Frontiers of Information Technology & Electronic Engineering
www.jzus.zju.edu.cn; engineering.cae.cn; www.springerlink.com
ISSN 2095-9184 (print); ISSN 2095-9230 (online)
E-mail: jzus@zju.edu.cn

Supplementary materials

Zhenyi XU, Ruibin WANG, Yang CAO, Xiushan XIA, Yu KANG 2022. Amending the COPERT model for heavy-duty vehicle emissions using a time frequency fusion network. *Front Inform Technol Electron Eng*. https://doi.org/10.1631/FITEE.2200218

1 Overview of PEMS and OBD Research

Quirama et al. (2020) used PEMS to construct an energy-based micro-trip operating mode and estimate the actual energy consumption and exhaust emissions of the fleet in a given area. Tsinghua University has developed the Beijing vehicle emission factor model in conjunction with other research units, which is known as the EMission factor model for the BEijing Vehicle fleet (EMBEV), and it has been formulated by selecting mature emission models considered from the international literature, and processing these appropriately to ensure adaptive compatibility with local conditions(Yang et al. (2019)). Wang et al. (2017) used a sequential decision-making strategy to estimate the speed of the road segment based on the low-frequency GPS trajectory of the vehicle, and combining this into the application of the micro-emission model, developed a method to estimate the vehicle carbon dioxide emissions. Feng et al. (2019) constructed an emission model for localized bus conditions in Haikou City based on cluster analysis of micro-trip segments, and used PEMS to verify the emission. WANG et al. (2016) considered the influence of the historical operating state of the vehicle, and constructed a microscopic emission model based on BP neural network by using short-term driving conditions. Xu et al. (2017) built a mobile source emission prediction model based on deep neural network to realize the relationship mapping between vehicle transient operating conditions and pollution emissions, and further , Xu et al. (2021a) proposes a method of gauging the exhaust gas telemetry data depth based on the COPERT emission factor.However, the COPERT model requires to be adapted to the local environment; and by constructing a three-layer self-encoder network, the feature extraction of multi-source heterogeneous data such as meteorological data, road network data, and data pertaining to traffic flow (especially in urban areas where functional means to gauge real-time emission data are available) can be realized, and accordingly the COPERT emission model can be corrected. Yang et al. (2016) used remote on-board diagnostic (OBD) data to analyze the actual CO_2 and NOx emissions of National IV buses, and an emission calculation method for these pollutants was established based on OBD data. Mechanisms for the estimation of pollution-causing emissions have immensely benefitted from the rapid development of artificial intelligence technology, since it has brought new theoretical knowledge and technical methods for research and analysis in traditional industries. In recent years, some scholars and research institutions have also conducted exploratory research based on OBD vehicle network online monitoring methods: Xu et al. (2021b) implemented transfer of models between different vehicles using a deep transfer network based on the effect of multiple sources of external factors on mobile source emissions. Xun et al. (2019) conducted experiments by reading and analyzing data from the OBD-II scanning tool in real time, the obtained data were processed using feature scaling and principal component analysis (PCA), and the algorithm employed to identify drivers used k-nearest neighbor (k-NN) and plain Bayesian techniques combined with a voting mechanism, the emergent experimental results show that the accuracy can reach 100%. Using the best attributes of the training model, which were identified from the OBD II data, Molina Campoverde (2019) employed random forest techniques to determine the highest inference of the driver variables. Wang et al. (2021) used OBD and artificial neural network (ANN) techniques to predict NOx and PM emissions of heavy-duty diesel trucks under actual road conditions, and by comparing the predicted values against the actual ones derived from PEMS test results, demonstrated the prevalence of a good correlation between them. Chen et al. (2015) proposed a new driving behavior analysis method based on vehicle on-board diagnostic (OBD) information and AdaBoost algorithm, and the experimental results showed the correctness of the proposed driving behavior analysis method.

2 Dataset Description

The names of the data attributes collected are the following: license plate, terminal number, date, engine speed (rpm), actual output torque percentage (%), engine water temperature(°C), engine fuel temperature(°C), post-treatment downstream NOx(ppm), post-treatment downstream oxygen(%), atmospheric pressure(kPa), environment temperature(°C), post-treatment exhaust gas mass flow rate(kg/h), urea tank level(%), urea tank temperature(°C), vehicle speed (km/h), gas pedal opening(%), single driving miles(km), total driving mileage(km), engine instantaneous fuel injection(L), engine instantaneous fuel consumption rate(L/100km), average engine fuel consumption rate, engine fuel consumption for a single trip, and total engine fuel consumption(L/100km), battery voltage(mv), fuel tank level, cumulative engine running time(h), longitude, latitude, SCR upstream temperature(°C), and SCR downstream temperature(°C).

The first step in the processing of the collected raw data was the deletion of irrelevant attributes. Then the abnormal records with values of 0 for vehicle speed, NOx value downstream of post-processing and instantaneous engine fuel consumption rate were deleted. The attributes and corresponding symbols contained in each data set after pre-processing are shown in Table S1.

Table. S1 Attribute comparison table

Name	Label	Name	Label
Engine speed	E_{speed}	Actual output torque percentage	AOTP
Engine water temperature	EWT	Engine fuel temperature	EFT
Post-treatment downstream NOx	NOx	Post-treatment downstream oxygen	O_2
Atmospheric pressure	AP	Environmental temperature	ET
Post-treatment exhaust gas mass flow rate	PT_{EGM}	Urea tank level	UTL
Urea tank temperature	UTT	Vehicle speed	V_{speed}
Gas pedal opening	GPO	Single driving miles	SDM
Engine fuel consumption rate(instantaneous)	EFCR	Average engine fuel consumption rate	$EFCR_{avg}$
Engine fuel consumption for single driving	EFC_{SD}	Total engine fuel consumption	EFC_{total}
Battery voltage	BV	Fuel tank level	FTL
Cumulative engine runtime	CER	Longitude	LNG
Latitude	LAT	/	/

3 Data Processing

3.1) Driving segment division Since the dataset is collected from the actual driving records of a diesel vehicle over multiple days, it would be comprised of multiple consecutive driving segments, resulting in the overall data records being presented in a manner that is not strictly continuous. Furthermore, in the preprocessing stage, we deleted some useless records, causing a split in the continuous driving segments and continuing the split in the records that were already not strictly continuous. However, the records immediately before and after the deleted invalid records can be considered continuous. Accordingly, we need to set the maximum value of a reasonable continuous time interval when dividing the whole dataset into continuous driving segments, so that it becomes possible to ignore the effect of deleting invalid data in the pre-processing stage without causing a compromise in the meaningfulness and usability of the overall dataset.

In the present study, we set the maximum value of a time interval as 180 s, and the adjacent records with an interval greater than 180 s are considered to belong to different driving segments; after the initial screening, it was observed that each driving segment contained a different number of records, and accordingly, to ensure that the requirements of the subsequent emission factor calculation based on the historical information pertaining to length would be satisfied, each selected driving segment was endeavored to be made as long as possible; thus, we set the minimum number of records of driving segments as 180, i.e., 15 min. The data volume of each segment is shown in Table S2.

Table. S2 Valid driving clip display

Clip number	Sample size	Clip number	Sample size	Clip number	Sample size	Clip number	Sample size
clip 1	580	clip 2	654	clip 3	314	clip 4	813
clip 5	342	clip 6	781	clip 7	258	clip 8	218
clip 9	261	clip 10	604	clip 11	303	clip 12	190
clip 13	371	clip 14	485	clip 15	362	clip 16	262
clip 17	242	clip 18	585	clip 19	301	clip 20	261
clip 21	226	clip 22	204	clip 23	287	clip 24	256
clip 25	236	clip 26	322	clip 27	281	clip 28	185
clip 29	435	clip 30	447	clip 31	453	clip 32	461
clip 33	206	clip 34	217	clip 35	398	clip 36	428
clip 37	378	clip 38	285	clip 39	203	clip 40	495
clip 41	470	clip 42	271	clip 43	258	clip 44	209
clip 45	522	clip 46	232	clip 47	198	clip 48	193

3.2) Obtaining the emission factor of NOx In carrying out the process of emission factor acquisition, it is necessary to keep in view the principle that the emission factor of the kth record we seek is jointly determined by records extending from the first record to the kth record, and the details of the selection for k will be further described in Section 4. The emission factors for records 1 to k - 1 are not available for each driving segment, and so we will start from the kth record of each driving segment and align the OBD emission factor EF_OBD and the COPERT emission factor EF_{COPERT} with the original attributes to obtain the complete data.The engine fuel volume flow rate Q_{FR} read by OBD is not recorded in the dataset of this paper; ths,we calculated its value with the help of EFCR and V_{Speed}, and its calculation formula is as follows.

$$Q_{FR} = 0.01 \cdot EFCR \cdot V_{Speed} \tag{S1}$$

By using Eqs(S1)-(S3), EF_{COPERT} and EF_OBD were calculated. The results of emission factors calculation for each driving segment are shown in Fig.S1 (with k = 48 as an example). It can be seen that most of the EF_{COPERT} values have lower peaks than EF_OBD and higher troughs than EF_OBD, i.e., the COPERT model is targeted by the COPER model and the NOx emission factors calculated by the OBD model are bfundamentally unequal in terms of values. Therefore, the emission factors calculated using the COPERT model are not applicable to the OBD data.

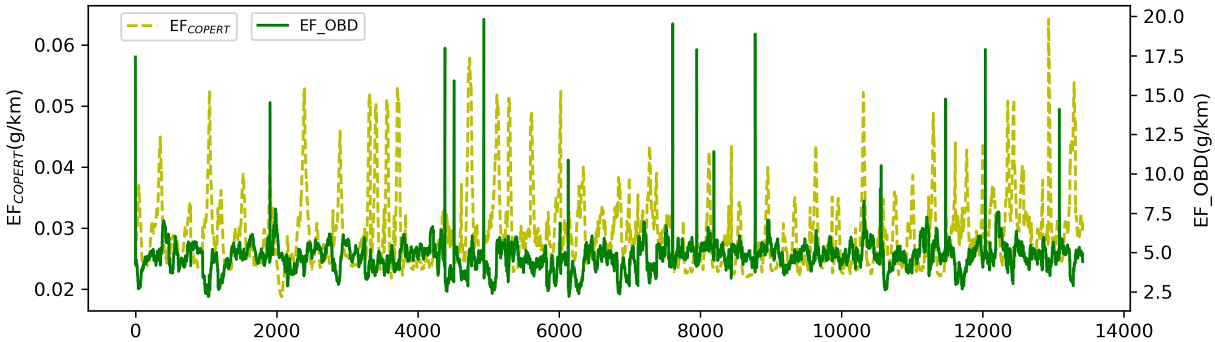

Fig. S1 Emission factors of NOx.

4 Formulaic description of Quality Index (QI_{avg})

$$QI_{avg} = \frac{QI_{sum}}{count(feature)+count(train,val)} = \frac{\sum_{train,val}\sum_{feaature}(QI)}{count(feature)+count(train,val)} \tag{S2}$$

$$QI = \frac{IE+LB}{2} \tag{S3}$$

$$IE = -\sum_{i=0}^{255} p(i)\log_2 p(i) \tag{S4}$$

$$LB = \sum\nolimits_x \sum\nolimits_y |G(x,y)| \tag{S5}$$

where $count(feature)$ represents the number of features, $count(train, val)$ represents the quantum of data in the training and validation sets, $p(i) = count(p = i)/(m \times n)$ represents the proportion of grayscale with pixel value i in a picture of size m*n with grayscale 256 (0 - 255), $count(p = i)$ is the frequency of pixel value i in the picture, $G(x, y)$ is the convolution of Laplacian operator at pixel point (x,y), and Laplacian operator $L = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}$.

References

Chen, S.-H., Pan, J.-S., Lu, K., et al., 2015. Driving behavior analysis based on vehicle obd information and adaboost algorithms. In: Proceedings of the international multiconference of engineers and computer scientists. Vol. 1. pp. 18–20.

Feng, L. I., Zhuang, J., Mengliang, L. I., 2019. Construction of driving cycle for investigation of public transport emission in haikou city. Journal of Chongqing University of Technology(Natural Science).

Molina Campoverde, J. J., 2019. Driving mode estimation model based in machine learning through pids signals analysis obtained from obd ii. In: International Conference on Applied Technologies. Springer, pp. 80–91.

Quirama, L. F., Giraldo, M., Huertas, J. I., Jaller, M., 2020. Driving cycles that reproduce driving patterns, energy consumptions and tailpipe emissions. Transportation Research Part D: Transport and Environment 82, 102294.

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. Vol. 2005. IOP Publishing, p. 012174.

Wang, S., Li, Z., Tan, J., Guo, W., Li, L., 2017. A method for estimating carbon dioxide emissions based on low frequency gps trajectories. In: 2017 Chinese Automation Congress (CAC). IEEE, pp. 1960–1964.

WANG, Z., GUAN, W., BI, J., MA, J., 2016. Microscopic emission model of motor vehicle based on short-time real driving cycle. Chinese Journal of Environmental Engineering 10 (10), 5803–5807.

Xu, Z., Kang, Y., Cao, Y., Li, Z., 2021a. Deep amended copert model for regional vehicle emission prediction. Science China Information Sciences 64 (3), 1–3.

Xu, Z., Kang, Y., Lv, W., 2017. Analysis and prediction of vehicle exhaust emission using ann. In: 2017 36th chinese control conference (ccc). IEEE, pp. 4029–4033.

Xu, Z., Wang, R., Kang, Y., Zhang, Y., Xia, X., Wang, R., 2021b. A deep transfer nox emission inversion model of diesel vehicles with multisource external influence. Journal of Advanced Transportation 2021.

Xun, Y., Sun, Y., Liu, J., 2019. An experimental study towards driver identification for intelligent and connected vehicles. In: ICC 2019-2019 IEEE International Conference on Communications (ICC). IEEE, pp. 1–6.

Yang, D., Zhang, S., Niu, T., Wang, Y., Xu, H., Zhang, K. M., Wu, Y., 2019. High-resolution mapping of vehicle emissions of atmospheric pollutants based on large-scale, real-world traffic datasets. Atmospheric Chemistry and Physics 19 (13), 8831–8843.

Yang, L., Zhang, S., Wu, Y., Chen, Q., Niu, T., Huang, X., Zhang, S., Zhang, L., Zhou, Y., Hao, J., 2016. Evaluating real-world co2 and nox emissions for public transit buses using a remote wireless on-board diagnostic (obd) approach. Environmental pollution 218, 453–462.