

Deep-MAPS: Machine Learning based Mobile Air Pollution Sensing

Jun Song^a Ke Han^{b,*} Marc Stettler^a

^a*Department of Civil and Environmental Engineering, Imperial College London, UK*

^b*Institute of System Science and Engineering, School of Transportation and Logistics, Southwest Jiaotong University, China*

*Corresponding author: kehan@swjtu.edu.cn

October 25, 2020

Abstract

Mobile and ubiquitous sensing of urban air quality has received increased attention as an economically and operationally viable means to survey atmospheric environment with high spatial-temporal resolution. This paper proposes a machine learning based mobile air pollution sensing framework, called *Deep-MAPS*, and demonstrates its scientific and financial values in the following aspects. (1) Based on a combination of fixed and mobile air quality sensors, we perform spatial inference of PM_{2.5} concentrations in Beijing (3,025 km², 19 Jun-16 Jul 2018) for a spatial-temporal resolution of 1km×1km and 1 hour, with under 15% SMAPE. (2) We leverage urban big data to generate insights regarding the potential cause of pollution, which facilitates evidence-based sustainable urban management. (3) To achieve such spatial-temporal coverage and accuracy, Deep-MAPS can save up to 90% hardware investment, compared with ubiquitous sensing that relies primarily on fixed sensors.

Keywords: air quality; ubiquitous sensing; big data; machine learning

1 INTRODUCTION

Severe deterioration of urban *air quality* (AQ) around the globe, especially in developing countries due to aggressive urbanization and motorization, has posed barriers to economic development and major threats to public health. Health Effects Institute [1] estimates that 95% of the world’s population is exposed to air

quality considered unsafe by the World Health Organization (WHO). On a global scale, air pollution is responsible for 5 million deaths each year and contributes to 9% of deaths, which varies by country from 2% to 15%. In China, PM_{2.5} pollution concentrations remain above the World Health Organization’s least-stringent target [2].

Detailed information about air pollution, including its sources and dynamics on a city scale, is of critical importance to public health and sustainable urban management [3]. For decades, fixed-location AQ monitoring stations have provided AQ information for various purposes including information dissemination, AQ inference/prediction, and policy appraisal. In recently years, mobile air quality sensing has received increased attention as an economically and operationally viable means to survey urban atmospheric environment [4, 5, 6]. Compared to fixed monitoring stations, the array of mobile sensors can offer high-resolution and dense coverage of urban atmospheric environment at comparatively low costs [7]. A number of initiatives have taken place world wide to collect AQ information based on mobile sensing and communication technologies, as reported by [8, 9, 10, 11, 12].

The global market for ubiquitous AQ sensing is rapidly expanding with a 14.3% compounded average growth rate, reaching \$530m by 2024 [13]. In China, public funding allocated to air pollution mitigation has reached 25bn RMB (3.6bn USD) in 2019, with a 25% increase over 2018 [14]. Given the major investment in hardware, however, the efficacy of mobile

sensing remains rather limited, as information available is restricted to sensor measurements without effective mining of data essential to decision making. In this paper, we demonstrate the scientific and financial values of Deep-MAPS, a machine learning framework in mobile air pollution sensing, in three main ways: (1) It allows air quality information to be inferred at fine granularity ($1\text{km} \times 1\text{km}$, hourly) at city scale, even with sparse fixed and mobile sensing coverage. (2) It leverages urban big data to generate insights regarding the potential cause of pollution, which facilitates evidence-based sustainable urban management. (3) It has the potential to substantially reduce the capital investment of high-density AQ monitoring in urban areas.

2 Related work

2.1 Bottom-up air pollution models

The physical/chemical interaction and regional transport of air pollutants are traditionally studied using bottom-up approaches, based on knowledge of the emission sources and a number of theoretical or empirical assumptions. [15] and [16] apply the Gaussian Plume Dispersion Model to roadside dispersion modeling. [17] and [18] have developed and applied the Operational Street Pollution Model to simulate atmospheric dynamics for street canyons based on the Gaussian Plume Model. The Gaussian-type models rely on the stationarity assumption, which limits its applications to fine-granular temporal dynamics. Computational Fluid Dynamic Models [19, 20] use differential equations and numerical schemes to simulate fluid or gas flows with given boundary conditions, which have been widely implemented in software [21], but suffers from significant computational burden. Chemical Transport Models [22, 23] simulate the chemical processes involving multiple species in the atmosphere. The CMAQ model and software [24] target multiple pollutants at multiple scales, and offer decision support tools to assess the impact of AQ related management measures. These bottom-up models offer high model fidelity pertaining to the physical and chemical processes of air pollution, but they tend to be computationally expensive and rely on assumptions that limit their integration with high frequency and ubiquitous AQ sensing data.

2.2 AQ inference and prediction based on machine learning

Another line of research aims to infer and predict AQ by learning its spatial and temporal correlation with explanatory factors such as urban layouts, traffic activities, and meteorological conditions, based on machine learning techniques. [27] propose a spatial-temporal classifier and a co-training framework to infer the air quality index in Beijing using fixed-location sensing data. [28] also consider fixed monitoring station data in Beijing to infer and predict PM_{2.5} concentrations based on semi-supervised learning. [29] propose a multi-group encoder-decoder approach to fuse heterogeneous features for next-day air quality prediction at station level.

Mobile air quality sensing has been recently introduced to AQ inference. [30, 31] use mobile data collected in Lausanne and built probabilistic models and deep learning models to estimate street-level air quality. However, these inference results have limited temporal resolution (daily, weekly, monthly), which cannot reveal detailed dynamics of urban air pollutants. [32] apply several regression methods to estimate CO concentration in Sydney based on fixed and mobile sensing data, but without using any additional explanatory data or urban features to reasonably generalize their results beyond statistical fitting.

Mobile air quality sensing has advanced the development and application of land use regression (LUR) models. [33] use over 85 hrs of data on a bicycle-based sampling platform in Minneapolis, and constructed LUR models for particle size, black carbon, and PM_{2.5} with modest goodness-of-fit. Vehicle-based mobile measurements were also applied to create LUR models to estimate the spatial variation of street-level PM_{2.5} and PM₁₀ in the downtown area of Hong Kong [34]. [35] develop LUR models with machine learning techniques for street-level PM_{2.5} concentration levels in Seoul. [36] analyze data collected from a Google Street View mapping vehicle equipped with air quality sensors that repeatedly sampled every street in a 30-km^2 area of Oakland, to model and reveal urban air pollution patterns at a very high spatial granularity. All these studies on LUR models, however, tend to have low temporal resolution (e.g. days, weeks), which limit their capabilities to reveal detailed pollution dynamics.

This paper proposes an AQ inference framework that is capable of producing fine-granular pollution maps ($1\text{km} \times 1\text{km}$, hourly) based on fixed and mobile

sensing data with limited spatial-temporal coverage, while ensuring the robustness and accuracy of the results by leveraging multi-source urban datasets.

2.3 Machine learning based on spatial-temporal data

The proliferation of urban big data in the past decade, enabled by information and communication technologies and pervasive sensing, has resulted in rapid development of spatial-temporal data analytics [37] and computational architectures [38]. Studies have proposed different ways to capture spatial and temporal correlations. Deep neural networks such as Long-Short-Term Memory (LSTM) allow long-term temporal dependencies as well as short-term variations to be simultaneously captured in time-series prediction [39]. [40] forecast the flow of crowds in an urban area using convolutional neural networks (CNN) by viewing traffic intensity in a region as a pixel on an image. [41] combine CNN and LSTM (ConvLSTM) to perform precipitation nowcasting. [42] predict spatial and temporal taxi demands using historical data based on a Multi-View Spatial-Temporal Network to simultaneously capture spatial correlations and temporal dependencies as well as semantic relations.

We note that in all these studies, either the training labels are available for the entire spatial domain at a particular time (e.g. precipitation, taxi demand), or they are available at several discrete locations with non-interrupted time series (e.g. fixed-location AQ sensing, traffic flow). The problem considered in this paper presents its own challenge, as the training labels (i.e. mobile data) are irregularly distributed in the space-time domain.

3 Problem and data description

3.1 Problem definition

We begin by presenting a set $\mathcal{S}$ of $1\text{km} \times 1\text{km}$ spatial grids in the study area, and set $\mathcal{T}$ of 1-hr periods. A *unit* is defined to be an element $u = (s, t) \in \mathcal{S} \times \mathcal{T}$. The main task undertaken in this work is *spatial inference* for PM2.5, as follows.

Definition 3.1. (Spatial inference) *Given the subset $\mathcal{U}_L \subset \mathcal{S} \times \mathcal{T}$ containing tabled units, where the label g_u denotes the PM2.5 concentration (obtained from fixed or mobile sensing) in $u \in \mathcal{U}_L$, the problem of PM2.5 spatial inference is to estimate the PM2.5 concentrations $g_{u'}$ for all $u' \in \mathcal{S} \times \mathcal{T} \setminus \mathcal{U}_L$.*

3.2 Fixed and mobile air quality data in Beijing

As shown in Fig 1(a), 21 fixed-location air pollution monitoring stations are located in a $55\text{km} \times 55\text{km}$ square area in Beijing, which generate hourly measurements of particulate matters (PM2.5, PM10) and gaseous pollutants (NOx, CO, SO2, O3).

Fig 1(a)-(b) show two adjacent fixed monitoring stations in Beijing and their PM2.5 measurements, which show similar trends yet with considerable differences for certain times. This is caused by local variation of pollution concentrations that cannot be adequately observed by sparsely distributed monitoring stations.

The mobile air quality measurements were collected by 15 probe vehicles traveling through the city in 28 days (2018.06.19-2018.07.16). The mobile sensing data include PM2.5 and PM10 concentrations as well as temperature and relative humidity. The trajectories of the 15 probe vehicles have been map-matched onto streets as shown in Fig 1(c). The mobile PM2.5 measurements were pre-processed, geo-meshed and aggregated into $1\text{km} \times 1\text{km}$ spatial grids and hourly intervals.

The spatial and temporal distribution of the fixed and mobile data (labels) is illustrated in Fig 2. In particular, Fig 2(a) shows the number of spatial grids covered by mobile sensing at every hour throughout the study period. The numbers are typically between 50 and 150 (grids), which is less than 5% of the 3,025 grids in total and hence demonstrates the spatial sparsity of the data. Furthermore, Figs 2(b-d) illustrate the temporal sparsity of mobile sensing measurements by showing that the majority of the grids receive less than 50 hrs (7.4% of the experiment period) of sensing coverage. The challenge posed by spatial-temporal sparsity is addressed in this paper by exploring their spatial-temporal correlations in conjunction with a wide variety of urban features that encapsulate the configuration and dynamics of the city.

3.3 Multi-source urban datasets

A detailed and balanced representation of urban features is critical for data-driven methods to reduce overfitting and increase the interpretability of the model. A total of 62 features are grouped into four categories, which are described below and summarized in Table 1.

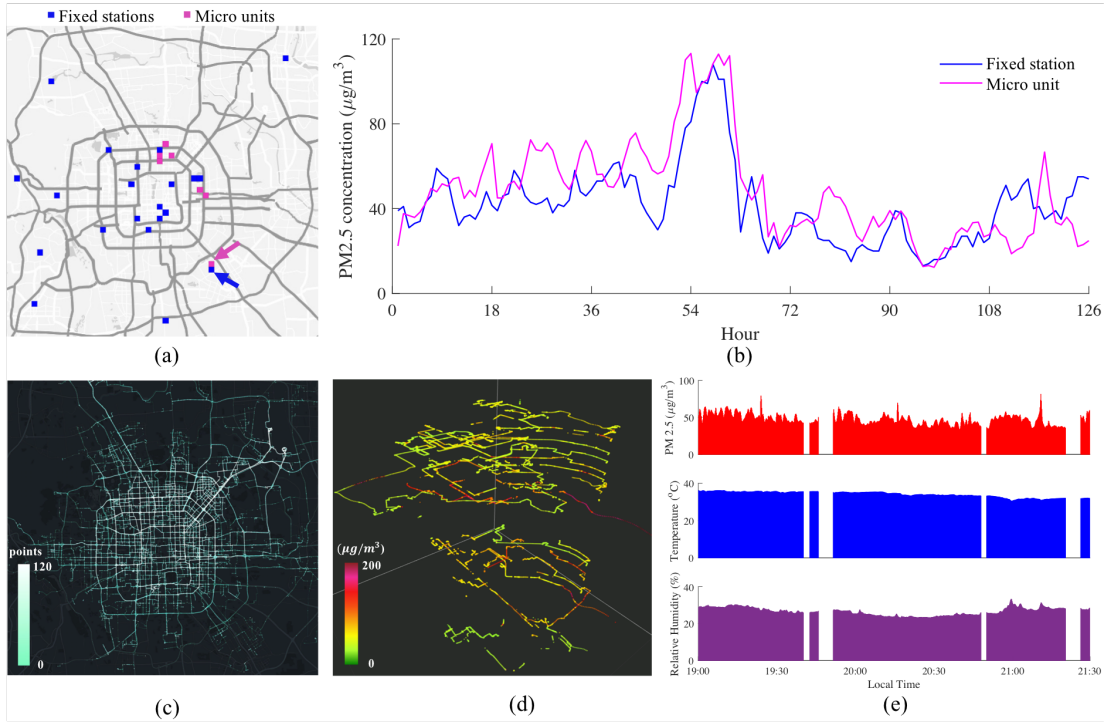


Figure 1: (a) The 55km×55km study area in Beijing with 21 fixed locations and 8 micro units. (b) Hourly PM2.5 measured by two adjacent stations identified in (a). (c) Street level coverage of mobile sensing during the study period. (d) Spatial-temporal trajectory of a probe vehicle between 28 June and 10 July. (e) Time series of PM2.5 concentration, temperature and relative humidity measured by a probe vehicle.

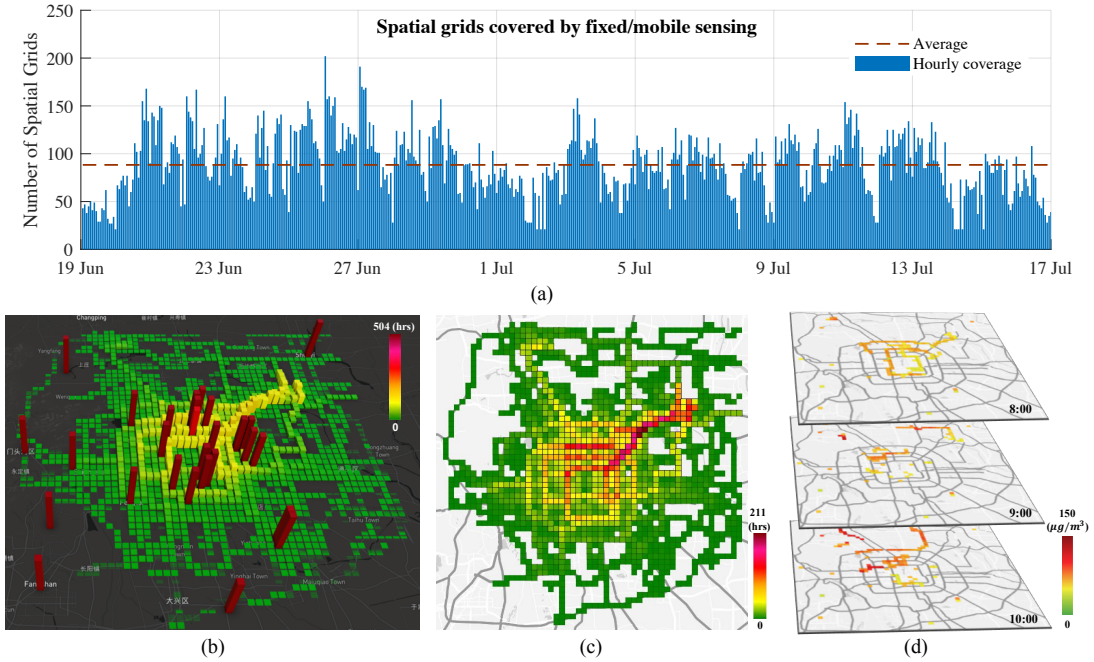


Figure 2: Temporal and spatial coverage of fixed/mobile sensing. (a): Number of grids covered by fixed/mobile sensing on an hourly basis. (b): Number of hourly measurements in each cell (fixed and mobile); (c): Number of hourly measurements in each cell (mobile only). (d): Fixed and mobile coverage (PM2.5) between 8:00-10:00 on 2018.06.24.

Geography & Land Use - includes points of interest (POI) (21 categories including building, factory, commercial area etc.). Within a 1km×1km spatial grid, the numbers of different types of POIs, as well as the percentages of landmass they occupy (Area of Interest, AOI) are calculated. In addition, the percentages of green and water covers, as well as digital elevation are included to further characterize the geographic features of the grid.

Transport - contains road network structure and dynamic traffic conditions. Within a 1km×1km spatial grid, the former contains number of signalized intersections and the lengths of roads of different categories (primary, secondary, tertiary and quaternary). These static features are relevant to traffic volume and fleet composition, as well as the stop and-go driving cycles, which are known to contribute to vehicle emissions. The dynamic traffic information is captured

by the percentages of roads with light, medium and heavy traffic conditions. Such a categorical feature is a direct indicator of congestion levels.

Public Vitality - refers to the intensity of social media activities enabled by Location Based Services (LBS). These include Wechat and Sina Weibo log-ins, posts and comments, which are indicators of public vitality in the area. The public vitality data are deemed relevant to instantaneous population density and indirect indicators of land use type (such as business districts and rural areas).

Meteorology - contains hourly information of local temperature, pressure, relative humidity, wind direction and speed, which are measured at 13 meteorological stations (9 within the study area). The hourly meteorological data were used to spatially interpolate grid-based quantities using the inverse distance weighting method.

Data Category	Data Type		Space-Time Resolution
Geographic & Land Use	Point of Interest (21 categories)		1km×1km Static (2018)
	Area of Interest (21 categories)		
	Area of Green Cover (%)		
	Area of Water Cover (%)		
	Digital Elevation (m)		
Transport	Network	No. of signalized intersections	1km×1km Static (2018)
		Total lengths of different types of roads	
	Traffic Conditions	Percentages of roads with light/medium/heavy conditions; categorized traffic conditions	1km×1km Hourly (2018)
Public Vitality	Social media Wechat & Sina Weibo	Number of users posts, and comments	1km×1km Hourly or annual average (2015-18)
Meteorology	Temperature (°C)		1km×1km Hourly (2018)
	Pressure (kPa)		
	Water vapor pressure (hPa)		
	Relative humidity (%)		
	Wind direction ([0°, 360°])		
	Wind speed (m/s)		
Air Quality	Fixed Stations	PM2.5 concentration ($\mu g/m^3$)	Point-based Hourly (2018)
		Location ([lat, lng])	
	Mobile Sensors	Time, location	Street level 2s - 15s (2018)
		Temperature (°C)	
		Relative humidity (%)	
		PM2.5 concentration ($\mu g/m^3$)	

Table 1: Supporting data used for fine-granular urban AQ inference in Beijing.

4 The Deep-MAPS framework

4.1 Preprocessing of mobile data

The mobile sensors used in the case study are first calibrated in a lab environment before deployment to ensure that they provide stable and reliable readings. Considering the variability of each individual low-cost mobile sensor, the mobile sensing box equipped on each probe vehicle is periodically (weekly) replaced by another box that has been placed near the fixed-location reference station recently.

The raw mobile data include time stamp, location (coordinates), PM2.5 concentration, temperature and relative humidity. These time series data are denoised using robust local regression to reduce temporal variations and eliminate extreme values. The resulting spatial-temporal data are then aggregated onto spatial grids and time intervals via arithmetic average.

The mobile and fixed sensors have distinct sensing environment; that is, the former are deployed on the street level while the latter usually on higher grounds or buildings. To reconcile their sensing data, we fur-

ther calibrate the mobile PM2.5 data using fixed-location data as the reference instruments [25, 26]. This is achieved by using multivariate regression involving the location (grid s), time (hourly interval t), temperature ($temp$) and relative humidity (RH). Specifically, the multivariate linear regression is expressed as:

$$y_f = \beta_0 + \beta_1 \cdot y_m + \beta_2 \cdot s + \beta_3 \cdot t + \beta_4 \cdot temp + \beta_5 \cdot RH$$

where y_f and y_m denote fixed and mobile PM2.5 measurements, respectively, which are collected from the same grid s at the same time interval t . The regression result is visualized in Figure 3, with a goodness-of-fit coefficient $R^2 = 0.70$.

Let β_i^* ($i = 1, \dots, 5$) be the coefficients obtained from the multi-variate linear regression, then for any mobile data $\hat{y}_m$ not sharing the same grid with any fixed station, its value is calibrated as:

$$\bar{y}_m = \beta_0^* + \beta_1^* \cdot \hat{y}_m + \beta_2^* \cdot s + \beta_3^* \cdot t + \beta_4^* \cdot temp + \beta_5^* \cdot RH$$

where $s, t, temp, RH$ are associated with the measurement $\hat{y}_m$.

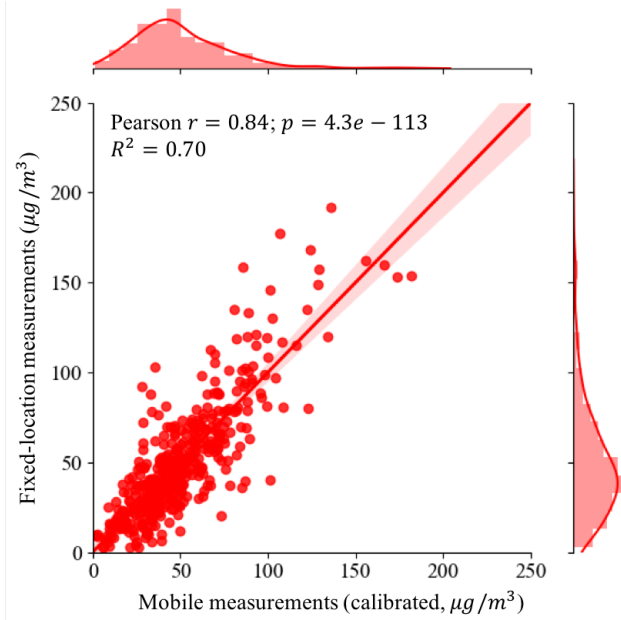


Figure 3: Multi-variate linear regression results with 1348 data points.

4.2 Machine learning for PM2.5 inference

Deep-MAPS is a machine learning framework for spatial inference of air pollutants based on fixed and mobile AQ sensing. This problem involves multi-source

and heterogenous datasets that significantly differ in granularity, numerical scale, and veracity. Therefore, it is crucial to define and extract relevant urban features with appropriate spatial and temporal struc-

tures to capture not only internal factors contributing to local AQ (such as transportation, buildings, meteorology), but also indicators of external influence (such as regional transport of pollutants). This section describes the feature engineering process and the machine learning model within Deep-MAPS;

The PM2.5 concentration of a given grid at a given time spatially and temporally related to its neighbors in the space-time domain. Unlike AQ inference based on fixed monitoring stations such as [27, 28], the la-

bels provided by mobile sensing are sparsely and irregularly distributed. To tackle this challenge, we propose a machine learning framework based on Gradient Boosting Decision Tree (GBDT), which relies on a specific feature extraction framework to account for the spatial-temporal correlations of the physical characteristics relevant to urban air quality. The feature extraction framework is shown in Fig 4 and detailed in subsequent sections.

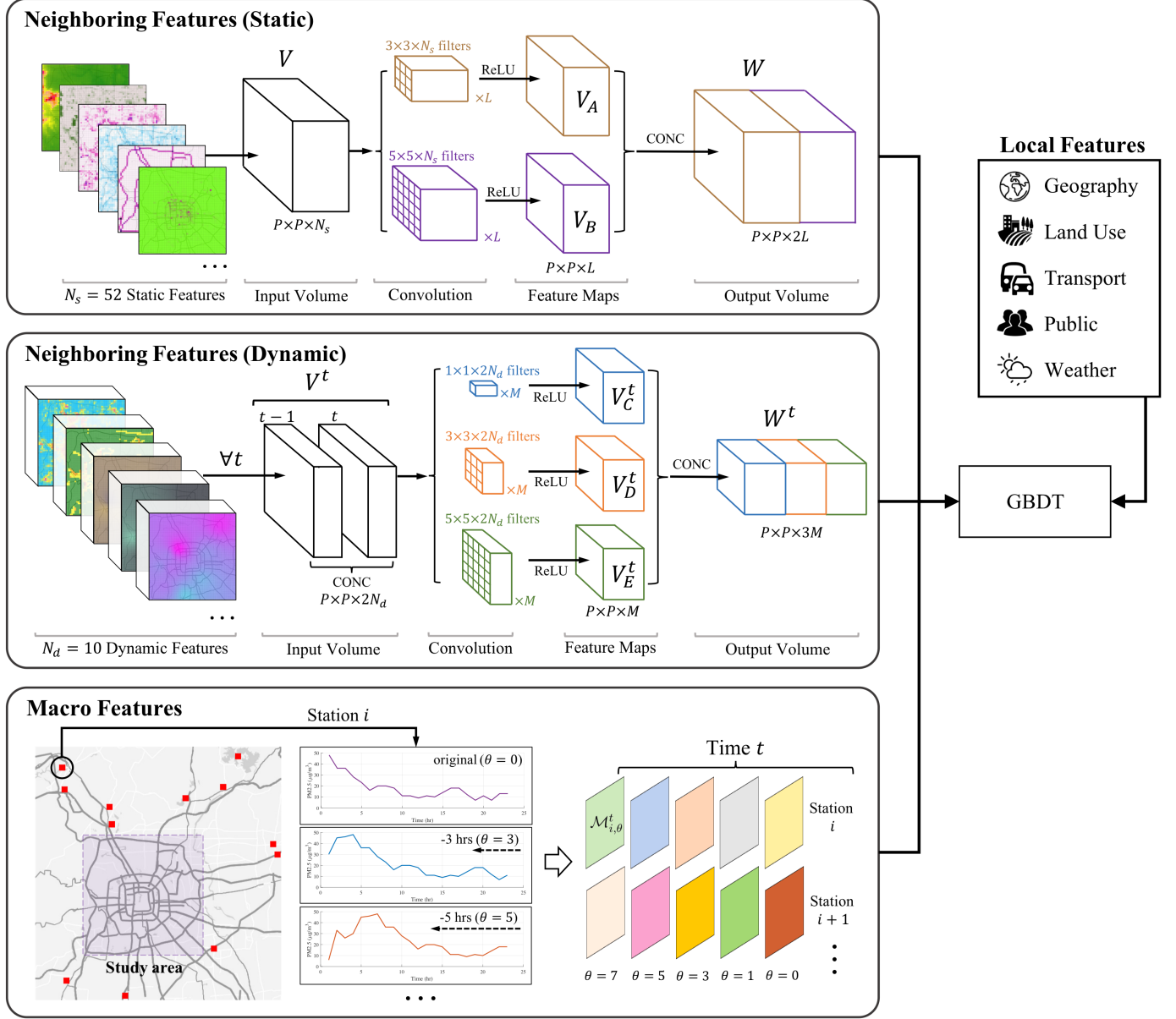


Figure 4: Feature extraction framework for Deep-MAPS.

4.2.1 Local features

Local features include geographic & land use characteristics, transport, public vitality, and meteorological conditions, which are defined for each grid s in the case of static feature, and grid-time tuple (s, t) in the case of dynamic feature.

4.2.2 Neighboring features

Neighboring features are defined for each grid-time tuple (s, t) , by including features in a neighborhood of (s, t) in the space-time domain. This allows us to capture the spatial-temporal correlations among the urban features. The extraction of neighboring features follows the technique of convolution, and we distinguish between static and dynamic features as follows.

The N_s static features are treated as N_s input images of size $P \times P$, forming an input volume $V \in \mathbb{R}^{P \times P \times N_s}$ (where $N_s = 52$, $P = 55$ in the case study). Two sets of filters, $\{A^1, \dots, A^L\} \subset \mathbb{R}^{3 \times 3 \times N_s}$ and $\{B^1, \dots, B^L\} \subset \mathbb{R}^{5 \times 5 \times N_s}$ are applied to V with stride 1 and padding 0, where the k -th channel A_k^i (B_k^i) of A^i (B^i) is a mean filter with a random multiplicative weight following standard Normal distribution. Convolution with these two sets of filters, followed by the rectifier $f(x) \doteq \max(0, x)$ as activation, results in two feature maps

$$V_A = \oplus (V_A^1, \dots, V_A^L), \quad V_B = \oplus (V_B^1, \dots, V_B^L) \quad \text{where} \\ V_A^i = f(V * A^i), \quad V_B^i = f(V * B^i), 1 \leq i \leq L \quad (4.1)$$

Here, $*$ is the convolution operator; $\oplus$ is the concatenation operator. The mean filters perform arithmetic average within their receptive fields, while the random weights are used to sample different combinations of the input features. Finally, the output is $W = \oplus(V_A, V_B) \in \mathbb{R}^{P \times P \times 2L}$; see Figure 4(top).

The N_d dynamic features ($N_d = 10$ in our case study) are convolved in a similar way. For every time t , the input volume V^t consists of the dynamic features of the present and previous time steps to account for the temporal dependencies. We then apply three sets of filters $\{C^1, \dots, C^M\} \subset \mathbb{R}^{1 \times 1 \times 2N_d}$, $\{D^1, \dots, D^M\} \subset \mathbb{R}^{3 \times 3 \times 2N_d}$, $\{E^1, \dots, E^M\} \subset \mathbb{R}^{5 \times 5 \times 2N_d}$, whose structures are similar to those of A^i and B^i , and obtain:

$$V_C^t = \oplus (V_C^{t,1}, \dots, V_C^{t,M}), \\ V_D^t = \oplus (V_D^{t,1}, \dots, V_D^{t,M}), \\ V_E^t = \oplus (V_E^{t,1}, \dots, V_E^{t,M}). \quad (4.2)$$

where

$$V_C^{t,i} = f(V^t * C^i), \\ V_D^{t,i} = f(V^t * D^i), \\ V_E^{t,i} = f(V^t * E^i), \quad (4.3)$$

$1 \leq i \leq M$. The output for time t is $W^t = \oplus(V_C^t, V_D^t, V_E^t) \in \mathbb{R}^{P \times P \times 3M}$; see Figure 4(middle).

4.2.3 Macro features

For many cities, severe air pollution may be caused by power production and industrial activities in surrounding areas. To account for the regional transport of pollutants without significantly expanding the study area, we utilize PM2.5 data from fixed monitoring stations outside the study area as additional features (coined *macro features*). The premise is that such data are likely to be temporally cross-correlated with the background pollutant concentration in Beijing. By introducing the macro features, we also reduce the risk of overfitting as the AQ inference results are less dependent on locally defined features.

We collect PM2.5 concentration data from 12 fixed monitoring stations outside the study area as shown in Figure 4(bottom). By analyzing these time series (with possible time shifts depending on the distance of the stations and average wind speed), we may indirectly account for the transport dynamics of PM2.5. Specifically, $\forall 1 \leq i \leq 12$, let $\{y_i(t), t \in \mathbb{Z}\}$ be the time series data from the i -th station. For a given time t , we define the macro features for all the grids at t to be $\{y_i(t - \theta) : 1 \leq i \leq 12, \theta \in \Theta\}$, where Θ is the set of backward time shifts, which is determined in conjunction with the distance from the study area to the i -th station and the average wind speed. Alternatively, the macro features can be viewed as a set of $P \times P$ images:

$$\{\mathcal{M}_{i,\theta}^t : 1 \leq i \leq 12, \theta \in \Theta, t \in \mathcal{T}\} \subset \mathbb{R}^{P \times P}$$

where $\mathcal{M}_{i,\theta}^t = y_i(t - \theta) \times \mathbf{1}(P, P)$; here, $\mathbf{1}(P, P)$ denotes a $P \times P$ square matrix of one's; see Figure 4(bottom).

4.2.4 Model training within Deep-MAPS

In this paper, we employ Gradient Boosting Decision Trees (GBDT) to fit air pollution data with local, neighboring and macro features discussed above. Other variants of GBDT, such as XGBoost [45] and hybrid boosted decision trees [46], were also considered, which showed similar inference accuracy. On the other hand, deep neural networks (such as ConvLSTM) typically require much longer training time

and are not particularly robust against the input data structure, which leads to non-convergence and over-fitting.

The overall training procedure of Deep-MAPS is summarized in Algorithm 1.

Algorithm 1: Training Procedure for Deep-MAPS

Input : Static features $\{\mathcal{S}_1, \dots, \mathcal{S}_{N_s}\} \subset \mathbb{R}^{P \times P}$
Dynamic features $\{\mathcal{D}_1^t : t \in \mathcal{T}\}, \dots, \{\mathcal{D}_{N_d}^t : t \in \mathcal{T}\} \subset \mathbb{R}^{P \times P}$
Macro features $\{\mathcal{M}_{i,\theta}^t : 1 \leq i \leq 12, \theta \in \Theta, t \in \mathcal{T}\} \subset \mathbb{R}^{P \times P}$

Output: Trained Deep-MAPS model for AQ inference

```

1 begin
2   Concatenate static features to form the input volume:  $V = \oplus (\mathcal{S}_1, \dots, \mathcal{S}_{N_s})$ ;
3   For  $1 \leq i \leq L, 1 \leq k \leq N_s$ , generate i.i.d. weights  $w_k^i, \omega_k^i \sim N(0, 1)$ .
4    $A_k^i \leftarrow w_k^i \times \frac{1}{9} \times \mathbf{1}(3, 3)$ ,  $B_k^i \leftarrow \omega_k^i \times \frac{1}{25} \times \mathbf{1}(5, 5)$ ;
5    $W = \oplus (V_A, V_B)$  where  $V_A, V_B$  are given by (4.1).
6 end
7 for  $\forall t \in \mathcal{T}$  do
8   Form the input volume  $V^t = \oplus (\mathcal{D}_1^t, \dots, \mathcal{D}_{N_d}^t, \mathcal{D}_1^{t-1}, \dots, \mathcal{D}_{N_d}^{t-1})$ ;
9   For  $1 \leq i \leq M, 1 \leq k \leq 2N_d$ , generate i.i.d. weights  $v_k^i, \nu_k^i, \mu_k^i \sim N(0, 1)$ ;
10   $C_k^i \leftarrow v_k^i \times \mathbf{1}(1, 1)$ ,  $D_k^i \leftarrow \nu_k^i \times \frac{1}{9} \times \mathbf{1}(3, 3)$ ,  $E_k^i \leftarrow \mu_k^i \times \frac{1}{25} \times \mathbf{1}(5, 5)$ ;
11   $W^t = \oplus (V_C^t, V_D^t, V_E^t)$  where  $V_C^t, V_D^t, V_E^t$  are given by (4.2)-(4.3).
12 end
13 Train the GBDT model using input features  $W, \{W^t : t \in \mathcal{T}\}$  and  $\{\mathcal{M}_{i,\theta}^t : 1 \leq i \leq 12, \theta \in \Theta, t \in \mathcal{T}\}$ ;

```

5 Results and discussion

5.1 Beijing case study

The study area is a 55km×55km square in Beijing, covering the Sixth Ring Road as shown in Fig 5. This is divided into 3,025 1km×1km spatial grids. The spatial inference is carried out for each grid with an hourly resolution in a four-week period (2018.06.19-2018.07.17). We aim to infer the hourly PM2.5 concentrations in the study area, with a total output volume of $3,025 \times 28 \times 24 = 2,032,800$.

5.2 Validation of spatial inference results for PM2.5

A total of 50,736 labels were obtained from fixed-location and mobile sensing data, where the mobile data have been calibrated following the procedure in Section 4.1. A five-fold cross validation is applied to the set of 50,736 data points, along with Root Mean Square Error (RMSE) and Symmetric Mean Absolute Percentage Error (SMAPE), are used to assess the validity and accuracy of the machine learning model (Deep-MAPS). The definitions of RMSE and SMAPE

appear below:

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (T_i - I_i)^2}$$

$$\text{SMAPE} = \frac{2}{N} \sum_{i=1}^N \frac{|T_i - I_i|}{|T_i + I_i|} \times 100\%$$

where T_i and I_i are ground-truth and inferred PM2.5 concentrations of a given test unit. Table 2 compares Deep-MAPS with several benchmark methods, including spatial interpolation (SI) methods inverse distance weighting (IDW) and Kriging, k-nearest neighbors (KNN), support vector regression (SVR), Gaussian process regression (GPR) [47], auto encoder (AE) [48], and Deep Air Learning (DAL) [28].

The results show that Deep-MAPS, with all combinations of urban features (local features, neighboring features and macro features), outperforms all other benchmark methods. The best performance is reached with 'N+M' and 'L+M+N', with SMAPEs below 15%. In addition, comparing 'L' with 'L+M' (or 'N' with 'N+M') indicates that the macro features significantly improves the accuracy of KNN, SVR and

Deep-MAPS. This means that it is important to consider regional transport of pollutants, which was not done in any of the machine learning models proposed to date. For Deep-MAPS, the neighboring features

‘N’ (or ‘N+M’) outperforms the local features ‘L’ (or ‘L+M’), suggesting that the spatial and temporal relationships among urban features should not be ignored in the fitting of AQ data.

	Feature	RMSE ($\mu g/m^3$)	SMAPE (%)
SI (IDW)	/	17.8	21.9
SI (Kriging)	/	20.8	26.6
KNN	L	23.4	25.9
	L+M	15.4	17.3
SVR	L	36.5	38.4
	L+M	27.5	31.4
GPR	L	19.8	24.5
	L+M	19.2	23.6
AE	L	17.1	20.4
	L+M	16.9	19.8
DAL	L	15.6	18.3
	L+M	15.3	17.7
Deep-MAPS	L	14.4	16.4
	L+M	13.9	15.5
	N	14.7	15.8
	N+M	14.0	14.9
	L+M+N	13.2	14.7

Table 2: Performance of different methods and feature combinations. L=local features; N=neighboring features; M=macro features.

The result of spatial inference is illustrated for 5 July 2018 in Fig 5. It is noted that between 8:00 and 15:00, the PM2.5 concentrations in the west half of Beijing is higher than the rest, which was likely caused by the higher relative humidity in the west side measured by meteorological stations. Moreover, the hourly PM2.5 concentration maps show relatively high and steady concentrations during most of the day before a sudden drop between 16:00-19:00, which was likely caused by the thundershower in late afternoon. This temporal trend agrees with the PM2.5 concentration in Beijing obtained by averaging fixed-location data, as shown on top of Fig 5. The fine-grained pollution concentration maps could offer critical insights into population exposure and potential cause of local concentrations.

The proposed inference method is further tested by restricting the test set to consist of only fixed sensing labels, while the training set consists of (i) fixed sensing data; or (ii) fixed and mobile sensing data. The first case can be seen as an extension to spatial inference based on fixed data only (e.g. [27, 28]), while the second case supplements the training set with mobile data. The test results based on a five-fold cross validation are:

- (i) SMAPE = 17.76%, (ii) SMAPE = 15.93%

The inclusion of mobile data reduces the SMAPE by nearly 2%, because a spatially diversified survey of urban features and pollutant concentrations allows the model to better capture their spatial and temporal correlations and reduce overfitting.

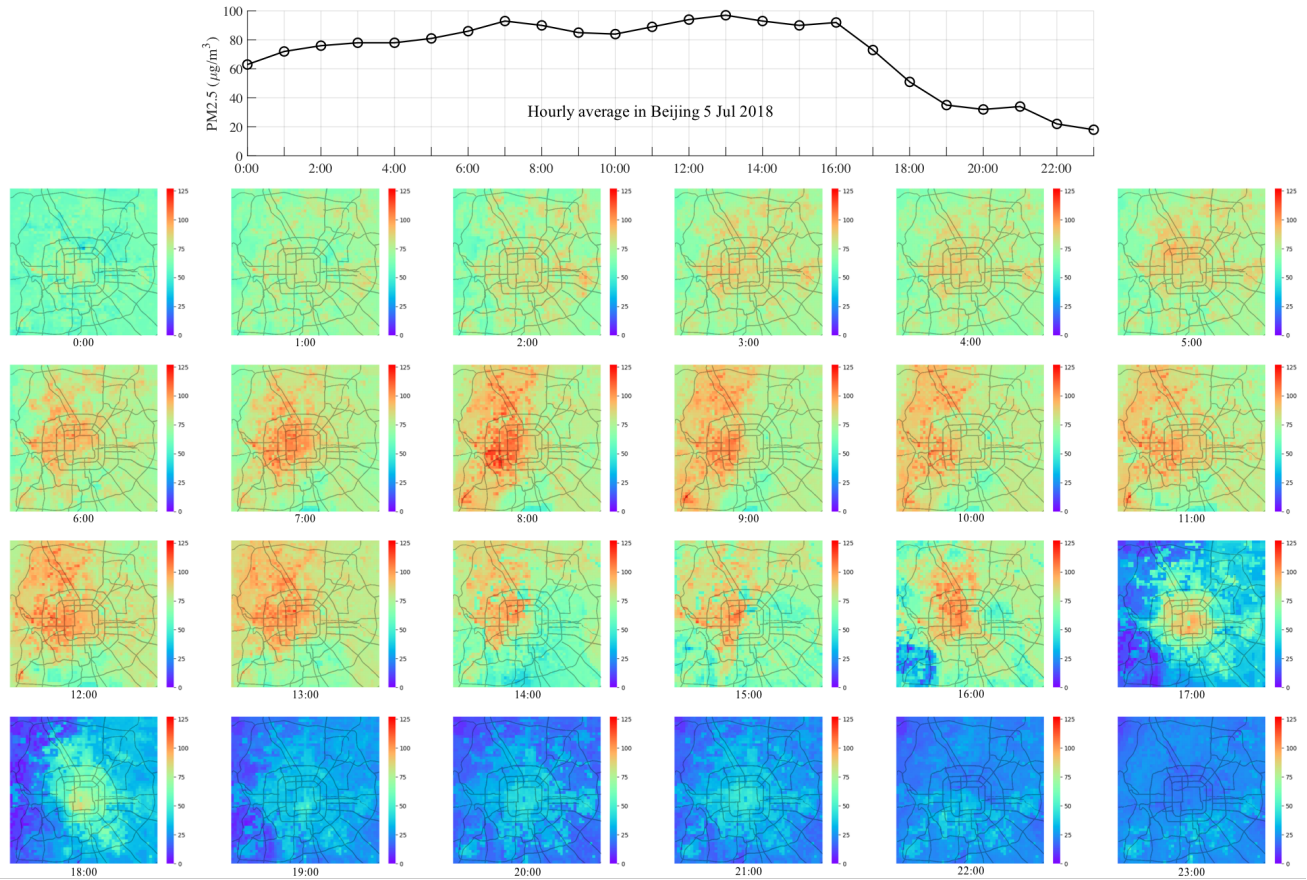


Figure 5: PM2.5 concentrations ($\mu\text{g}/\text{m}^3$) in Beijing on 5 July 2018.

5.3 Impact of mobile sensing coverage

One of the practical concerns of mobile sensing is the amount of mobile coverage required to reach certain level of inference accuracy. To address this problem, we set aside an independent test set of size 9,200 consisting of randomly selected fixed and mobile labels. For the training set, we consider six different cases: all the fixed labels plus $x\%$ of mobile labels, $x = 0, 20, 40, 60, 80, 100$.

Fig 6 shows the performance of Deep-MAPS with these six training sets. All three performance indices exhibit improving trends as more mobile data points are included in the training set. In particular, the ‘elbow’ point occurs where mobile data first appear in the training set, even in low quantity. Nevertheless, with every increment of mobile data the margin of improvement declines. This is interesting as it suggests that in order to achieve certain level of accuracy,

Deep-MAPS only requires limited amount of mobile sensing data. In our case, less than 5% of the spatial coverage suffices to achieve SMAPE below 15%.

In the next experiment, we test the robustness of the proposed method in a sensitivity analysis, by analyzing its performance on grids with varying temporal coverage. This is driven by the observation that the temporal coverage of different grids differs significantly, as shown in Figure 2(b)-(c), thus it is necessary to examine the variability of the method’s performance at these grids. Table 3 shows the SMAPE of Deep-MAPS at grids whose temporal coverage is below y hrs. It can be seen that the performance slightly improves as y increases, suggesting that the method provides a better fit for spatial units with more sample labels. Yet, the difference of RMSE is below $1.4 \mu\text{g}/\text{m}^3$ and difference of SMAPE is 1.5%, which shows the robustness of the proposed method in handling grids with varying temporal coverage.

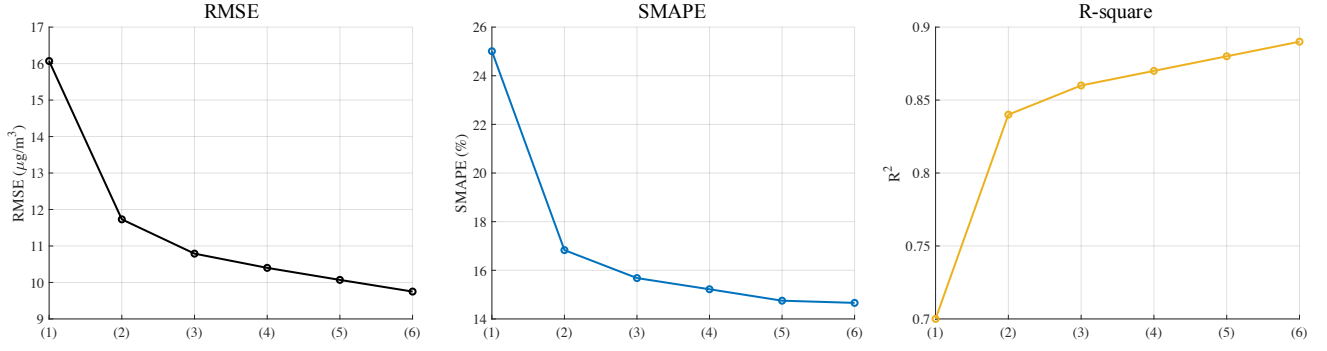


Figure 6: Performance of Deep-MAPS on an independent test set with different training data. (1): Fixed-location data only; (2): Fixed data + 20% mobile data; (3): Fixed data + 40% mobile data; (4): Fixed data + 60% mobile data; (5): Fixed data + 80% mobile data; (6): Fixed data + 100% mobile data.

y	10	20	30	60	90	120	150	180	210
No. of grids	983	1170	1270	1438	1509	1552	1588	1601	1613
RMSE ($\mu g/m^3$)	15.73	15.43	14.91	14.54	14.46	14.43	14.40	14.35	14.34
SMAPE (%)	16.48	16.07	15.84	15.37	15.20	15.15	15.09	15.01	14.98

Table 3: Performance of Deep-MAPS on grids with $\leq y$ hrs of temporal coverage.

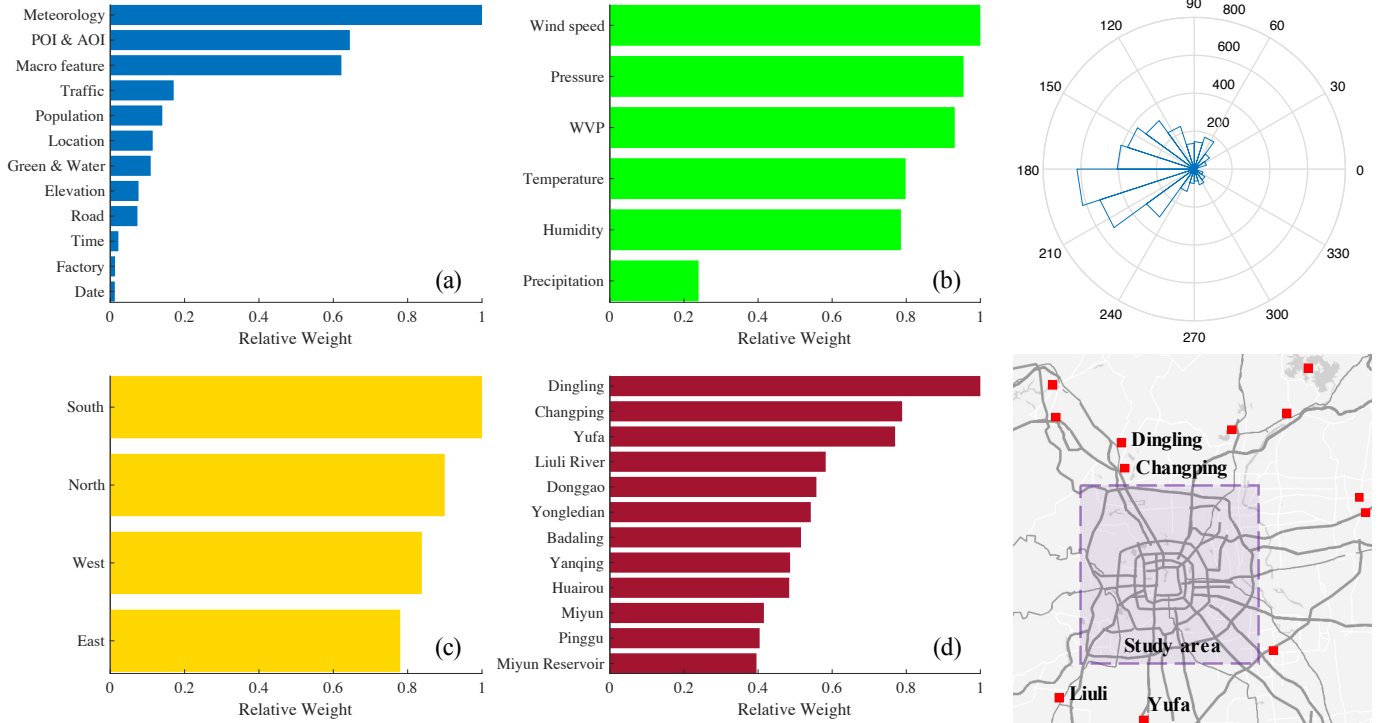


Figure 7: Feature analysis. (a)-(b) show the relative weights/importance of features internal to the study area. (c)-(d) indicate the regional transport of PM2.5

5.4 Feature analysis and interpretation of the inference results

We analyze the relative importance of different urban features by investigating their weights generated by GBDT, in order to generate insights regarding poten-

tial causes and relevant factors of PM2.5 concentrations. The weight of an input feature measures its contribution to the reduction in loss by rendering a better fit between model outputs and observations.

Based on results generated from Deep-MAPS , Fig

7(a) shows the relative importance of main feature categories, among which the top four are meteorology, POI&AOI, macro feature (i.e. regional transport of pollution), and traffic conditions. Similar studies, such as [27], confirm the importance of meteorological parameters, traffic and POI in explaining PM concentrations. Indeed, local restaurants, auto services, residential buildings etc. (represented by POI, AOI) are the main cause of PM_{2.5} formation; this is followed by traffic, which is responsible for fugitive dust and secondary reactions. However, [27] do not consider the regional transport of pollutants and their relative weights compared with some local features.

The PM_{2.5} concentration during the summer season is crucially related to meteorological parameters, the relative weights of which are shown in panel (b). It is shown that wind speed, pressure, water vapor pressure, temperature and humidity are all highly related to PM_{2.5} concentrations. This is reasonable as stronger winds disperse pollutants more quickly, and pressure is highly correlated with wind speed in the summer season. It is also known that higher humidity is favorable for PM_{2.5} formation. In [28], pressure and humidity are deemed irrelevant features for PM_{2.5} concentrations. This is because their study focuses on the winter season, which experiences little change of pressure and humidity.

Figs 7(c)-(d) focus on macro features and the regional transport of PM_{2.5}. The relative weights of individual fixed-location stations outside the study area indicate likely paths of regional transport. In particular, Dingling and Changping, located to the Northwest, have the highest weights, followed by Yufa and Liuli River, which are located to the South Southwest. This highlights two main directions of PM_{2.5} transport during 19 June - 16 July 2018. The wind rose diagram in Fig 7 provides supportive information regarding the physical transport of PM_{2.5}. This is also partially confirmed by [28], whose findings suggest that South of Beijing (Hebei Province) is a main source of industrial pollutants that affects Beijing.

6 Discussion

6.1 Interpretation of urban features

A crucial step in achieving the aforementioned AQ inference and managerial insights is the consideration of urban features, which should ideally have a balanced representation of urban configuration and dynamics contributing to the emission and formation of pollu-

nants. In this paper we consider 62 types of urban data, covering geography, land use, transport, public, and meteorology, to capture the spatial and temporal variation of urban features related to air pollution. Furthermore, a novel approach involving macro features is employed to capture the regional transport dynamics outside the study area, which not only improves the inference accuracy but also suggest likely paths of regional transport.

As we have shown in Section 5.4, Deep-MAPS allows us to interpret the inference results in terms of the contribution of each input feature. However, this is a preliminary assessment of the correlation between PM concentration and urban features. It should not be misconstrued with causality analysis, i.e. pollution attribution.

6.2 Potential impact of Deep-MAPS

In Aug 2018, the Chinese government launched the Qianliyan Initiative [43], which aims to monitor the atmospheric environment (in particular, PM_{2.5}) of the JingJinJi Area (Beijing, Tianjin and Hebei) and 28 surrounding cities. The whole area was divided into 36,793 spatial grids of size 3km×3km, and in major cities like Beijing, the spatial resolution can reach 1km×1km. The capital investment for such a large-scale AQ monitoring project, assuming at least one fixed monitoring station per spatial grid, could be drastically reduced by mobile sensing infrastructure, including Deep-MAPS. Take the study area of this paper for instance. To achieve full coverage of the 3,025 grids, the minimum capital investment is 3,025×300k, which is over 900m RMB excluding annual maintenance and admin fees, where 300k RMB/station is the price estimate for a reliable micro-station. In contrast, Deep-MAPS can achieve satisfactory inference accuracy (with SMAPE below 15%) with under 20 mobile and 30 fixed stations, the cost of which is less than 10% of the 900m investment. This is a considerable saving even with the additional costs of securing and administrating the urban big data required to run Deep-MAPS. Moreover, the benefit of Deep-MAPS goes beyond financial savings by offering additional data and insights as we conveyed in this paper.

7 Conclusion

In this paper, we propose a machine learning framework (Deep-MAPS) to for fine-granular PM_{2.5} inference based on fixed and mobile AQ sensing data.

While ubiquitous AQ sensing with low-cost and mobile sensors has become widely accessible, this paper is the first to use fixed and mobile AQ data as well as relevant urban datasets to produce fine-granular, grid-based PM_{2.5} maps for large urban space. This involves identifying relevant factors affecting the emission, dispersion, transport and dissipation of air pollutants via a wide source of urban datasets, including macro features that captures the regional transport of pollutants.

The Deep-MAPS framework is applied to a case study in Beijing with satisfactory accuracy (SMAPE < 15%), which outperforms several benchmarks. The results also show the effectiveness of using local, neighboring and macro features. Furthermore, the results are interpreted to extract insights into the cause and dynamics of PM_{2.5} pollution in Beijing. Specifically, the following findings are made:

- Including macro features improves not only the accuracy but also the interpretability of the inference results.
- Using neighboring features (as opposed to local features) is an effective way to account for the spatial and temporal correlations of urban features and PM_{2.5} data.
- Adding more mobile sensing labels to the training set improves the inference accuracy, but the margin of improvement decreases.
- Deep-MAPS exhibits robustness in dealing with grids with sparse temporal coverage, with limited performance deterioration compared to grids with dense temporal coverage (such as fixed-location stations).
- In the Beijing case study, less than 5% of fixed/mobile sensing coverage (see Figure 2) is sufficient to reach an SMAPE below 15% using the proposed method. This suggests the great potential of Deep-MAPS to reduce hardware investment for high-density AQ monitoring in urban areas.
- For the Beijing case study between 19 June and 16 July 2018, the top three most relevant features of PM_{2.5} concentration are meteorology, POI&AOI, and macro features (i.e. regional transport). The primary paths of regional transport are from Dingling & Changping (Northwest) and Yufa & Liuli River (Southwest).

As a direction for future work, the feature set will be expanded to include more and detailed explanatory variables, such as emission inventory, detailed urban dynamics, and power consumption for factories and buildings, to better represent spatial-temporal variability of emission sources. Moreover, macroscopic indicators, such as satellite images, could be considered to capture background concentrations and regional transport dynamics more accurately.

References

- [1] Health Effects Institute, 2018. State of Global Air report. URL, accessed on 11 Nov 2018.
- [2] Health Effects Institute, 2019. State of global air 2019: Air pollution a significant risk factor worldwide. URL, accessed 24 Oct 2020.
- [3] Lü, J., Liang, L., Feng, Y., Li, R. and Liu, Y., 2015. Air pollution exposure and physical activity in China: current knowledge, public health implications, and future research needs. *International journal of environmental research and public health*, 12(11), pp.14887-14897.
- [4] Chong, C.Y. and Kumar, S.P., 2003. Sensor networks: evolution, opportunities, and challenges. *Proceedings of the IEEE*, 91(8), pp.1247-1256.
- [5] White, R.M., Paprotny, I., Doering, F., Cascio, W.E., Solomon, P.A. and Gundel, L.A., 2012. Sensors and 'apps' for community-based. *EM: Air and Waste Management Association's Magazine for Environmental Managers*, (MAY), pp.36-40.
- [6] Hasenfratz, D., Saukh, O., Walser, C., Hueglin, C., Fierz, M., Arn, T., Beutel, J., Thiele, L., 2015. Deriving high-resolution urban air pollution maps using mobile sensor nodes. *Pervasive and Mobile Computing*, 16, 268-285.
- [7] Kumar, P., Morawska, L., Martani, C., Biskos, G., Neophytou, M., Di Sabatino, S., Bell, M., Norford, L. and Britter, R., 2015. The rise of low-cost sensing for managing air pollution in cities. *Environment international*, 75, pp.199-205.
- [8] Wallace, J., Corr, D., Deluca, P., Kanaroglou, P. and McCarry, B., 2009. Mobile monitoring of air pollution in cities: the case of Hamilton, Ontario, Canada. *Journal of Environmental Monitoring*, 11(5), pp.998-1003.
- [9] Wang, M., Zhu, T., Zheng, J., Zhang, R.Y., Zhang, S.Q., Xie, X.X., Han, Y.Q., Li, Y., 2009. Use of a mobile laboratory to evaluate changes in on-road air pollutants during the Beijing 2008 Summer Olympics. *Atmos. Chem. Phys.* 9, 8247-8263.
- [10] Mead, M.I., Popoola, O.A.M., Stewart, G.B., Landshoff, P., Calleja, M., Hayes, M., Baldovi, J.J., McLeod, M.W., Hodgson, T.F., Dicks, J., Lewis, A., Cohen, J., Baron, R., Saffell, J.R., Jones, R.L., 2013. The use of electrochemical sensors for monitoring urban air quality in low-cost, high-density networks. *Atmospheric Environment* 70, 186-203.
- [11] Maag, B., Zhou, Z. and Thiele, L., 2018. W-Air: Enabling Personal Air Pollution Monitoring on Wearables. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 2(1), p.24.
- [12] Alvear, O., 2018. Mobile Sensing Architecture for Air Pollution Monitoring (Doctoral dissertation), University of Valencia.

- [13] Air Quality Sensor Market: 2019 Global Industry Trends, Growth, Share, Size and 2024 Forecast Research Report.
- [14] The Xinhua News Agency, 2019. Government report on central and local public funding in 2018 and 2019. URL: http://www.gov.cn/xinwen/2019-03/05/content_5371085.htm (accessed on 25 Feb 2020).
- [15] Gibson, M.D., Kundu, S. and Satish, M., 2013. Dispersion model evaluation of PM_{2.5}, NO_x and SO₂ from point and major line sources in Nova Scotia, Canada using AERMOD Gaussian plume air dispersion model. *Atmospheric Pollution Research*, 4(2), pp.157-167.
- [16] Zhang, L., Yin, Y. and Chen, S., 2013. Robust signal timing optimization with environmental concerns. *Transportation Research Part C: Emerging Technologies*, 29, pp.55-71.
- [17] Hertel, O., Berkowicz, R. and Larssen, S., 1991. The operational street pollution model (OSPM). In *Air Pollution Modeling and Its Application VIII* (pp. 741-750). Springer, Boston, MA.
- [18] Ottosen, T.B., Kakosimos, K.E., Johansson, C., Hertel, O., Brandt, J., Skov, H., Berkowicz, R., Ellermann, T., Jensen, S.S. and Ketzel, M., 2015. Analysis of the impact of inhomogeneous emissions in the Operational Street Pollution Model (OSPM). *Geoscientific Model Development*, 8(10), pp.3231-3245.
- [19] Chu, A.K.M., Kwok, R.C.W. and Yu, K.N., 2005. Study of pollution dispersion in urban areas using Computational Fluid Dynamics (CFD) and Geographic Information System (GIS). *Environmental Modelling & Software*, 20(3), pp.273-277.
- [20] Parra, M.A., Santiago, J.L., Martín, F., Martilli, A. and Santamaría, J.M., 2010. A methodology to urban air quality assessment during large time periods of winter using computational fluid dynamic models. *Atmospheric Environment*, 44(17), pp.2089-2097.
- [21] Glatzel, T., Litterst, C., Cupelli, C., Lindemann, T., Moosmann, C., Niekrawietz, R., Streule, W., Zengerle, R. and Koltay, P., 2008. Computational fluid dynamics (CFD) software tools for microfluidic applications-A case study. *Computers & Fluids*, 37(3), pp.218-235.
- [22] Simpson, D., Benedictow, A., Berge, H., Bergström, R., Emberson, L.D., Fagerli, H., Flechard, C.R., Hayman, G.D., Gauss, M., Jonson, J.E. and Jenkin, M.E., 2012. The EMEP MSC-W chemical transport model-technical description. *Atmospheric Chemistry and Physics*, 12(16), pp.7825-7865.
- [23] Chen, Y.H. and Prinn, R.G., 2006. Estimation of atmospheric methane emissions between 1996 and 2001 using a three-dimensional global chemical transport model. *Journal of Geophysical Research: Atmospheres*, 111(D10).
- [24] US Environmental Protection Agency, 2019. CMAQ: The Community Multiscale Air Quality Modeling System. URL, accessed on 17 Apr 2019.
- [25] Saukh, O., Hasenfratz, D., Walser, C., Thiele, L., 2014. On rendezvous in mobile sensing networks. *Lecture Notes in Electrical Engineering*, 281, 29-42.
- [26] Maag, B., Zhou, Z., Saukh, O., Thiele, L., 2017. SCAN: Multi-hop calibration for mobile sensor arrays. In *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, No. 19.
- [27] Zheng, Y., Liu, F. and Hsieh, H.P., 2013, August. U-Air: When urban air quality inference meets big data. In *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 1436-1444). ACM.
- [28] Qi, Z., Wang, T., Song, G., Hu, W., Li, X. and Zhang, Z.M., 2018. Deep air learning: Interpolation, prediction, and feature analysis of fine-grained air quality. *IEEE Transactions on Knowledge and Data Engineering*.
- [29] Zhang, Y., Lv, Q., Gao, D., Shen, S., Dick, R., Hannigan, M., Liu, Q., 2019. Multi-group encoder-decoder networks to fuse heterogeneous data for next-day air quality prediction. In *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence*.
- [30] Marjovi, A., Arfire, A. and Martinoli, A., 2015, June. High resolution air pollution maps in urban environments using mobile sensor networks. In *2015 International Conference on Distributed Computing in Sensor Systems* (pp. 11-20). IEEE.
- [31] Marjovi, A., Arfire, A. and Martinoli, A., 2017. Extending urban air quality maps beyond the coverage of a mobile sensor network: data sources, methods, and performance evaluation. In *Proceedings of the International Conference on Embedded Wireless Systems and Networks* (No. CONF).
- [32] Hu, K., Rahman, A., Bhugubanda, H. and Sivaraman, V., 2017. HazeEst: machine learning based metropolitan air pollution estimation from fixed and mobile sensors. *IEEE Sensors Journal*, 17(11), pp.3517-3525.
- [33] Hankey, S., Marshall, J.D., 2015. Land use regression models of on-road particulate air pollution (particle number, black carbon, pm_{2.5}, particle size) using mobile monitoring. *Environmental Science and Technology*, 49(15), 9194-202.
- [34] Shi, Y., Lau, K. L., Ng, E., 2016. Developing street-level pm_{2.5} and pm₁₀ land use regression models in high-density hong kong with urban morphological factors. *Environmental Science and Technology*, 8178-8187.
- [35] Lim, C. C., Kim, H., Vilcassim, M. J. R., Thurston, G. D., Kim, S. Y., 2019. Mapping urban air quality using mobile sampling with low-cost sensors and machine learning in seoul, south korea. *Environment International*, 131, 105022.
- [36] Apte, J. S., Messier, K. P., Gani, S., Brauer, M., Kirchstetter, T. W., Lunden, M. M., Marshall, J. D., Portier, C. J., Vermeulen, R.C., Hamburg, S.P., 2017. High-resolution air pollution mapping with google street view cars: Exploiting big data. *Environmental Science and Technology*, 51(12), 6999-7008.
- [37] Birant, D. and Kut, A., 2007. ST-DBSCAN: An algorithm for clustering spatial-temporal data. *Data & Knowledge Engineering*, 60(1), pp.208-221.
- [38] Tang, B., Chen, Z., Heffernan, G., Wei, T., He, H. and Yang, Q., 2015, October. A hierarchical distributed fog computing architecture for big data analysis in smart cities. In *Proceedings of the ASE BigData & SocialInformatics 2015* (p. 28). ACM.
- [39] Zhao, Z., Chen, W., Wu, X., Chen, P.C. and Liu, J., 2017. LSTM network: a deep learning approach for short-term traffic forecast. *IET Intelligent Transport Systems*, 11(2), pp.68-75.
- [40] Zhang, J., Zheng, Y. and Qi, D., 2017, February. Deep spatio-temporal residual networks for citywide crowd flows prediction. In *Thirty-First AAAI Conference on Artificial Intelligence*.
- [41] Shi, X., Chen, Z., Wang, H., Yeung, D.Y., Wong, W.K. and Woo, W.C., 2015. Convolutional LSTM network: A machine learning approach for precipitation nowcasting. In *Advances in neural information processing systems* (pp. 802-810).
- [42] Yao, H., Wu, F., Ke, J., Tang, X., Jia, Y., Lu, S., Gong, P., Ye, J. and Li, Z., 2018, April. Deep multi-view spatial-temporal network for taxi demand prediction. In *Thirty-Second AAAI Conference on Artificial Intelligence*.

- [43] The Quanliyan Initiative, 2018. Ministry of Ecological Environment, People's Republic of China. URL: http://www.gov.cn/xinwen/2018-08/26/content_5316749.htm (accessed on 23 Feb 2020).
- [44] Nielsen, M.A., 2015. Neural networks and deep learning (Vol. 25). USA: Determination press.
- [45] Chen, T., Guestrin, C., 2016 XGBoost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pages 785-794.
- [46] He, X., Pan, J., Jin, O., Xu, T., Liu, B., Xu, T., Shi, Y. Atallah, A., Herbrich, R., Bowers, S., Quinonero Candela, J., 2014. In ADKDD'14 Proceedings of the Eighth International Workshop on Data Mining for Online Advertising, pp. 1-9.
- [47] Cheng, Y., Li, X., Li, Z., Jiang, S., Li, Y., Jia, J., Jiang, X.. 2014. AirCloud: a cloud-based air-quality monitoring system for everyone. In Proceedings of the 12th ACM Conference on Embedded Network Sensor Systems. ACM, 251-265.
- [48] Ma, R., Liu, N., Xu, X., Wang, Y., & Zhang, L., 2020. Fine-grained air pollution inference with mobile sensing systems: a weather-related deep autoencoder model. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies.