleration	Cruise[1]	Idle[2]
30.74%	28.75%	33.21%	7.29%
Emission state			
High emission		Low emission	
6.913%		93.09%	

[1]The driving state of a vehicle moving at a relatively steady speed

[2]The driving state of a vehicle that is not starting but its engine is idling

Table 2 The results of the proposed model and other baselines on OBD dataset

Model	For all			For high emission samples		
	MAPE(%)	RMSE	MAE	MAPE(%)	RMSE	MAE
FNN-no history [37]	44.96±2.882	80.78±1.362	57.00±1.175	48.18±1.964	257.7±8.602	240.9±9.337
FNN [37]	24.45±2.499	55.78±1.507	39.97±1.969	27.29±1.684	167.4±7.500	138.3±8.539
LSTM [25]	18.01±2.360	26.42±1.362	17.69±0.8054	11.12±0.8496	75.49±6.371	55.18±4.299
Bi-LSTM [38]	17.23±2.942	26.35±0.7204	17.62±0.4208	11.04±0.6255	75.34±4.872	54.68±3.054
GRU [28]	15.75±4.556	26.26±0.6194	17.13±0.3916	11.50±1.060	78.35±5.105	56.50±5.179
DL-VEM [11]	9.663±0.9352	19.09±2.187	9.445±1.794	8.237±0.8146	51.26±7.702	32.11±4.705
Transformer [39]	8.866±0.5006	25.99±1.162	17.12±0.8410	8.354±0.8241	38.25±4.132	24.98±2.566
Transformer+ITL	6.652±0.6281	20.27±1.098	13.12±1.148	6.238±0.6085	29.62±2.523	19.09±1.810
Our model	**2.608** ±0.2246	**7.182** ±1.013	**4.402** ±0.3291	**1.568** ±0.1721	**10.08** ±1.338	**7.246** ±0.8146

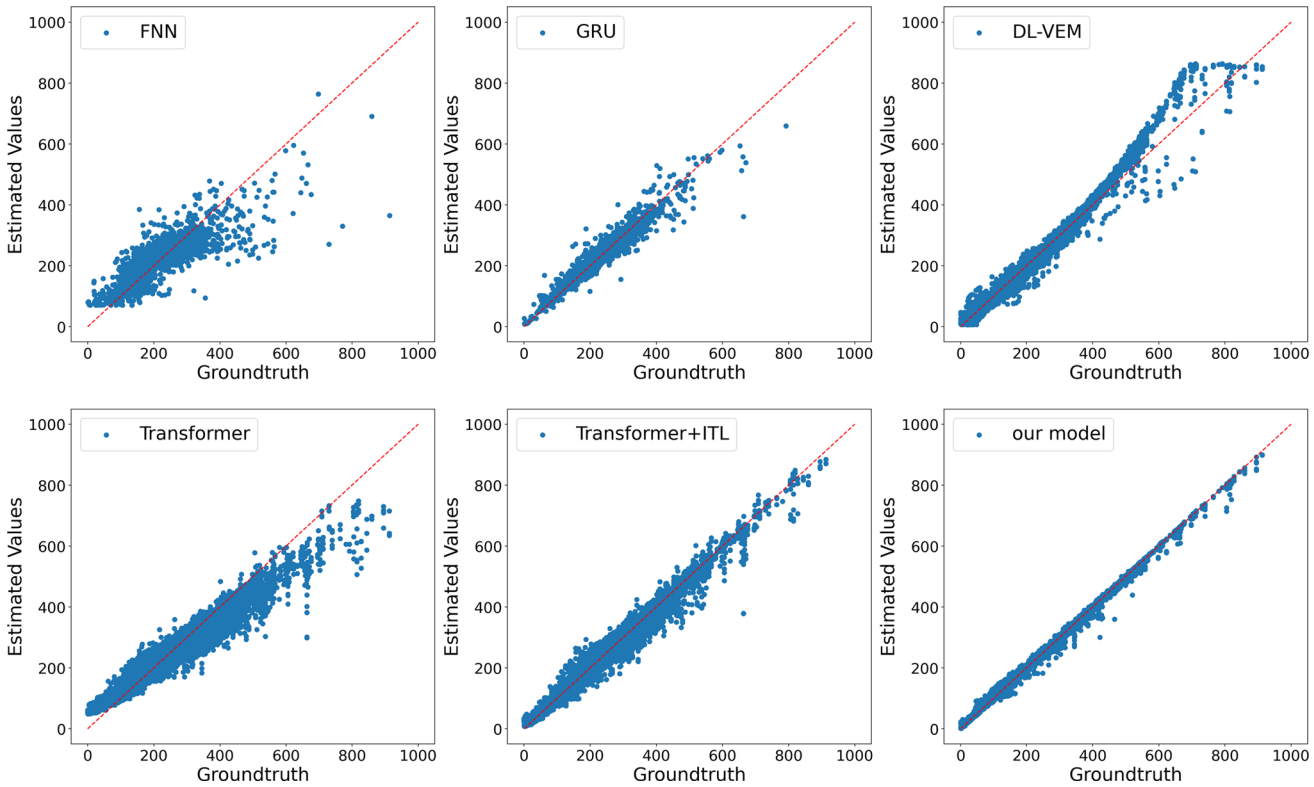

Fig. 5 Visualization of the estimation results of our proposed model and baselines on the test dataset

5.2 Results

5.2.1 Baselines and performance comparison

To verify the performance of our proposed model, we compared the following six models that are widely adopted deep learning models in existing research:

1. FNN [37]: Feedforward Neural Network consists of fully connected layers stacked on each other. In this paper, we adopt FNN to estimate emission values without and with the historical driving states.
2. LSTM [25]: Long Short-Term Memory is a variant of recurrent neural network (RNN) that captures long and temporal dependence of sequences.
3. Bi-LSTM [38] and GRU [28]: Bi-LSTM and Gated Recurrent Unit(GRU) are variants of RNN that can additionally capture the inverse temporality of time series.

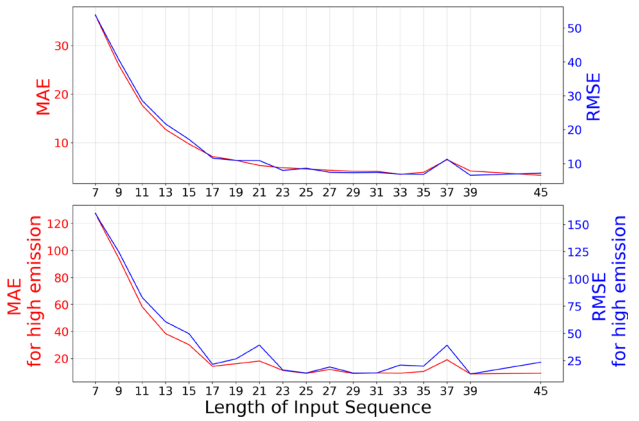

Fig. 6 Estimation errors MAE and RMSE of the proposed model with input sequences of different lengths as input

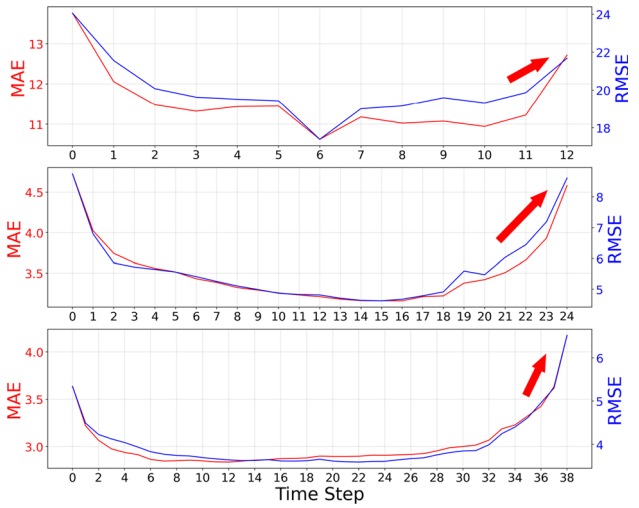

Fig. 7 The estimation errors MAE and RMSE for each time step of the emission sequence generated by the proposed model when the length of input sequences are 13, 25, and 39

4. DL-VEM [11]: A Deep Learning-based Vehicle Emission Model (DL-VEM) is an end-to-end deep learning CO_2 estimation framework that combines LSTM and FNN units.
5. Transformer [39]: Transformer is a sequence-to-sequence model based on a self-attention mechanism entirely. There are 2-layer encoders, 1-layer decoder, and 8 heads in experiments.

We first compare the performance of the baselines and our proposed model in emission estimation of vehicles on the OBD dataset. To measure and evaluate the performance of different methods, mean absolute percentage errors (MAPE), root mean squared errors (RMSE), and mean absolute errors (MAE) are adopted. Below are the formulas for the three metrics mentioned:

$$\begin{aligned} \text{MAPE} &= \frac{1}{m}\sum_{i=1}^{m}|F_t^i - \hat{F}_t^i|, \\ \text{RMSE} &= \sqrt{\frac{1}{m}\sum_{i=1}^{m}(F_t^i - \hat{F}_t^i)^2}, \\ \text{MAE} &= \frac{100\%}{m}\sum_{i=1}^{m}\left|\frac{F_t^i - \hat{F}_t^i}{F_t^i}\right|. \end{aligned} \tag{9}$$

Table 2 shows the error indexes of all models, where bold font indicates the minimum values. It observes that the proposed model significantly outperforms other baselines on test data, reducing the RMSE by 62% compared to DL-VEM. In particular, the estimation accuracy of the proposed model reduces the RMSE by 73.6% for the high-emission samples (about 10% of the test data) compared to Transformer. Figure 5 shows the estimation results of our proposed model and baselines. Baselines do not perform well in the high-emission samples, but the estimation results of our model are more accurate, thus providing excellent performance in both smooth and non-smooth driving states of vehicles.

Particularly, transformer fails to achieve competitive results on our dataset, including the high-emission samples. That enables global feature aggregation of sequences, thus realizing long-range temporal feature aggregation. However, the estimation for the instantaneous high emissions pays more attention to dynamic neighborhood features aggregation of driving state sequences. It needs to rely on a large amount of data for Transformer to realize this, yet high-emission samples are inadequate due to data imbalance. The effective aggregation of neighborhood features for transformer is further hindered.

5.2.2 Effects of different lengths of input sequence

Figure 6 shows the estimation results of the proposed model for input sequences with different lengths. When the

Table 3 The results of the proposed model with different input attributes

Attributes	For all			For high-emission samples		
	MAPE	RMSE	MAE	MAPE	RMSE	MAE
v+a	4.086%	11.71	6.992	2.804%	19.38	12.86
v+a+ewt	3.352%	10.14	6.343	3.051%	22.91	14.27
v+a+tpd	2.806%	8.661	4.639	1.787%	12.21	8.317
v+a+ewt+tpd	**2.608%**	**7.182**	**4.406**	**1.568%**	**10.08**	**7.246**

Table 4 The results of the proposed model with different modules

state2state	seq2seq	GRU	Attention	ITL	For all			For high-emission samples		
					MAPE	RMSE	MAE	MAPE	RMSE	MAE
√		√			15.75%	26.26	17.13	11.50%	78.35	56.50
	√	√			6.868%	15.60	10.78	8.215%	49.52	40.09
	√	√		√	4.381%	11.27	6.853	6.858%	44.80	33.95
	√	√	√		3.385%	9.129	5.832	2.242%	14.83	10.39
	√	√	√	√	**2.608%**	**7.182**	**4.406**	**1.568%**	**10.08**	**7.246**

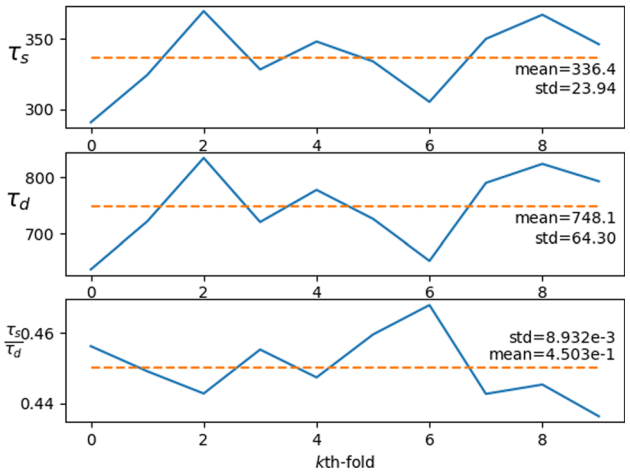

Fig. 8 The weight values τ_s and τ_d of loss function in K-fold experiment and ratio values of two weights

length is short, the result is not satisfactory. That is since the information in driving states is insufficient to analyze driving patterns and emission patterns. The longer the input sequence, the driving state information is saturated to make the accuracy of emission estimation no longer rise further. The time steps that have positive effects on the last moment are limited and the forward states would be forgotten by the model.

In addition, we find that the model has the best performance of emission estimation for the intermediate time steps even if the length of the input sequence is different, as shown in Fig. 7, since the intermediate moments are supervised by historical and future driving states. With the accumulation of the error of the incremental estimation, the emission errors of the last moment become larger.

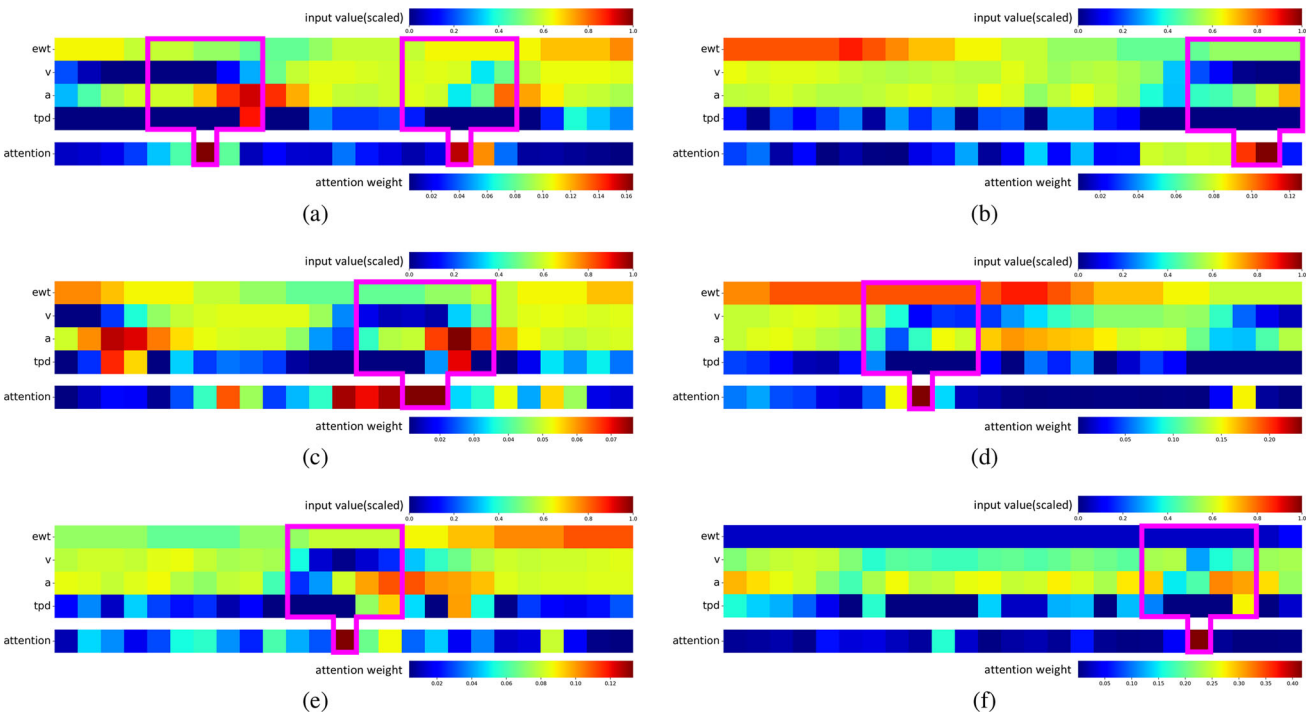

Fig. 9 Visualization of scaled values of the input sequence and generated attention weights for six high-emission samples. The attention mechanism assigns greater attentional weights to time steps in segments where the driving states are in flux

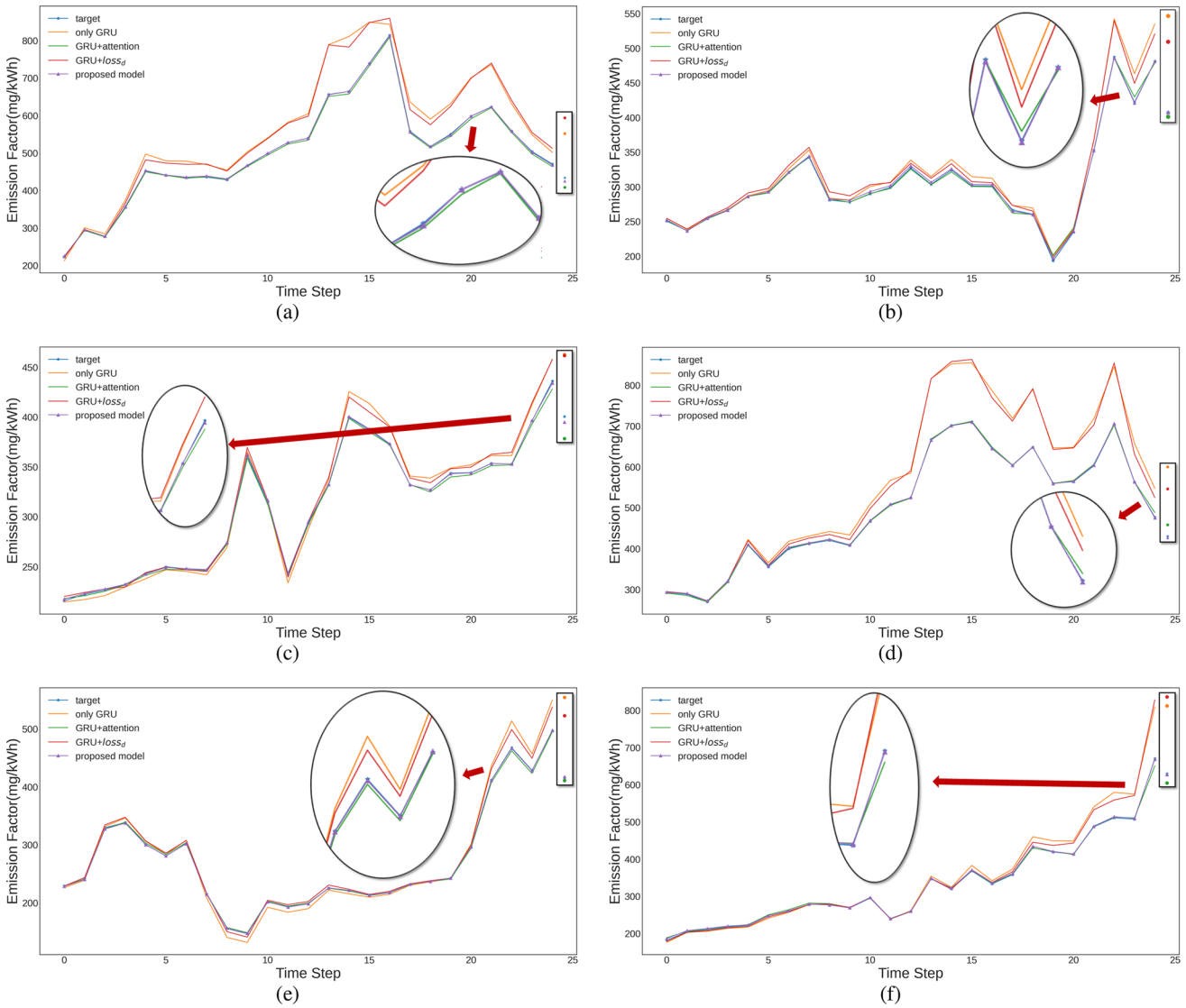

Fig. 10 Visualization of the ground-truth and generated emission sequence with different modules for high-emission samples consistent with Fig. 9. The blue curve indicates the groundtruth of the emission sequence; the orange curve indicates the generated emission sequence by the model with only GRU; the green curve indicates the generated emission sequence by the model with GRU and attention mechanism; the red curve indicates the generated emission sequence by the model with GRU and ITL; the purple curve indicates the generated emission sequence by the proposed model

5.2.3 Effects of different input attributes

To verify the contribution of different input attributes to emission estimation, the estimation accuracy with different attributes as inputs is shown in Table 3, where bold font indicates the minimum values. Particularly, the proposed model still performs well with no engine attributes. Moreover, the throttle pedal degree (*tpd*) is the most highly correlated with vehicle emissions, and the model obtains a 26% reduction in RMSE after adding the *tpd*, which is more than adding other attributes. The comparative results show that the model has wide generalizability and is not strict with data attributes. Even with only speed and acceleration, the model still performs well. In addition, more attributes related to the engine assist to determine the driving pattern more effectively and thus map out more reasonable emission patterns.

5.2.4 Effects of different modules

To verify the effectiveness of each component in the model proposed in this paper, i.e., seq2seq framework, attention mechanism, and ITL, we conducted comparative experiments shown in Table 4, where bold font indicates the minimum values.

1. **Attention Mechanism**: The seq2seq framework with only GRU is the baseline model, and the model

performance improves by 41.5% after introducing the attention mechanism (comparison of RMSE results in the 2nd and 4th rows of Table 4). In particular, the estimation error for high-emission samples is reduced by 70.0%. Obviously, the attention mechanism can effectively achieve high-emission pattern analysis of vehicles. Six high-emission samples are in visualization, and Fig. 9 shows the attention weights for the sequence of driving states, with the pink boxes framing the driving states and segments with the highest weights. The objects that the model focuses on are usually in the segments where the driving state changes, which indicates that high emissions of vehicles usually occur when the driving state is in large fluctuations.

2. **Incremental Tracking Loss**: When only GRU is in the framework, the model performance improves by 27.8% after introducing ITL (comparison of RMSE results in the 2nd and 3rd rows of Table 4). Meanwhile, the estimation error for high-emission samples decreases by only 9.53%. In contrast, when the loss function is introduced after the attention mechanism is adopted, the model performance improves by 21.3%, and the estimation error for the high-emission samples decreases by 32.0% (comparison of RMSE results in the 4th and 5th rows of Table 4). The values of τ_s and τ_d in each fold experiment are shown in Fig. 8, and their mean values are 336.4 and 748.1, respectively. The value of τ_s/τ_d is constant in different training datasets, although τ_s and τ_d fluctuate slightly.

The estimated emission sequences corresponding to the six high-emission samples in Fig. 9 are shown in Fig. 10. The model with no attention mechanism tracks the spike and rise phenomenon of the emission sequence unsatisfactorily, even if the ITL function is added since the supervisory information is not rich enough. With the introduction of the attention mechanism, the differential constraint can more accurately adjust the emission increments at adjacent times, thus achieving more accurate estimation. Especially, the accuracy of transformer after adding ITL function improves by 22% (comparison of RMSE results in the 7th and 8th rows of Table 2), and the weight distribution becomes concentrated. ITL function is more general and can effectively supervise the network to capture the mutations in the sequence.

6 Conclusion

In this paper, we address the problem of estimating instantaneous emissions for the vehicle under actual history and current driving conditions. Our proposed model introduces a sequence-to-sequence learning framework and an incremental tracking loss function to construct emission sequences of vehicles by mining the driving pattern, and supervising the emission estimation using the statistical properties of emission sequences and adjacency increment constraints. Finally, we evaluate the performance of the proposed model on OBD emission data of a heavy-duty diesel vehicle. The proposed model outperforms the existing state-of-the-art models for emission estimation and has an excellent performance in estimating high-emission values for vehicles. The results also show that the estimation accuracy of the proposed model is still considerable even when only the underlying driving state attributes are utilized without the engine attributes, which indicates the excellent generalization of the model.

In addition, this paper shows a U-shaped phenomenon due to error accumulation in emission sequence estimation. In the future, we will try to overcome the above difficulty and be able to further improve the performance of the proposed model.

Acknowledgements This work was supported in part by the National Natural Science Foundation of China (62033012, 61725304, 62103124), Major Special Science and Technology Project of Anhui, China (201903a07020012, 202003a07020009), and Postdoctoral Science Foundation of China (2021M703119).

Data availability The data that support the findings of this study are available from Department of Ecology and Environment of Anhui Province but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are, however, available from the authors upon reasonable request and with permission of Department of Ecology and Environment of Anhui Province.

Declarations

Conflict of interest The authors declare that they have no conflict of interests regarding the publication of this paper.

References

1. Smit R, Ntziachristos L, Boulter P (2010) Validation of road vehicle and traffic emission models-a review and meta-analysis. Atmos. Environ. 44:2943–2953
2. Muhammad AN, Aseere AM, Chiroma H, Shah H, Gital AY, Hashem IAT (2021) Deep learning application in smart cities: recent development, taxonomy, challenges and research prospects. Neural Comput. Appl. 33:2973–3009
3. Davis N, Lents J, Osses M, Nikkila N, Barth M (1939) Development and application of an international vehicle emissions model. Transport. Res. Record 2005:156–165
4. Koupal J, Cumberworth M, Michaels H, Beardsley M, Brzezinski D (2003) Design and implementation of moves: Epa's new generation mobile source emission model. Ann Arbor 1001:105
5. Cappiello A, Chabini I, Nam E. K. Lue A, Zeid M Abou (2002) A statistical model of vehicle emissions and fuel consumption, in:

Proceedings. The IEEE 5th International Conference on Intelligent Transportation Systems, IEEE, pp. 801–809
6. Scora G, Barth M (2006) Comprehensive modal emissions model (cmem), version 3.01, User guide. Centre for environmental research and technology. University of California, Riverside 1070, 79
7. Rakha H, Ahn K, Trani A (2004) Development of vt-micro model for estimating hot stabilized light duty vehicle and truck emissions. Transport Res Part D Transport Environ 9:49–74
8. Le Cornec CM, Molden N, van Reeuwijk M, Stettler ME (2020) Modelling of instantaneous emissions from diesel vehicles with a special focus on nox: insights from machine learning techniques. Sci Total Environ 737:139625
9. Seo J, Yun B, Park J, Park J, Shin M, Park S (2021) Prediction of instantaneous real-world emissions from diesel light-duty vehicles based on an integrated artificial neural network and vehicle dynamics model. Sci Total Environ 786:147359
10. Zhongqi W, Wei G, Jun B, Jun M (2016) Microscopic emission model of motor vehicle based on short-time real driving cycle, Chinese. J Environ Eng 10:5803–5807
11. Jia T, Zhang P, Chen B (2022) A microscopic model of vehicle CO2 emissions based on deep learning–a spatiotemporal analysis of taxicabs in Wuhan, China, IEEE Trans Intell Transport Syst
12. Lyu P, Wang PS, Liu Y, Wang Y (2021) Review of the studies on emission evaluation approaches for operating vehicles. J Traffic Transp Eng (English Edition) 8:493–509
13. De Haan P, Keller M (2004) Modelling fuel consumption and pollutant emissions based on real-world driving patterns: the hbefa approach. Int J Environ Pollut 22:240–258
14. Madziel M, Jaworski A, Kuszewski H, Wo's P, Campisi T, Lew K (2021) The development of co2 instantaneous emission model of full hybrid vehicle with the use of machine learning techniques. Energies 15:142
15. Pan Y, Zhang W, Niu S (2021) Emission modeling for new-energy buses in real-world driving with a deep learning-based approach. Atmos Pollut Res 12:101195
16. Moradi E (2021) A machine learning methodology for developing microscopic vehicular fuel consumption and emission models for local conditions using real-world measures, McGill University (Canada)
17. Motallebiaraghi F, Rabinowitz A, Holden J, Fong A, Jathar S, Bradley T, Asher ZD (2021) High-fidelity modeling of light-duty vehicle emission and fuel economy using deep neural networks, SAE Technical Paper 01–0181
18. Hong W, Chakraborty I, Wang H, Tao G (2021) Co-optimization scheme for the powertrain and exhaust emission control system of hybrid electric vehicles using future speed prediction. IEEE Trans Intell Vehi 6:533–545
19. Gao C, Zhang N, Li Y, Bian F, Wan H (2022) Self-attention-based time-variant neural networks for multi-step time series forecasting. Neural Comput Appl 34:8737–8754
20. Tian Y, Huang L, Yu H, Wu X, Li X, Wang K, Wang Z, Wang F (2022) Context-aware dynamic feature extraction for 3d object detection in point clouds. IEEE Trans Intell Transp Syst 23:10773–10785. https://doi.org/10.1109/TITS.2021.3095719
21. Almomani IM, Alkhalil NY, Ahmad EM, Jodeh RM (2011) Ubiquitous gps vehicle tracking and management system. In: 2011 IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEECT), IEEE, pp. 1–6
22. Rubino L, Bonnel P, Carriero M, Krasenbrink A (2010) Portable emission measurement system (pems) for heavy duty diesel vehicle pm measurement: the European pm pems program, SAE International Journal of Engines 2
23. Rimpas D, Papadakis A, Samarakou M (2020) Obd-ii sensor diagnostics for monitoring vehicle operation and consumption. Energy Reports 6:55–63
24. Moradi E, Miranda-Moreno L (2022) A mixed ensemble learning and time-series methodology for category-specific vehicular energy and emissions modeling. Sustainability 14:1900
25. Singh M, Dubey R (2021) Deep learning model based co2 emissions prediction using vehicle telematics sensors data, IEEE Trans Intell Veh
26. Baltusis P (2004) On board vehicle diagnostics. Technical Report, SAE Technical Paper
27. Zhang S, Zhao P, He L, Yang Y, Liu B, He W, Cheng Y, Liu Y, Liu S, Hu Q et al (2020) On-board monitoring (obm) for heavy-duty vehicle emissions in China: regulations, early-stage evaluation and policy recommendations. Sci Total Environ 731:139045
28. Dey R, Salem FM (2017) Gate-variants of gated recurrent unit (gru) neural networks. In: 2017 IEEE 60th international midwest symposium on circuits and systems (MWSCAS), pp. 1597–1600
29. Niu Z, Zhong G, Yu H (2021) A review on the attention mechanism of deep learning. Neurocomputing 452:48–62
30. Kendall A, Gal Y, Cipolla R (2018) Multi-task learning using uncertainty to weigh losses for scene geometry and semantics. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 7482–7491
31. Loshchilov I, Hutter F (2017) Decoupled weight decay regularization, arXiv preprint arXiv:1711.05101
32. Pei L, Cao Y, Kang Y, Xu Z, Zhao Z (2022) Uj-flac: Unsupervised joint feature learning and clustering for dynamic driving cycles construction. IEEE Trans Intell Transp Syst 23:10970–10982
33. Hansson A, Korsberg E, Maghsood R, Nordén E (2021) Lane-level map matching based on hmm. IEEE Trans Intell Veh 6:430–439
34. GB17691-2018: limits and measurement methods for emissions from disel fuelled heavy-duty vehicles(CHINA VI), Technical Report, China Environmental Science Press, 2018
35. Bai S, Han J, Liu M, Qin S, Wang G, Li G-X (2018) Experimental investigation of exhaust thermal management on nox emissions of heavy-duty diesel engine under the world harmonized transient cycle (whtc). Appl Therm Eng 142:421–432
36. Paszke A, Gross S, Massa F., Lerer A, Bradbury J, Chanan G, Killeen T, Lin Z, Gimelshein N, Antiga L, et al. (2019) Pytorch: An imperative style, high-performance deep learning library, Advances in neural information processing systems 32
37. Wang J, Wang L, Ji Z, Qi S, Xie Z, Yang Z, Zhang X (2021) Research on actual road emission prediction model of heavy-duty diesel vehicles based on obd remote method and artificial neural network. In: Journal of Physics: Conference Series, volume 2005, IOP Publishing, p. 012174
38. Huang Z, Xu W, Yu K (2015) Bidirectional lstm-crf models for sequence tagging, arXiv preprint arXiv:1508.01991
39. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser Ł, Polosukhin I (2017) Attention is all you need, Adv Neural Inf Process Syst 30