Deciphering urban traffic impacts on air quality by deep learning and emission inventory

Wenjie Du[a,d,#], Lianliang Chen[b,c,#], Haoran Wang[a,d], Ziyang Shan[a,d], Zhengyang Zhou[c,d], Wenwei Li[e,f], Yang Wang[a,c,*]

[a]School of Software Engineering, University of Science and Technology of China, Hefei 230026, China.

[b] Alibaba Inc., Hangzhou 310052, China.

[c]School of Computer Science and Technology, University of Science and Technology of China, Hefei 230026, China.

[d]188 RenAi Road, Suzhou Institute for Advanced Research, University of Science and Technology of China, Suzhou, Jiangsu 215123, China.

[e]CAS Key Laboratory of Urban Pollutant Conversion, Department of Applied Chemistry, University of Science and Technology of China, Hefei 230026, China.

[f]USTC-CityU Joint Advanced Research Center, Suzhou, 215123, China.

[*] Corresponding author at 188 RenAi Road, Suzhou Industry Park, Suzhou, Jiangsu, 215123, China.

[#] Wenjie Du and Lianliang Chen contributed equally to this manuscript.

E-mail address: angyan@ustc.edu.cn (Y. Wang).

ABSTRACT

Air pollution is a major obstacle to future sustainability, and traffic pollution has become a large drag on the sustainable developments of future metropolises. Here, combined with the large volume of real-time monitoring data, we propose a deep learning model, iDeepAir, to predict surface-level $PM_{2.5}$ concentration in Shanghai megacity and linked with emission inventory creatively to decipher urban traffic impacts on air quality. Our model exhibits high-fidelity in reproducing pollutant concentrations and reduces the MAE by 20% compared with other models and identifies the ranking of major factors. Local meteorological conditions have become a nonnegligible factor. Layer-wise relevance propagation (LRP) is used here to enhance the interpretability of the model and we visualized and analyzed the reasons for the different correlation between traffic density and $PM_{2.5}$ concentration in various regions of Shanghai. Meanwhile, As the strict and effective industrial emission reduction measurements implementing in China, the contribution of urban traffic to $PM_{2.5}$ formation is gradually increasing from 21.62% in 2010 to 35.67% in 2017 in Shanghai, and the impact of traffic emissions would be ever-prominent in 2030 according to our prediction. We also infer that the promotion of vehicular electrification would achieve further alleviation of $PM_{2.5}$ about 11.72% and reduce 11.72% by 2030 gradually. These insights are of great significance to provide the decision-making basis for accurate and high-efficient traffic management and urban pollution control, and eventually benefit people's lives and high-quality sustainable developments of cities.

Keywords: PM2.5 concentration forecast; Traffic emissions; Deep learning; Attention mechanism; New energy vehicles.

1. Introduction

Air pollution is a large obstacle to the world's future sustainable developments, and millions of people die from air pollution-related diseases every year around the world (Zheng *et al.*, 2017). This is seriously severe in some developing countries like China (Lelieveld *et al.*, 2015), which has the highest country-level values globally for the population-weighted annual average concentration of $PM_{2.5}$ (Tichenor and Sridhar, 2019; Zhang *et al.*, 2012)and has been a major public health concern in recent years (Li *et al.*, 2019a). Shanghai, one of the most developed and populous cities in China, has suffered severe increasing haze episodes mostly attributed to the severe particle pollution especially fine particles (particles ≤2.5μm in aerodynamic diameter; $PM_{2.5}$) (Li *et al.*, 2019a) since 1990s with the rapid urbanization and industrialization (Wang *et al.*, 2015a). Under these circumstances, air pollution-related diseases have emerged gradually, such as respiratory diseases in the elderly and preterm birth and low birth weight for birth when maternal exposure to $PM_{2.5}$ in Shanghai (Liu *et al.*, 2017).

Going further, the total number of vehicles in China has exceeded 200 million in 2019 and increases by more than 20 million annually in the urban. Vehicular traffic is a principal source of air pollutants such as nitrogen oxides (NO_x), carbon monoxide (CO) and carbonaceous particles (Zhang and Batterman, 2010). Traffic emission has become one of the important factors affecting air quality due to the extensive motor vehicles in China (Yan *et al.*, 2020). Meanwhile, some secondary pollutants discharged like O_3 , SO_2 and a major portion of $PM_{2.5}$ are generally diverse in different regions and time (Kroll *et al.*, 2020). When they involve different changes or conditions, they could promote or alleviate the formation of $PM_{2.5}$ in varying degrees such as high relative humidity promoting the formation of $PM_{2.5}$ (Benas *et al.*, 2013), higher temperature enhancing the photochemical reaction in the atmosphere (Dumka *et al.*, 2015), wind contributing greatly to diffuse particulate matter (Xiao *et al.*, 2011), which makes it difficult to trace and analyze the causes of local air pollution and the major drivers (Le *et al.*, 2020) in a specific area. Generally speaking, atmospheric chemical reactions serve as essential nonlinear links between traffic emissions and atmospheric composition (Yang *et al.*, 2021a; Zhu *et al.*, 2021). Meanwhile, local meteorological factors, for instance, air temperature, wind-field (Zhou *et al.*, 2021), humidity, and so on also strongly regulate photochemical formation of ozone and PM (Le, Wang, Liu, Yang, Yung, Li and Seinfeld, 2020; Wu *et al.*, 2020; Yang, Wen, Wang, Zhang, Pinto, Pennington, Wang, Wu, Sander, Jiang, Hao, Yung and Seinfeld, 2021a). Here, we disentangle the complex factors involving emissions inventory, transport emission, and meteorological conditions to evaluate the effect of different factors on air quality in urban area by deep-learning.

To disentangle the complex factors, we focus on the issue of pollution tracing by deep learning. Relatively speaking, traditional source apportionment (SA) methods which mainly based on receptor-oriented model and source-oriented model (Huang *et al.*, 2014; Zhang *et al.*, 2017), are flawed for $PM_{2.5}$ source identification and traffic emission impact evaluation for the following reasons: a) most parameters in these models, which are determined from the laboratory (Julie *et al.*, 2016), cannot accurately reflect the real scene, and it is quite time-consuming to get these parameters; b) many important influential factors such as meteorological conditions, pollutant discharge, secondary chemical substances formed during pollutants diffusion are not considered, hence results in poor accuracy of the simulating results (Kumar *et al.*, 2018; Li *et al.*, 2017a; Souri *et al.*, 2016); c) the validity of quantifying impacts of different anthropogenic emissions remains uncertain (Daskalakis *et al.*, 2016; Gao *et al.*, 2018; Kumar, Peuch, Crawford and Brasseur, 2018). Compared with traditional chemical transport modeling, the data-driven based data analysis method has more flexibility in leveraging real-world data and could better fit nonlinear relationship (Yang *et al.*, 2021a; Zhu *et al.*, 2021) which has been considered as a new perspective to conduct environment-related research in recent years (Alfaseeh *et al.*, 2020; Wang *et al.*, 2020) since this technology is able to simulate complex pollution formation mechanisms by focusing on data itself.

Some researchers think that deep learning is suitable for analyzing air pollution(Hino *et al.*, 2018; Xing *et al.*, 2020), such as $PM_{2.5}$ concentration prediction via interpretable convolutional neural networks (Park *et al.*, 2020; Zhou, Zhang, Du and Liu, 2021), assessing traffic impacts by random forest (Yang, Wen, Wang, Zhang, Pinto, Pennington, Wang, Wu, Sander, Jiang, Hao, Yung and Seinfeld, 2021a), air quality prediction in an image-based deep learning model (Zhang *et al.*, 2020). Actually, deep learning technologies has demonstrated its strong ability and applied in multiple fields including intelligent driving (Zhang *et al.*, 2018), intelligent medical (Li *et al.*, 2019b; Lindsey *et al.*, 2018) , life science studies (Anonymous, 2019; Ham *et al.*, 2019; Yuan and Bar-Joseph, 2019).

In order to enhance the interpretability of the model, Layer-wise relevance propagation is used here. LRP (Bach *et al.*, 2015a), a method for explaining the predictions of a broad class of machine learning models, has been widely applied and well verified in many scenarios, such as machine translation, emotion analysis and text classification text classification (Arras *et al.*, 2017a; Arras *et al.*, b; Ding *et al.*, 2017). According to the contribution of neurons in the former layer to the latter layer, all relevant values in the latter layer are allocated to the former layer and pushed back to the input layer. In the process of pushing back, this follows the conservation principle (Lapuschkin *et al.*, 2019).

In this paper, we propose a novel deep learning network, iDeepAir, to decipher the impact of urban traffic on air quality by combining the multi source real-time data such as traffic information, in situ surface-level pollutant concentrations and meteorology. We assess the sensitivity of $PM_{2.5}$ in Shanghai to traffic emission changes at different stages by comparing predicted concentrations under different traffic emission scenarios. Then, we utilize iDeepAir model to account for the nonlinear interactions among different input parameters to fit the complex chemical reaction and temporal accumulation procedure of $PM_{2.5}$ formation. Furthermore, with the embedded LRP algorithm (Bach, Binder, Montavon, Klauschen, Mueller and Samek, 2015a; Lapuschkin, Waeldchen, Binder, Montavon, Samek and Mueller, 2019), the contribution of each pollutant on the formation of $PM_{2.5}$ can be quantified clearly and separately which enhance the interpretability of the model. Moreover, to quantify the contribution of anthropogenic emissions to $PM_{2.5}$ development for each year from 2010 to 2017 we couple the emission inventories with the iDeepAir. And finally, considering the development of new energy transportation policy and traffic emissions in the future, we assess the possible benefits of future traffic evolution on $PM_{2.5}$ reductions and derive some potential impacts of new energy transportation policy in 2030.

2. Material and methods

2.1. Study area and datasets

Thanks to the development of environmental monitoring technology, we can obtain a large number of historical monitoring data for academic research. The iDeepAir model based on observation values connects multiple feature time series data to predict $PM_{2.5}$ concentration. The study focuses on Shanghai region, which is one of the most developed area in China including 16 districts with different terrain and population density. The location information of monitoring points and road network are present in Fig.1.

The dataset used in this research consists of four parts: transportation-related data, air quality data, meteorological data and pollutant emission load related data. The detailed composition of the dataset and related statistical results are presented in Table 1. (1) Transportation-related data. This dataset includes two parts: Total numbers of both petrol and new energy vehicles and Traffic State

Indexes (TSIs). Total numbers of petrol and new energy vehicles, which are from Shanghai Statistical Yearbook (http://tjj.sh.gov.cn) and Shanghai Traffic Comprehensive Annual Report. TSIs, which can be obtained from Shanghai Traffic Information Platform (http://www.jtcx.sh.cn), includes 68 traffic state indexes reflecting the real-time traffic status in different regions of the city. Regarding one specific region, this index is calculated with real-time road traffic status that is collected from intraregional road segments every 2 minutes (Text S1). (2) Air quality data. This data, which contains the hourly average monitored air quality data of Shanghai, is collected from the Real-time Air Quality Reporting System (http://219.233.250.38:8087/AQI/siteAQI.aspx). The detailed elements of monitored data include Time, $PM_{2.5}$, PM_{10}, O_3, SO_2, NO_2, CO, and all contaminant concentrations are recorded in micrograms per cubic meter ($\mu g/m^3$). (3) Meteorological data. The meteorological data of Shanghai is collected from the platform of Weather Underground, by taking the Shanghai Hongqiao international Airport Monitoring Station as the reference point (https://www.wunderground.com/weather/ZSSS). The detailed categories of each meteorological record are shown in Table 1 and the sampling interval of meteorological data is 30 minutes. (4) Pollutant emission load related data. The annual total pollutant emission loads from the sectors of industry, resident, and transportation can be found from Shanghai Municipal Statistics Bureau (http://tjj.sh.gov.cn, http://www.stats.gov.cn). The annual generated electricity in Shanghai, which can be used to infer the emissions of the power sector.

All data are one-hour interval, with a total of 5969 records. The datasets are divided randomly into a training set (80%), and a test set (20%).

2.2. iDeepAir architecture

The iDeepAir adopts an encoder-decoder architecture, which consists of three kinds of modules: Traffic Fusion Module (TFM), Feature Interaction Module (FIM) and Time Interaction Module (TIM) (Fig. 2). (a) Traffic Fusion Module. We employ an independent TFM to exploit the spatial dependence between the road traffic status and air pollution with its multiple embedded convolutional layers, and the results are combined with geographic information to present the spatial distribution of the dependencies within traffic status and air pollution. (b) Feature Interaction Module. In this module, we simulate the complex chemical reactions between different contaminants by employing the Multi-Head Attention structure(Vaswani *et al.*, 2017) to achieve the interactions between different features. Considering that not all of these features can contribute to the prediction of air pollution, we then introduce traditional attention mechanism into this module to highlight key features. By combining the self-attention and attention mechanism, this module can not only enable the interactions between different features, but also can address sequential forecasting problems with long-term dependency. Besides, this can achieve higher accuracy forecasting. (c) Time Interaction Module. We use Temporal Convolutional Network (TCN) to capture the long-term and deep interactions between pollutant ingredients in the temporal dimension and simulate the pollutant accumulation processes over time in chemical reactions with the inputs of fine-grained meteorological records and learned feature interactions, and finally generate the sequence of $PM_{2.5}$ concentration in next 24 hours (Text S3). TCN contains three sub-structures: causal convolution, dilation convolution and residual connection (Text S3).

2.3. iDeepAir training algorithm

We implement and train the iDeepAir network with the deep learning toolkits Keras (version 2.2.4) and Tensorflow (version 1.10.0) in Python (version 3.6.6). The training process is performed on Tesla V100-PCIE GPU, running under the CentOS Linux 7 server. During the training phase, the batch size is 128 and the learning rate is 0.001. Adam optimizer is adopted and the objective loss function is formulated as follows:

$$Loss = \|\hat{y} - y\|_2^2$$

where $\hat{y}$ and y are the predicted vector and the ground truth vector respectively.

2.4. Evaluation methods

We use the mean absolute error (MAE), and Root Mean Squared Error (RMSE) as the measurement to evaluate the prediction performance of our proposed iDeepAir. Given a test set $Y = \{y_1, y_2, \cdots, y_n\}$, the MAE and RMSE is as follows:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|\widehat{y_i} - y_i|$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}|\widehat{y_i} - y_i|^2}$$

where $\widehat{Y} = \{\widehat{y_1}, \widehat{y_2}, \cdots, \widehat{y_n}\}$ is the set of the predicted value.

2.5. Layer-wise relevance propagation

Regarding a deep learning neural network, the input vector can be denoted as $V = \{v_1, v_2, \cdots, v_n\}$, and the prediction result of the network is $f(V)$, LRP produces a decomposition $R = \{r_1, r_2, \cdots, r_n\}$ of that prediction on the input variables satisfying:

$$\sum_{i=1}^{n} r_i = f(x)$$

The LRP method is based on a backward propagation mechanism applying uniformly to all neurons in the network. By employing the LRP method on our iDeepAir network, we then obtain the relevance between all 5 pollutants and the output, and denote these as $\{r_0, r_1, r_2, r_3, r_4\}$.

2.6. Calculations of emission inventory

Based on the annual total pollutant emission loads of industrial, residential and transportation sectors, the emission loads of SO_2, NO_2, CO, PM_{10}, and $PM_{2.5}$ from these three sectors can be directly calculated with considering the proportional relationship among sulphur content, nitrogen content, and carbon content within fuel. Further, based on the annual generated electricity in Shanghai, the emission load of power sector can be calculated by:

$$\varepsilon_{energy}^{year} = e_g^{year} * \lambda$$

where $\varepsilon_{energy}^{year}$ corresponds to the emission loads from power sector in a specific year, e_g^{year} indicates the generated electricity in that year, and λ is the emission factor of generated electricity.

2.7. Quantifying contribution of emission sources

Assuming c_i is the contribution of the i_th emission source (Here $i \in \{0,1,2,3\}$), and we have

$$c_i = \frac{\sum_{k=0}^{4} p_i^k r_k}{\sum_{i=0}^{3} \sum_{k=0}^{4} p_i^k r_k}$$

Here p_i^k is the value of k_th pollutant component discharged by the i_th emission source from the emission inventories.

2.8. Prediction of future contribution of traffic emissions

Based on the total numbers of petrol vehicles in Shanghai from 2010 to 2019, we first learn the growth regularity of the total number of urban petrol vehicles with regression analysis, and the regression equation can be written as (Fig.S1):

$$v_{year}^p = 26.569 \times (year - 2009) + 111.81$$

where variable $year$ indicates the year for prediction, and v_{year}^p corresponds to the predicted number of petrol vehicles in the year for prediction. Also, we can predict the number of Electrified vehicles in Shanghai based on regression analysis on the historical numbers of new energy vehicles during 2013-2019, the regression equation is (Fig.S2):

$$v_{year}^n = 0.4634 \times (year - 2012)^2 + 1.162 \times (year - 2012) - 2.0102$$

where variable $year$ also means the year for prediction, and v_{year}^n is the predicted number of new energy vehicles in the year for prediction. By combining the contributed average $PM_{2.5}$ concentration of transportation emissions from 2010 to 2017 with the actual number of petrol vehicles in

Shanghai during the same time, we then learned the regression equation which represents the correlations among these two kinds of dataset (Fig.S3):

$$p_{transportation}^{year} = 0.3054 \times \left(v_{year}^{p}\right)^{0.6885}$$

where $p_{transportation}^{year}$ is the predicted average $PM_{2.5}$ concentration that traffic emission contributes in the year for prediction. And if we replace the variable v_{year}^{p} with v_{year}^{n}, this equation can be easily used to calculate the $PM_{2.5}$ reductions caused by the promotion of new energy vehicles in a corresponding year.

3. Results and discussion

3.1. Evaluations of iDeepAir on $PM_{2.5}$ predictions

In this subsection, the performance of iDeepAir on simulating the dynamic spatiotemporal generation and evolution processes of urban air pollution can be evaluated by measuring its accuracy on future $PM_{2.5}$ concentration prediction (Fig. 3a). Besides, we compared our model with several alternative neural networks for sequence forecasting including ARIMA (Autoregressive integrated moving average (Box and Pierce, 1970)), GBDT (Gradient Boosting Regression Tree(Friedman, 2001)) and emerging deep learning based methods including LSTM (Long-short-term-memory network(Hochreiter and Schmidhuber, 1997)), GRU (Gated Recurrent Unit(Cho *et al.*, 2014)), Seq2seq (Sequence to sequence(Sutskever *et al.*, 2014)), DA-RNN (Dual-stage Attention-based Recurrent Neural network(Yao *et al.*, 2017)), ADAIN (Neural Attention Model for Urban Air Quality Inference (Cheng *et al.*, 2018)), Geo-MAN (Multi-level-attention-based RNN Model for Time Series Prediction(Liang *et al.*, 2018)). These methods use the same dataset, but the input data could be adjusted for different models (Text S2).

The results of models for $PM_{2.5}$ concentration prediction in different time periods (+6h, +12h, +24h, +48h) are presented in Table 2. The best results are marked in bold. Obviously, the results demonstrate that our iDeepAir outperforms other alternative solutions in each prediction task, particularly in short-term predictions, iDeepAir has a great prediction effect with the least MAR and RMSE. Totally, these models tend to have greater inaccuracies in long-term prediction (+24h, +48h) but iDeepAir has a better prediction accuracy compared with other baselines. Specifically, compared to the baseline of ARIMA, iDeepAir can reduce the MAE from 30.252 μg/m^3 to 16.961μg/m^3 (Table 2 and Fig. 3b). Besides, we further discover that those hierarchical structured networks such as seq2seq, DA-RNN, ADAIN, and Geo-MAN, can significantly surpass those non-hierarchical structured networks, and this explains the superiority of our hierarchical iDeepAir which is carefully designed based on the prior knowledge of the dynamic formation process of $PM_{2.5}$ (firstly the vehicle exhaust is discharged into the atmosphere to affect the pollutant concentration data of the atmosphere, and then the generation of $PM_{2.5}$ is enhanced or offset under different meteorological conditions). To verify the effectiveness of each module and the robustness of the model, we conduct a series of ablation studies by removing and replacing each module of the integrated iDeepAir framework. It can be observed that existed modules can effectively improve the performance of the integrated model independently. These experiments verify the predictive ability and robustness of the algorithm under different time lengths, and reflects the practicality and stability (Fig. 3c).

In terms of verifying the accuracy of future $PM_{2.5}$ concentration prediction, iDeepAir has a distinct advantage with high fidelity. In addition, an important output of iDeepAir is a ranking of the contributions of single parameter to the prediction in LRP method (take the 24-hour prediction results as an example) (Bach *et al.*, 2015b) (Fig. 3d). For $PM_{2.5}$, the five major governing factors are air quality, wind direction, weather conditions, pressure and wind speed which are mainly some meteorological factors. Secondary pollutants like O_3, and NO_2 and traffic emission are also important precursors of $PM_{2.5}$ (Wang *et al.*, 2015b) which are also important influence factors. Some open source such as urban dust, soil dust, and cement dust could directly affect the air quality, and then they remain in the air or are not completely removed, which would still promote or inhibit the formation of $PM_{2.5}$ in the future (L *et al.*, 2014). Besides, the formation of $PM_{2.5}$ is correlated to wind direction and wind speed due to the dilution or aerial migration effect and slightly decreased as pressure increases because of the change of air pressure will lead to the movement of air flow, which would contribute to the diffusion of pollutants (Li *et al.*, 2017b; Liu *et al.*, 2020). As the variation of meteorological conditions, secondary pollutants or precursor gases (e.g. sulfate, nitrate, ammonium and carbonaceous matters) are essential to participate in the generation of $PM_{2.5}$ through photo-chemical reactions forming ozone and biogenic VOC (volatile organic compound) to cause the severe $PM_{2.5}$ pollution in Shanghai megacity (Wang, Qiao, Lou, Zhou, Chen, Wang, Tao, Chen, Huang, Li and Huang, 2015b). Furthermore, weather condition is a key factor by changing solar irradiance that is a limiting factor that influences ozone-related photochemistry (Parker *et al.*, 2020; Pusede *et al.*, 2014; Yang *et al.*, 2021b). Such a ranking of influencing factors of PM2.5 formation is comparatively consistent with current research (Su *et al.*, 2020; Yang, Wen, Wang, Zhang, Pinto, Pennington, Wang, Wu, Sander, Jiang, Hao, Yung and Seinfeld, 2021a).

3.2. Spatial contribution from crucial domain on air quality

Combined with the spatial location of Shanghai, we analyzed the correlation between traffic flow and $PM_{2.5}$ in the mesh division of 16 × 16. With the embedded Layer-wise Relevance Propagation (LRP) algorithm, the spatial traffic emissions are located and tracked clearly. The overall spatial patterns of urban traffic follow a core-peripheral distribution, and Huangpu district, which lies in the core area of Shanghai, is the most congested district in the city. In addition, there are two minor cores in Baoshan District and Pudong New District respectively (Fig. 4a). Regarding Huangpu district which is highly-developed, the more traffic emissions due to its developed economy and the less diffusion caused by its internal skyscrapers will definitely lead to more serious air pollution.

The spatial correlation between urban traffic and $PM_{2.5}$ concentration is shown with a citywide heat-map in Fig. 4b. As observed, there exist strong positive spatial correlations within traffic patterns and $PM_{2.5}$ concentration (Circled blue and marked **A**). And most significant correlations are focused on the west side of Huangpu River since this area has more focused traffic flows. Besides, in the central area of the city, there exist two interesting sub-regions (Circled green and marked B) where the traffic flow is relatively high while the correlations are inconspicuous. Through practical field investigations, we discover most of these sub-regions are parklands and the heights of most buildings in these two sub-regions are relatively low, and infer that this kind of specific land properties could effectively accelerate the diffusion of air pollution, hence reduce the impacts of traffic emissions on pollution. Furthermore, there exist some particular sections (Circled black and marked C) with conflicting heavy traffic and scarce correlations. Considering the industrial distribution of Shanghai, we believe the correlations between urban traffic and $PM_{2.5}$ concentration have been masked by the dominant industrial emissions (Wang *et al.*, 2014a).

3.3. Calculate Emission inventory from 2010 to 2017

Based on the released emission loads of different pollutants from anthropogenic sources and the generated electricity in Shanghai from 2010 to 2017, we calculate and reconstruct the emission inventories for Shanghai for each year from 2010 to 2017 (Fig. S4). During these years, the industrial emissions were always the source with the highest emission load, and the transportation sector was the second-highest emission source. Moreover, the emission loads of the industrial and power sectors decreased significantly with the strictest industrial emission limitation measurements ever such as simultaneous desulfurization and denitrification since 2013 as the consequence of clean air actions (Tang *et al.*, 2019; Zhang *et al.*, 2019). However, even though strict license plate limitation policy and vehicle emission standard had been employed during these years, the emission load of transportation sector decreased laxly and limitedly since the absolute total number of vehicles increased with great rapidity. Based on the calculated emission inventories, we also discovered that the total emission loads of the key pollutants of CO, NO_2, and SO_2 have been reduced by 59.81%, 72.03%, and 60.93% respectively, and the air pollution limitation measurements that Shanghai had employed were quite effective. Given the unsatisfactory transportation emission reductions, we should increase our efforts in this regard (Fig. 5a).

3.4. Quantified contribution of anthropogenic emissions to $PM_{2.5}$ formation

Based on the surveilled air quality data of Shanghai, we decompose the final prediction results in terms of the input features with the embedded LRP algorithm of iDeepAir, and cooperating this learned relevance with the calculated emission inventories, the contribution of anthropogenic emissions to $PM_{2.5}$ formation could be directly and separately calculated (Fig. 5b). As demonstrated, the emission contribution of the industrial, residential and power sectors to $PM_{2.5}$ formation decreased steadily and continuously, while the contribution of traffic emissions increased independently and smoothly (Fig. 5a). We believe that the effectiveness of the strictest industrial and power emission limitation measurements has been verified. Given the fact that industry and power will be stick to these limitation measurements chronically, we should definitely take the issue of reducing traffic emissions as the priority in subsequent efforts on sustainably improving urban air quality.

To verify the correctness of our quantified contribution, we compared our results with the official announced results. From an analysis report of $PM_{2.5}$ source which was publicly released by Shanghai Municipal Bureau of Ecology and Environment, the traffic emissions accounted for 22.6%~33.6% of the total impacts of all local emissions on $PM_{2.5}$ formation in Shanghai, this report is generated with some previous mentioned methods(Wang *et al.*, 2014b; Wang *et al.*, 2006; Yao *et al.*, 2002) which distinguish the impacts of different emissions. In our calculation, transportation emissions accounted for 22.69% and 24.35% of the total impacts of all local emissions on $PM_{2.5}$ formation (Fig. 5a) respectively in 2012 and 2013, and this is quite consistent with the official released report.

3.5. Prediction of contribution of transportation emissions in future and Air quality benefit for future new energy vehicle promotion

By learning the current growth rates of the total petrol and new energy vehicle numbers in Shanghai from historical dataset, the current vehicle license plate limitation and new energy promotion policies, we first estimate the possible total numbers of both petrol and new energy vehicles in future (Fig. 6a). As predicted, the numbers of both traditional petrol and new energy vehicles increase significantly during the next ten years, and even though the absolute number of new energy vehicles may be up to 1.69 million, they may only account for a small part of 21.63% of total urban vehicles in 2030. Meanwhile, we use a power function to approximate the relationships between with the actual numbers of petrol vehicles and the real contributed $PM_{2.5}$ concentration of transportation emissions during 2010-2017, and use this approximated function to predict the possible contributed $PM_{2.5}$ concentration of urban traffic in the following ten years. Afterward, to be fair, we assume that no further limitation measurements will be employed on reducing industrial and power emissions and the contributed concentration of these two kinds of emissions are fixed for the next ten years, and the contribution of urban petrol vehicles to $PM_{2.5}$ formation are calculated for the next ten years (Fig. 6b). In 2030, the emissions from urban petrol vehicles may account for more than 50% of the total impacts of all local emissions on $PM_{2.5}$ formation, and petrol vehicle emissions will be the primary source of urban $PM_{2.5}$, and our initial conjecture that the issue of reducing traffic emissions should be considered as the priority in improving urban air quality is verified positively. Fortunately, we are delighted to discover that the promotion of new energy vehicles may bring us an obvious reduction on $PM_{2.5}$ formation in 2030, considering the current fleet electrification promotion policy and the possible contribution of petrol vehicle emissions to $PM_{2.5}$, which is a very valuable solution.

3.6. Effectiveness evaluation of new energy vehicle promotion policies

Given the effectiveness of the promotion of new energy vehicles on $PM_{2.5}$ reductions and the giant prospective total number of petrol vehicles in 2030, it is a realistic way to alleviate urban air pollution by improving the promotion of new energy vehicles. We here evaluate the effectiveness of improving the promotion of new energy vehicles on $PM_{2.5}$ reductions (Fig. 6c). As shown, if 50% of all petrol vehicles are replaced by new energy vehicles in 2030, the contribution of transportation emissions to $PM_{2.5}$ can be reduced by 11.72%, and the absolute value that traffic emissions contribute to $PM_{2.5}$ can be reduced from 25.33μg/m^3 to 15.72μg/m^3 (Fig. 6d). And the reductions of transportation contribution on $PM_{2.5}$ and the absolute traffic contributed $PM_{2.5}$ concentration can reach 25.24% and 8.36μg/m^3 respectively in case that 80% of all petrol vehicles are replaced by new energy vehicles. Based on this analysis, we suggest that more forceful policies on enhancing the promotion of new energy vehicles should be considered for greener and sustainable future developments of modern cities.

4. Conclusions

In this paper, we propose a novel method to quantify the influence of anthropogenic emissions on $PM_{2.5}$ concentration with a novel and traceable deep learning model. Our experiment results indicate that the proposed model could achieve better fitting and prediction performances than the LSTM, GBRT, Seq2Seq, DA-RNN model and other deep learning models. Furthermore, we output the contributions of input parameter to the prediction in LRP method and visualize and analyze the spatial correlation between traffic flow and $PM_{2.5}$.

In addition, we discover that transportation emissions will play the most dominate role in future urban air pollution, and how to reduce traffic emissions are an unavoidable issue on achieving sustainable developments in modern cities. Meanwhile, to some extent, new energy vehicles can be considered as an effective way to reduce traffic emissions. However, the current promotion policies and efforts are far from enough. To further improve urban air quality, we need some more effective and powerful measurements in response to the rapid growth of urban traffic emissions.

Due to its powerful ability on simulating the complex chemical reaction and temporal accumulation procedure of $PM_{2.5}$ formation by integrating multi-source data, the proposed deep learning network may be easily applied in other metropolises to address the challenging pollution tasks. With the traceable deep learning network, we can quantify the impacts of different anthropogenic emissions on different urban air pollutants by computing jointly with local emission inventories. To support more efficient and sustainable city planning, it is very important to deeply and sufficiently understand the impacts of different human activities on air pollution. The insights and observations obtained in this paper are of great significance to provide the qualitative and quantitative decision-making basis for citywide traffic management and urban pollution control, and may eventually benefit people's lives and high-quality sustainable developments of cities.

Last but not least, to fully address the challenge of urban air pollution, multiple ingredients such as urban layout, industrial planning and population distribution should be considered comprehensively in analyzing and tackling air pollution, and the proposed deep learning network is capable of embedding these heterogeneous features and learning the fine-grained influences of them on urban air pollutions. For future research, if the organics of the emissions of different human activities can be further considered (Guo *et al.*, 2020), the accuracy of quantifying the impacts of different human activities can be subsequently improved.

Acknowledgements

This paper is partially supported by the Anhui Science Foundation for Distinguished Young Scholars (No.1908085J24), Natural Science Foundation of China (No.62072427), Jiangsu Natural Science Foundation (No. BK20191193).

Declaration of Competing Interest

Wenjie Du and Lianliang Chen contributed equally to this manuscript. The authors declared that there is no conflict of interest.

Appendix A. Supplementary material

Supplementary material has been submitted with this article.

Data statement

All datasets supporting this study are available from the corresponding authors on reasonable request.

Table 1. Composition of the dataset and the related statistics.

Data Feature	Variable	Mean	Std	Min	Max	Time interval
Traffic data	TSI	28.39	9.13	5.83	50.64	2 min
Air quality	$P_{2.5}$(μg/m^3)	55.56	42.01	4	356.67	60 min
	PM_{10}(μg/m^3)	78.97	51.64	4.78	386.44	
	O_3/(μg/m^3)	64.33	38.5	5.11	271	
	SO_2/(μg/m^3)	20.14	13.91	6	125.56	
	NO_2/(μg/m^3)	49.19	26.63	3.78	173.89	
	CO/(μg/m^3)	0.89	0.37	0.35	2.96	
Meteorological data	Weather conditions[a]	2.87	4.49	0	15	30 min
	Dew point temperature /°C	9.04	9.48	-17	27	
	Humidity/%	69.56	18.19	14	100	
	Pressure/kPa	1019.23	8.36	994	1037	
	Temperature/°C	15.12	8.27	-3	37	
	Wind direction[b]	3.84	2.59	0	8	
	Wind speed/(m*s^{-1})	13.95	6.53	3.6	43.2	
	Air quality[c]	2.02	1.02	1	6	
Time period	From 2014/8/11 to 2015/4/30					

a. The weather conditions has 16 features (sunny, overcast, sunny to cloudy, fog, sleet, thunder shower, light rain, heavy rain, moderate rain, rainstorm, heavy snow, light snow, moderate snow, rain, hail, cloudy), which are encoded into numerical variable [0,15].

b. The wind direction has 9 directions (no wind, north wind, west wind, east wind, south wind wind, northwest wind, northeast wind, etc.), which are encoded into numerical variable [0,8].

c. The air quality has six levels (I, II, III, IV, V, VI). The higher the index, the more serious the air pollution is, which are encoded into numerical variable [1,6].

Table 2. The result for hourly prediction values of PM2.5 of different models in different time periods.

Method	+6h		+12h		+24h		+48h	
	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE
ARIMA	38.649	28.562	38.634	27.805	41.150	30.252	37.753	27.593
GBDT	26.771	21.573	26.988	21.809	27.148	22.156	28.234	22.764
LSTM	25.830	20.176	27.164	21.721	31.690	23.779	33.495	25.466
GRU	24.321	20.017	26.248	20.963	30.546	22.979	32.879	22.856
Seq2Seq	19.685	14.946	24.966	19.653	28.708	22.816	29.419	23.012
DA-RNN	17.303	14.131	22.769	17.958	27.138	21.223	32.572	24.857
ADAIN	16.254	13.176	24.567	19.468	27.479	21.441	34.073	25.994
Geo-MAN	21.744	17.795	21.632	18.046	27.792	22.617	32.506	24.713
iDeepAir	**5.849**	**11.825**	**20.285**	**15.890**	**23.233**	**16.961**	**24.595**	**19.891**

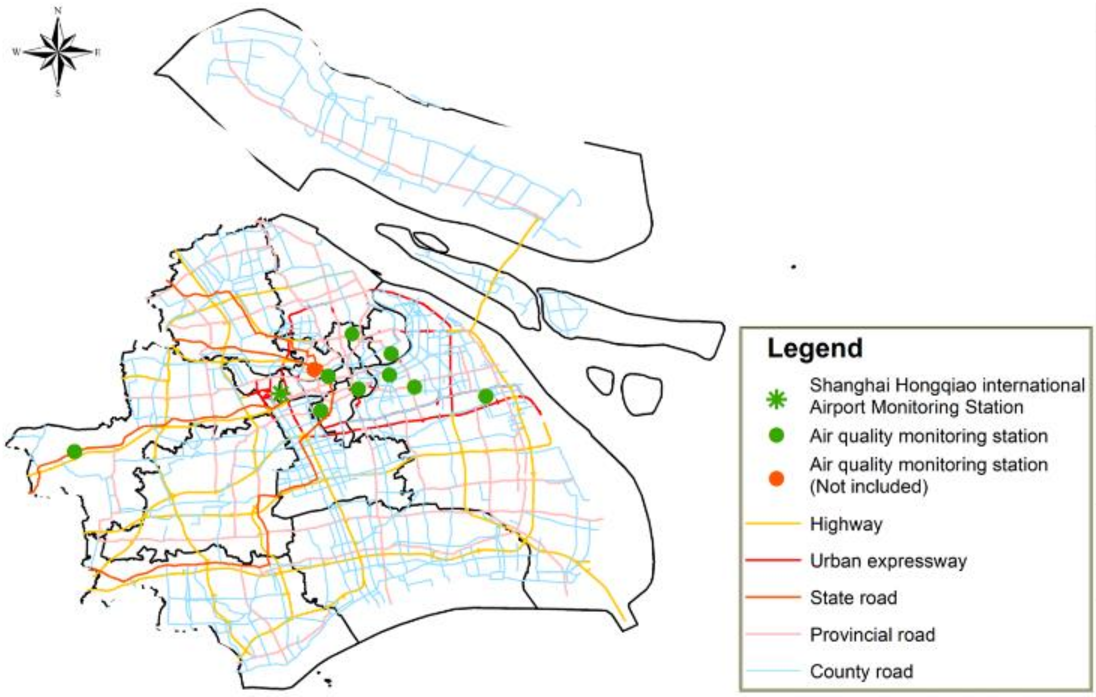

Figure 1. The spatial distribution of monitoring stations and road traffic network in Shanghai.

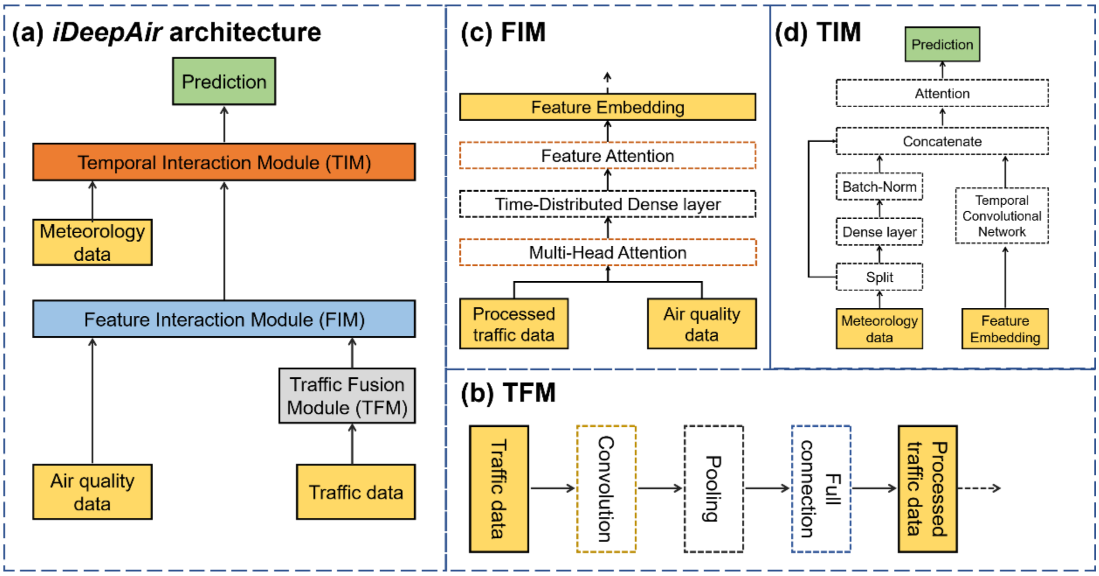

Figure 2. Overview of iDeepAir architecture and three important modules. **(a)** Overall hierarchical structure of iDeepAir, **(b)** Traffic Fusion Module (TFM), (**c**) Feature Interaction Module (FIM), and (**d**) Time Interaction Module (TIM).

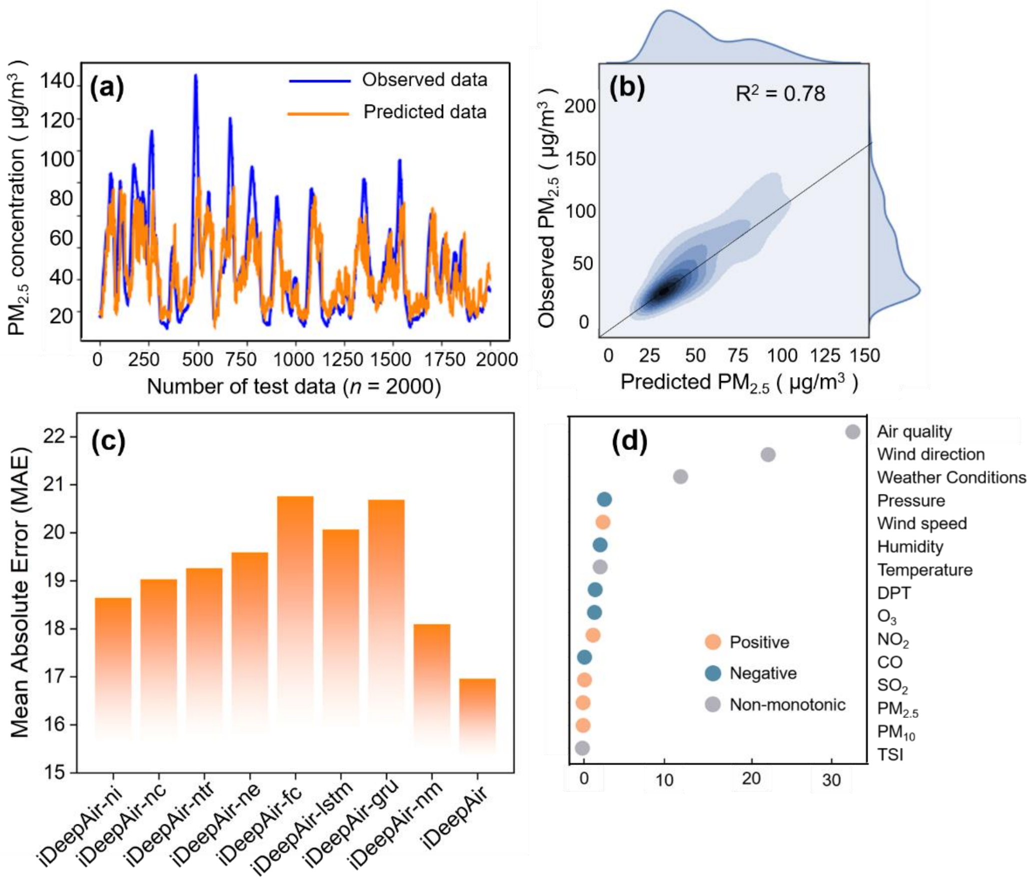

Figure 3. Model performance, ablative evaluations and variable contribution for $PM_{2.5}$ prediction. (**a**) and (**b**) The iDeepAir Model performance for $PM_{2.5}$ prediction. (**c**) Ablative evaluations of iDeepAir. (**d**) The Contribution assessment of variables (DPT: Dew point temperature).

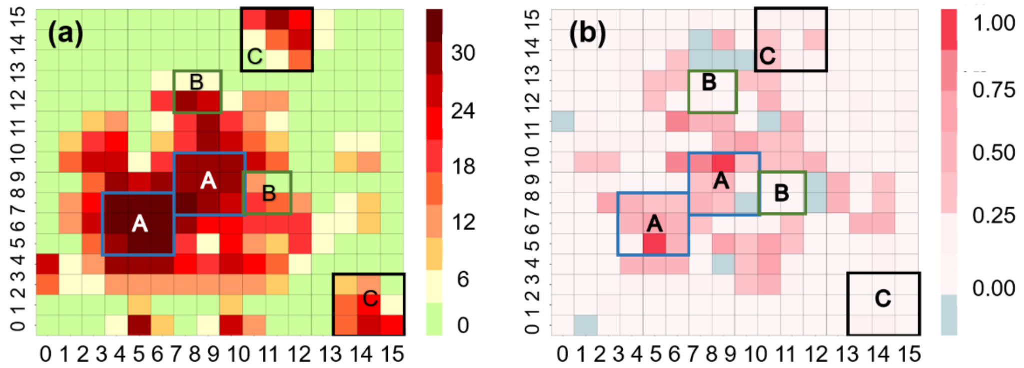

Figure 4. Spatial distributions of average TSIs and correlations between traffic and $PM_{2.5}$ in Shanghai. **(a)** Spatial distribution of average TSIs in Shanghai. **(b)** Spatial distribution of correlations between urban traffic and $PM_{2.5}$ Concentration.

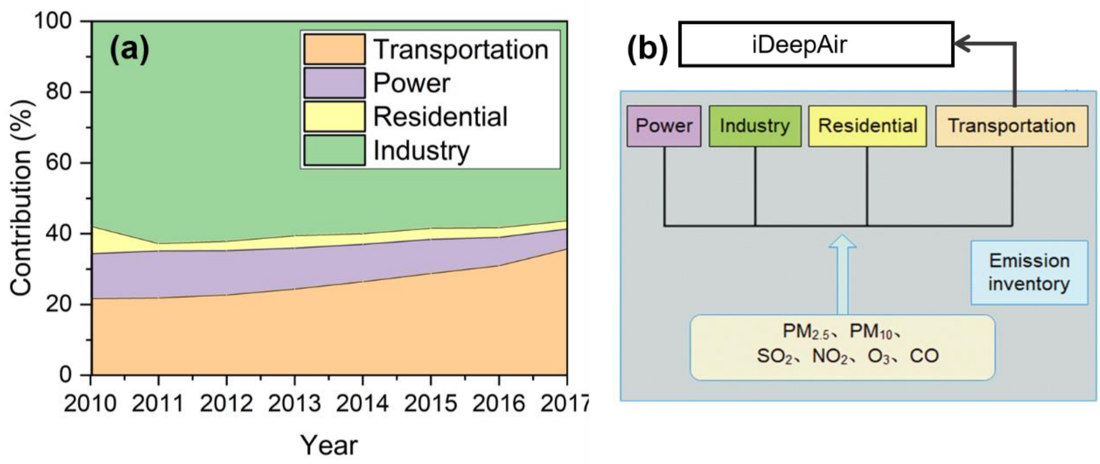

Figure 5. The contribution of anthropogenic emissions to $PM_{2.5}$ formation in Shanghai. **(a)** $PM_{2.5}$ emission contribution rate of each industry (Transportation, Power, Residential and Industry). **(b)** the overall framework of emission inventory and iDeepAir model.

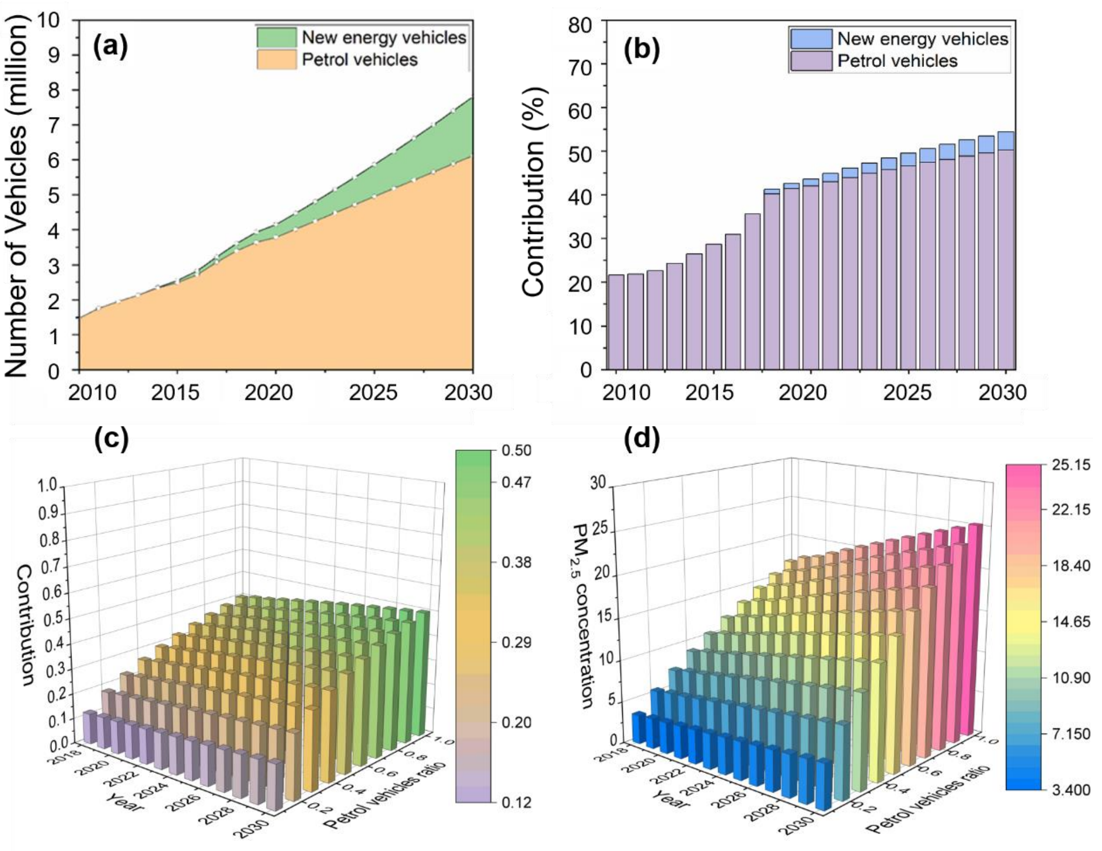

Figure 6. Predictions of the total numbers of urban vehicles and evaluations of new energy vehicle promotion from 2020 to 2030 in Shanghai. (a) Predictions of the total numbers of both petrol and new energy vehicles from 2020 to 2030. **(b)** Predictions of the contribution of transportation emissions to $PM_{2.5}$. **(c)** Contribution of transportation emissions to $PM_{2.5}$ with different percentages of petrol vehicles replaced by new energy vehicles. **(d)** $PM_{2.5}$ concentration that transportation emissions contributed with different percentages of petrol vehicles replaced by new energy vehicles.

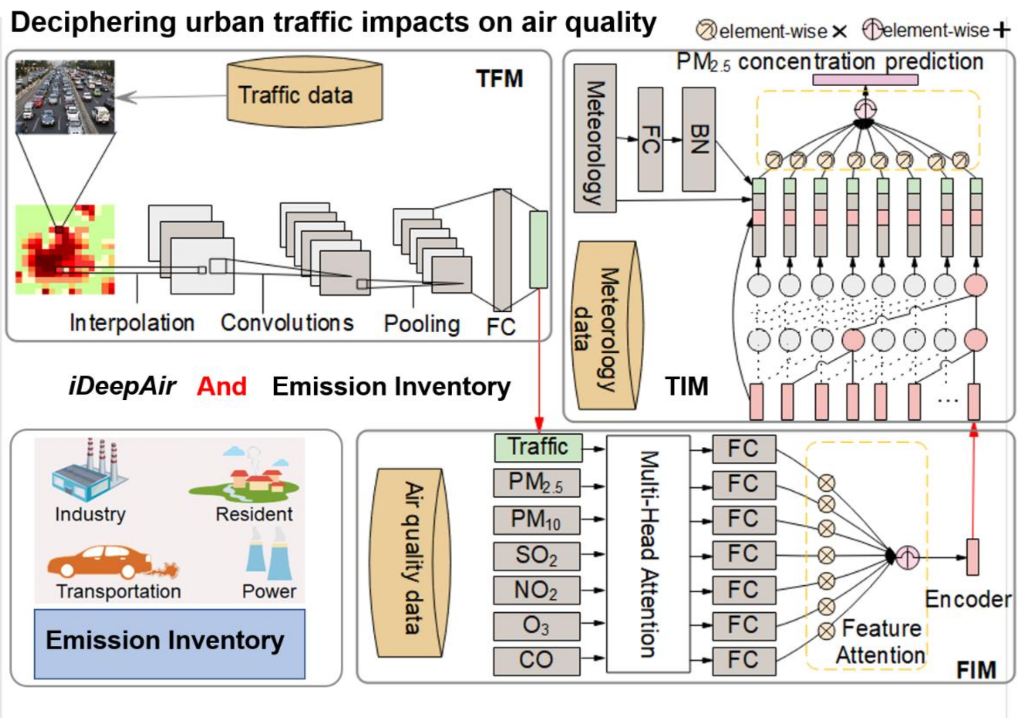

Table of Contents (TOC) graphic

参考文献

Alfaseeh, L, Tu, R, Farooq, B, Hatzopoulou, M, 2020. Greenhouse gas emission prediction on road network using deep sequence learning. *Transport Res D-Tr E* 88(102593).

Anonymous, 2019. Deep learning monitors human activity based on sound alone. *Nature* 570(7759).

Arras, L, Horn, F, Montavon, G, Mueller, K, Samek, W, 2017a. "What is relevant in a text document?": An interpretable machine learning approach. *Plos One* 12(e01811428).

Arras, L, Montavon, G, Muller, K R, Samek, W, 2017b. Explaining Recurrent Neural Network Predictions in Sentiment Analysis [arXiv]. *arXiv*:9.

Bach, S, Binder, A, Montavon, G, Klauschen, F, Mueller, K, Samek, W, 2015a. On Pixel-Wise Explanations for Non-Linear Classifier Decisions by Layer-Wise Relevance Propagation. *Plos One* 10(e01301407).

Bach, S, Binder, A, Montavon, G, Klauschen, F, Müller, K, Samek, W, 2015b. On Pixel-Wise Explanations for Non-Linear Classifier Decisions by Layer-Wise Relevance Propagation. *Plos One* 10(7):e130140.

Benas, N, Beloconi, A, Chrysoulakis, N, 2013. Estimation of urban PM10 concentration, based on MODIS and MERIS/AATSR synergistic observations. *Atmos Environ* 79:448-454.

Box, G E P, Pierce, D A, 1970. Distribution of Residual Autocorrelations in Autoregressive-Integrated Moving Average Time Series Models. *Publications of the American Statistical Association* 65(332):1509-1526.

Cheng, W, Shen, Y, Zhu, Y, Huang, L, 2018. A Neural Attention Model for Urban Air Quality Inference: Learning the Weights of Monitoring Stations, pp. 2151-2158.

Cho, K, van Merrienboer, B, Gulcehre, C, Bahdanau, D, Bougares, F, Schwenk, H, Bengio, Y, 2014. Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation. *Computer Science*.

Daskalakis, N, Tsigaridis, K, Myriokefalitakis, S, Fanourgakis, G S, Kanakidou, M, 2016. Large gain in air quality compared to an alternative anthropogenic emissions scenario. *Atmos Chem Phys* 16(15):9771-9784.

Ding, Y, Liu, Y, Luan, H, Sun, M, 2017. Visualizing and Understanding Neural Machine Translation, in Barzilay R, Kan M Y (Eds.), pp. 1150-1159.

Dumka, U C, Kaskaoutis, D G, Srivastava, M K, Devara, P C S, 2015. Scattering and absorption properties of near-surface aerosol over Gangetic-Himalayan region: the role of boundary-layer dynamics and long-range transport. *Atmos Chem Phys* 15(3):1555-1572.

Friedman, J H, 2001. Greedy function approximation: A gradient boosting machine. *Ann Stat* 29(5):1189-1232.

Gao, M, Han, Z, Liu, Z, Li, M, Xin, J, Tao, Z, Li, J, Kang, J, Huang, K, Dong, X, Zhuang, B, Li, S, Ge, B, Wu, Q, Cheng, Y, Wang, Y, Lee, H, Kim, C, Fu, J S, Wang, T, Chin, M, Woo, J, Zhang, Q, Wang, Z, Carmichael, G R, 2018. Air quality and climate change, Topic 3 of the Model Inter-Comparison Study for Asia Phase III (MICS-Asia III) - Part 1: Overview and model evaluation. *Atmos Chem Phys* 18(7):4859-4884.

Guo, S, Hu, M, Peng, J, Wu, Z, Zamora, M L, Shang, D, Du, Z, Zheng, J, Fang, X, Tang, R, Wu, Y, Zeng, L, Shuai, S, Zhang, W, Wang, Y, Ji, Y, Li, Y, Zhang, A L, Wang, W, Zhang, F, Zhao, J, Gong, X, Wang, C, Molina, M J, Zhang, R, 2020. Remarkable nucleation and growth of ultrafine particles from vehicular exhaust. *P Natl Acad Sci Usa*.

Ham, Y, Kim, J, Luo, J, 2019. Deep learning for multi-year ENSO forecasts. *Nature* 573(7775):568.

Hino, M, Benami, E, Brooks, N, 2018. Machine learning for environmental monitoring. *NATURE SUSTAINABILITY* 1(10):583-588.

Hochreiter, S, Schmidhuber, J, 1997. Long short-term memory. *Neural Comput* 9(8):1735-1780.

Huang, R, Zhang, Y, Bozzetti, C, Ho, K, Cao, J, Han, Y, Daellenbach, K R, Slowik, J G, Platt, S M, Canonaco, F, Zotter, P, Wolf, R, Pieber, S M, Bruns, E A, Crippa, M, Ciarelli, G, Piazzalunga, A, Schwikowski, M, Abbaszade, G, Schnelle-Kreis, J, Zimmermann, R, An, Z, Szidat, S, Baltensperger, U, El Haddad, I, Prevot, A S H, 2014. High secondary aerosol contribution to particulate pollution during haze events in China. *Nature* 514(7521):218-222.

Julie, Y Z, Chao, Z, Shi, Z, Li, V O K, Jiawei, H, Yu, Z, 2016. p-causality: identifying spatiotemporal causal pathways for air pollutants with urban Big Data [arXiv]. *arXiv*:10.

Kroll, J H, Heald, C L, Cappa, C D, Farmer, D K, Fry, J L, Murphy, J G, Steiner, A L, 2020. The complex chemical effects of COVID-19 shutdowns on air quality. *Nat Chem* 12(9):777-779.

Kumar, R, Peuch, V, Crawford, J H, Brasseur, G, 2018. Five steps to improve air-quality forecasts. *Nature* 561(7721):27-29.

L, R, Z, Z, X, Z, W, Y, B, Y, Z, B, Y, J, 2014. Source Apportionment of PM10 and PM2.5 in Urban Areas of Chongqing. *Research of Environmental Sciences* 27(12):1387-1394.

Lapuschkin, S, Waeldchen, S, Binder, A, Montavon, G, Samek, W, Mueller, K, 2019. Unmasking Clever Hans predictors and assessing what machines really learn. *Nat Commun* 10(1096).

Le, T, Wang, Y, Liu, L, Yang, J, Yung, Y L, Li, G, Seinfeld, J H, 2020. Unexpected air pollution with marked emission reductions during the COVID-19 outbreak in China. *Science (American Association for the Advancement of Science)* 369(6504):702-706.

Lelieveld, J, Evans, J S, Fnais, M, Giannadaki, D, Pozzer, A, 2015. The contribution of outdoor air pollution sources to premature mortality on a global scale. *Nature* 525(7569):367.

Li, M, Liu, H, Geng, G, Hong, C, Liu, F, Song, Y, Tong, D, Zheng, B, Cui, H, Man, H, Zhang, Q, He, K, 2017a. Anthropogenic emission inventories in China: a review. *Natl Sci Rev* 4(6):834-866.

Li, T, Guo, Y, Liu, Y, Wang, J, Wang, Q, Sun, Z, He, M Z, Shi, X, 2019a. Estimating mortality burden attributable to short-term PM2.5 exposure: A national observational study in China. *Environ Int* 125:245-251.

Li, X, Feng, Y J, Liang, H Y, 2017b. The Impact of Meteorological Factors on PM2.5 Variations in Hong Kong. *IOP Conference Series: Earth and Environmental Science* 78:1-10.

Li, X, Wang, H, He, H, Du, J, Chen, J, Wu, J, 2019b. Intelligent diagnosis with Chinese electronic medical records based on convolutional neural networks. *Bmc Bioinformatics* 20(62).

Liang, Y, Ke, S, Zhang, J, Yi, X, Zheng, Y, 2018. GeoMAN: Multi-level Attention Networks for Geo-sensory Time Series Prediction.

Lindsey, R, Daluiski, A, Chopra, S, Lachapelle, A, Mozer, M, Sicular, S, Hanel, D, Gardner, M, Gupta, A, Hotchkiss, R, Potter, H, 2018. Deep neural network improves fracture detection by clinicians. *P Natl Acad Sci Usa* 115(45):11591-11596.

Liu, A, Qian, N, Yu, H, Chen, R, Kan, H, 2017. Estimation of disease burdens on preterm births and low birth weights attributable to maternal fine particulate matter exposure in Shanghai, China. *Sci Total Environ* 609:815-821.

Liu, T, Wang, X, Hu, J, Wang, Q, An, J, Gong, K, Sun, J, Li, L, Qin, M, Li, J, Tian, J, Huang, Y, Liao, H, Zhou, M, Hu, Q, Yan, R, Wang, H, Huang, C, 2020. Driving Forces of Changes in Air Quality during the COVID-19 Lockdown Period in the Yangtze River Delta Region, China. *Environ. Sci. Technol. Lett.* 7(11):779-786.

Park, Y, Kwon, B, Heo, J, Hu, X, Liu, Y, Moon, T, 2020. Estimating PM2.5 concentration of the conterminous United States via interpretable convolutional neural networks. *Environ Pollut* 256:113395.

Parker, H A, Hasheminassab, S, Crounse, J D, Roehl, C M, Wennberg, P O, 2020. Impacts of Traffic Reductions Associated With COVID‐19 on Southern California Air Quality. *Geophys Res Lett* 47(23).

Pusede, S E, Gentner, D R, Wooldridge, P J, Browne, E C, Rollins, A W, Min, K E, Russell, A R, Thomas, J, Zhang, L, Brune, W H, Henry, S B, DiGangi, J P, Keutsch, F N, Harrold, S A, Thornton, J A, Beaver, M R, St. Clair, J M, Wennberg, P O, Sanders, J, Ren, X, VandenBoer, T C, Markovic, M Z, Guha, A, Weber, R, Goldstein, A H, Cohen, R C, 2014. On the temperature dependence of organic reactivity, nitrogen oxides, ozone production, and the impact of emission controls in San Joaquin Valley, California. *Atmos Chem Phys* 14(7):3373-3395.

Souri, A H, Choi, Y, Jeon, W, Li, X, Pan, S, Diao, L, Westenbarger, D A, 2016. Constraining NOx emissions using satellite NO2 measurements during 2013 DISCOVER-AQ Texas campaign. *Atmos Environ* 131:371-381.

Su, T, Li, Z, Zheng, Y, Luan, Q, Guo, J, 2020. Abnormally Shallow Boundary Layer Associated With Severe Air Pollution During the COVID‐19 Lockdown in China. *Geophys Res Lett* 47(20).

Sutskever, I, Vinyals, O, Le, Q V, 2014. Sequence to Sequence Learning with Neural Networks, in Ghahramani Z, Welling M, Cortes C, Lawrence N D, Weinberger K Q (Eds.), Advances in Neural Information Processing Systems.

Tang, L, Qu, J, Mi, Z, Bo, X, Chang, X, Anadon, L, Wang, S, Xue, X, Li, S, Wang, X, Zhao, X, 2019. Substantial emission reductions from Chinese power plants after the introduction of ultra-low emissions standards. *Nat Energy* 4:1-10.

Tichenor, M, Sridhar, D, 2019. Metric partnerships: global burden of disease estimates within the World Bank, the World Health Organisation and the Institute for Health Metrics and Evaluation. *Wellcome open research* 4:35.

Vaswani, A, Shazeer, N, Parmar, N, Uszkoreit, J, Jones, L, Gomez, A N, Kaiser, L, Polosukhin, I, 2017. Attention Is All You Need, in Guyon I, Luxburg U V, Bengio S, Wallach H, Fergus R, Vishwanathan S, Garnett R (Eds.), Advances in Neural Information Processing Systems.

Wang, A, Xu, J, Tu, R, Saleh, M, Hatzopoulou, M, 2020. Potential of machine learning for prediction of traffic related air pollution. *Transport Res D-Tr E* 88(102599).

Wang, H L, Qiao, L P, Lou, S R, Zhou, M, Chen, J M, Wang, Q, Tao, S K, Chen, C H, Huang, H Y, Li, L, Huang, C, 2015a. PM 2.5 pollution episode and its contributors from 2011 to 2013 in urban Shanghai, China. *Atmos Environ* 123:298-305.

Wang, H L, Qiao, L P, Lou, S R, Zhou, M, Chen, J M, Wang, Q, Tao, S K, Chen, C H, Huang, H Y, Li, L, Huang, C, 2015b. PM 2.5 pollution episode and its contributors from 2011 to 2013 in urban Shanghai, China. *Atmos Environ* 123:298-305.

Wang, Y, Li, L, Chen, C, Huang, C, Huang, H, Feng, J, Wang, S, Wang, H, Zhang, G, Zhou, M, Cheng, P, Wu, M, Sheng, G, Fu, J, Hu, Y, Russell, A G, Wumaer, A, 2014a. Source apportionment of fine particulate matter during autumn haze episodes in Shanghai, China. *J. Geophys. Res. Atmos.* 119(4):1903-1914.

Wang, Y, Li, L, Chen, C, Huang, C, Huang, H, Feng, J, Wang, S, Wang, H, Zhang, G, Zhou, M, Cheng, P, Wu, M, Sheng, G, Fu, J, Hu, Y, Russell, A G, Wumaer, A, 2014b. Source apportionment of fine particulate matter during autumn haze episodes in Shanghai, China. *J Geophys Res-Atmos* 119(4):1903-1914.

Wang, Y, Zhuang, G, Zhang, X, Huang, K, Xua, C, Tang, A, Chen, J, An, Z, 2006. The ion chemistry, seasonal cycle, and sources of PM2.5 and TSP aerosol in Shanghai. *Atmos Environ* 40(16):2935-2952.

Wu, J, Bei, N, Hu, B, Liu, S, Wang, Y, Shen, Z, Li, X, Liu, L, Wang, R, Liu, Z, Cao, J, Tie, X, Molina, L T, Li, G, 2020. Aerosol – photolysis interaction reduces particulate matter during wintertime haze events. *Proc Natl Acad Sci USA* 117(18):9755-9761.

Xiao, Z, Zhang, Y, Hong, S, Bi, X, Jiao, L, Feng, Y, Wang, Y, 2011. Estimation of the Main Factors Influencing Haze, Based on a Long-term Monitoring Campaign in Hangzhou, China. *Aerosol Air Qual Res* 11(7):873-882.

Xing, J, Zheng, S, Ding, D, Kelly, J T, Wang, S, Li, S, Qin, T, Ma, M, Dong, Z, Jang, C, Zhu, Y, Zheng, H, Ren, L, Liu, T, Hao, J, 2020. Deep Learning for Prediction of the Air Quality Response to Emission Changes. *Environ Sci Technol* 54(14):8589-8600.

Yan, L, Luo, X, Zhu, R, Santi, P, Wang, H, Wang, D, Zhang, S, Ratti, C, 2020. Quantifying and analyzing traffic emission reductions from ridesharing: A case study of Shanghai. *Transport Res D-Tr E* 89(102629).

Yang, J, Wen, Y, Wang, Y, Zhang, S, Pinto, J P, Pennington, E A, Wang, Z, Wu, Y, Sander, S P, Jiang, J H, Hao, J, Yung, Y L, Seinfeld, J H, 2021a. From COVID-19 to future electrification: Assessing traffic impacts on air quality by a machine-learning model. *Proc Natl Acad Sci USA* 118(26):e2102705118.

Yang, K, Kong, L, Tong, S, Shen, J, Chen, L, Jin, S, Wang, C, Sha, F, Wang, L, 2021b. Double High-Level Ozone and PM2.5 Co-Pollution Episodes in Shanghai, China: Pollution Characteristics and Significant Role of Daytime HONO. *Atmosphere-Basel* 12(5):557.

Yao, Q, Dongjin, S, Haifeng, C, Wei, C, Guofei, J, Cottrell, G, 2017. A Dual-Stage Attention-Based Recurrent Neural Network for Time Series Prediction [arXiv]. *arXiv*:7.

Yao, X H, Chan, C K, Fang, M, Cadle, S, Chan, T, Mulawa, P, He, K B, Ye, B M, 2002. The water-soluble ionic composition of PM2.5 in Shanghai and Beijing, China. *Atmos Environ* 36(PII S1352-2310(02)00342-426):4223-4234.

Yuan, Y, Bar-Joseph, Z, 2019. Deep learning for inferring gene relationships from single-cell expression data. *P Natl Acad Sci Usa* 116(52):27151-27158.

Zhang, K, Batterman, S, 2010. Near-road air pollutant concentrations of CO and PM2.5: A comparison of MOBILE6.2/CALINE4 and generalized additive models. *Atmos Environ* 44(14):1740-1748.

Zhang, Q, Fu, F, Tian, R, 2020. A deep learning and image-based model for air quality estimation. *Sci Total Environ* 724.

Zhang, Q, He, K, Huo, H, 2012. Cleaning China's air. *Nature* 484(7393):161-162.

Zhang, Q, Zheng, Y, Tong, D, Shao, M, Wang, S, Zhang, Y, Xu, X, Wang, J, He, H, Liu, W, Ding, Y, Lei, Y, Li, J, Wang, Z, Zhang, X, Wang, Y, Cheng, J, Liu, Y, Shi, Q, Yan, L, Geng, G, Hong, C, Li, M, Liu, F, Zheng, B, Cao, J, Ding, A, Gao, J, Fu, Q, Huo, J, Liu, B, Liu, Z, Yang, F, He, K, Hao, J, 2019. Drivers of improved PM2.5 air quality in China from 2013 to 2017. *P Natl Acad Sci Usa* 116(49):24463-24469.

Zhang, X, Gao, H, Zhao, J, Zhou, M, 2018. 基于深度学习的自动驾驶技术综述. *清华大学学报. 自然科学版* 58(1000-0054(2018)58:4<438:JYSDXX>2.0.TX;2-Z4):438-444.

Zhang, Y, Cai, J, Wang, S, He, K, Zheng, M, 2017. Review of receptor-based source apportionment research of fine particulate matter and its challenges in China. *Sci Total Environ* 586:917-929.

Zheng, M, Yan, C, Wang, S, He, K, Zhang, Y, 2017. Understanding PM2.5 sources in China: challenges and perspectives. *Natl Sci Rev* 4(6):801-803.

Zhou, H, Zhang, F, Du, Z, Liu, R, 2021. Forecasting PM2.5 using hybrid graph convolution-based model considering dynamic wind-field to offer the benefit of spatial interpretability. *Environ Pollut* 273:116473.

Zhu, J, Deng, F, Zhao, J, Zheng, H, 2021. Attention-based parallel networks (APNet) for PM2.5 spatiotemporal prediction. *Sci Total Environ* 769:145082.