

RESEARCH ARTICLE

Prediction of daily $PM_{2.5}$ concentration in China using partial differential equations

Yufang Wang^{1,2}*, Haiyan Wang³, Shuhua Chang^{2*}, Adrian Avram³

1 Department of Statistics, Tianjin University of Finance and Economics, Tianjin, China, **2** Coordinated Innovation Center for Computable Modeling in Management Science, Tianjin University of Finance and Economics, Tianjin, China, **3** School of Mathematical and Natural Sciences, Arizona State University, Tempe, United States of America

* These authors contributed equally to this work.

* shuhua55@126.com



OPEN ACCESS

Citation: Wang Y, Wang H, Chang S, Avram A (2018) Prediction of daily $PM_{2.5}$ concentration in China using partial differential equations. PLoS ONE 13(6): e0197666. <https://doi.org/10.1371/journal.pone.0197666>

Editor: Pan-Ping Liu, North University of China, CHINA

Received: January 5, 2018

Accepted: May 7, 2018

Published: June 6, 2018

Copyright: © 2018 Wang et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data is available from figshare: <https://figshare.com/s/ea760256b9850255eed8> (DOI: 10.6084/m9.figshare.6243254).

Funding: This work was supported by the Major Research Plan of the National Natural Science Foundation of China, (91430108) to SHC; the National Basic Research Program (2012CB955804) and the National Natural Science Foundation of China (11171251, 11771322, 11571324) to SHC; the National Science Foundation (DMS-1737861) to HYW; the Major

Abstract

Accurate reporting and forecasting of $PM_{2.5}$ concentration are important for improving public health. In this paper, we propose a partial differential equation (PDE) model, specially, a linear diffusive equation, to describe the spatial-temporal characteristics of $PM_{2.5}$ in order to make short-term prediction. We analyze the temporal and spatial patterns of a real dataset from China's National Environmental Monitoring and validate the PDE-based model in terms of predicting the $PM_{2.5}$ concentration of the next day by the former days' history data. Our experiment results show that the PDE model is able to characterize and predict the process of $PM_{2.5}$ transport. For example, for 300 continuous days of 2016, the average prediction accuracy of the PDE model over all city-regions is 93% or 83% based on different accuracy definitions. To our knowledge, this is the first attempt to use PDE-based model to study $PM_{2.5}$ prediction in both temporal and spatial dimensions.

Introduction

$PM_{2.5}$ (particulate matter smaller than 2.5um), has been linked to deducing visibility, negatively affecting human's health and impacting global climate and therefore $PM_{2.5}$ pollution has aroused unprecedent public concern in the world [1]. In some places where there is critically high atmospheric pollution, government will enforce extreme actions as closing of schools and industries and restriction of motor vehicles circulation. If it were possible to predict these episodes one or two days in advance, public authorities will have more time to take more efficient actions to protect citizens, for instance, by increasing public transports in case of traffic block or by issuing an early warning to inform citizens to take preventing- $PM_{2.5}$ measures.

The originate of $PM_{2.5}$ is complex. Many factors such as area topography, geomorphology, emission source location and weather may have influence on $PM_{2.5}$ concentration [2] and regional transported aerosols are also an important factor for $PM_{2.5}$ pollution [3–5]. Similarly like stock market, $PM_{2.5}$ concentrations are dynamic and exhibit wide variation, and accurate prediction of them thus becomes a highly challenging task because of the highly nonlinear nature and the complex characteristics.

Program of Tianjin University of Finance and Economics (ZD1302) to SHC; and Science Research Project of Tianjin Education Commission (2017SK108) to YFW. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Presently, studies can be divided into three groups: deterministic methods, statistical approaches and satellite remote sensing techniques. Deterministic methods, such as 3-D chemistry transport models (CTMs) [6, 7] mainly address the formation mechanism of $PM_{2.5}$ from the view of physico-chemical and meteorological processes by a knowledge of the temporal dynamics of the emission quantites of various pollutants. This kind of method needs perfect representation of the physico-chemical processes, therefore it is difficult to guarantee the real-time forecast. Statistical approaches, as empirical prediction methods, aim to detect some certain correlated patterns between air quality data and various selected predictors, thus to predict the pollutant concentrations in future. Common statistical approaches, such as linear regression models [8, 9], neural networks [10], nonlinear regression models [11, 12], neuro-fuzzy models, are easier to implement but only limited to specific geographical locations. Satellite remote sensing techniques [13] have the advantages of spatially seamless and long-term coverage, so they have been widely employed to predict $PM_{2.5}$ by considering satellite-derived aerosol optical depth empirically correlated with $PM_{2.5}$ in recent years, but the equipment expense is high. Each approach addresses problems from different perspectives.

Differential equation models have been extensively used in biology, sociology, economics, physics and other fields to describe various time evolution and dynamic systems. Logistic models have been widely used to describe various population dynamics and predict the growth of bacteria and tumors over time in Biology Mathematics and the spatial spread of infectious disease in epidemiology [14]. Reaction-diffusion equations describe population development in [15]. Navier-Stokes equations describe the motion of viscous fluid substances in physics [16]. In particular, leveraging geocoded data dramatically expands application of PDE models to many social problems. Recently, some researchers proposed using PDE models to forecast information diffusion in online social networks by introducing social distance [17–19].

The main contribution of this study is to introduce a PDE model of $PM_{2.5}$ transport which describes the temporal and spatial characteristics of $PM_{2.5}$ in China from a global view. Further we use the model to predict the next day's $PM_{2.5}$ concentrations. The PDE-based model is linear-diffusion equation built on intuitive distance metric between clusters in China. These clusters are obtained by a higher-order network analysis in [20, 21]. This is the first work to apply PDE model to predict $PM_{2.5}$ concentration. As a original work, this method can help Chinese government make prediction about $PM_{2.5}$ concentrations one or two days in advance.

Methods

Data

The research data used in this study covers 300 days from January 1st, 2016 to October 26th, 2016. The average $PM_{2.5}$ daily level of each cluster is calculated from the daily $PM_{2.5}$ level of all cities included in this cluster. All relevant data is available from figshare: <https://figshare.com/s/ea760256b9850255eed8> (DOI: 10.6084/m9.figshare.6243254). The daily $PM_{2.5}$ data can be publicly deposited from this link. Specifically, This $PM_{2.5}$ data set contains 189 priority pollution monitoring cities in China's mainland, which cover all the thirty-four provincial-level regions of China. The most polluted and most interested cities are all included, such as Beijing, Shanghai and Guangzhou. The 189 cities have been divided into 9 clusters by a higher-order complex network method, considering various weather conditions and geography conditions in [20]. The original $PM_{2.5}$ concentrations of each city are normalized to a discrete level value 1, 2, ..., or 6, according to "Ambient Air Quality Standards" (GB3095-1996) of China, where $PM_{2.5}$ concentrations are divided into 0–35, 36–75, 76–115, 116–150, 151–250 and greater than 250 μg and these different concentration ranges are leveled from 1 to 6, describing that air quality is good, mild, moderate, severe, highly severe or seriously severe. The linear

scaling makes concentration value to a specific range, which ensures the larger value input attributes do not overwhelm smaller value inputs thus in turn helps decrease prediction errors.

Clustering and embedding

In our previous work [20], we have divided 189 main cities into 9 clusters by a higher-order spectral method. Specifically, the higher order organization, motif M_8 in Fig 1, reflects the $PM_{2.5}$ movements from source to target in $PM_{2.5}$ -city network. We applied motif M_8 as the basic building block of complex network and applied the higher-order spectral method in [22] to divide 189 main cities of China into 9 clusters in Fig 2. Cities in each cluster are more affected with each other about $PM_{2.5}$ pollution (local emission) and $PM_{2.5}$ of cities among different clusters also possibly influences each other through air flow (regional transport). Therefore in this paper we propose a partial differential equation in spatio-temporal dimension to describe the dynamic system of $PM_{2.5}$ transport. Similarly as information diffusion on social network, for a city cluster, one part of its $PM_{2.5}$ is produced by itself from its road dust, vehicle exhaust, industrial emission, agriculture activities, which can be seen as local emission. The other part of its $PM_{2.5}$ is from other city clusters through regional transport [3–5]. Therefore, the change of $PM_{2.5}$ concentration in a city cluster equals the amount from regional transport plus the $PM_{2.5}$ produced by the city itself (Fig 3).

As we demonstrate in the last paragraph, we could divide 189 cities of China into 9 clusters, i.e., $U = \{U_1, U_2, \dots, U_9\}$ through a higher-order spectral clustering method [20, 22]. Once the entire group of 189 cities are partitioned into disjoint sets U_i , $i = 1, 2, \dots, 9$, we will order U_i in a meaningful way. For general clustering partition, the spatial arrangement of U_i can be based on specific modeling goal and social or geographical characteristics of the underlying network. In [23, 24], the level of democracy, diaspora size, international economic relations and geographical proximity are used to order U_i . In [17], friendship hops are used to define distance metric, then U_x is embed at location x based on that x -axis is used as the social distance. For

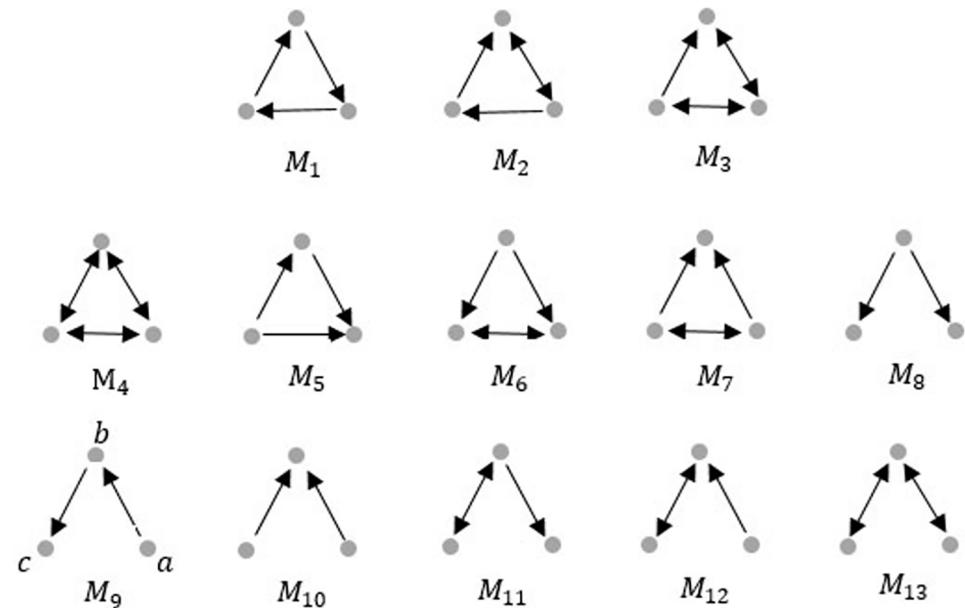


Fig 1. Triangle motifs.

<https://doi.org/10.1371/journal.pone.0197666.g001>

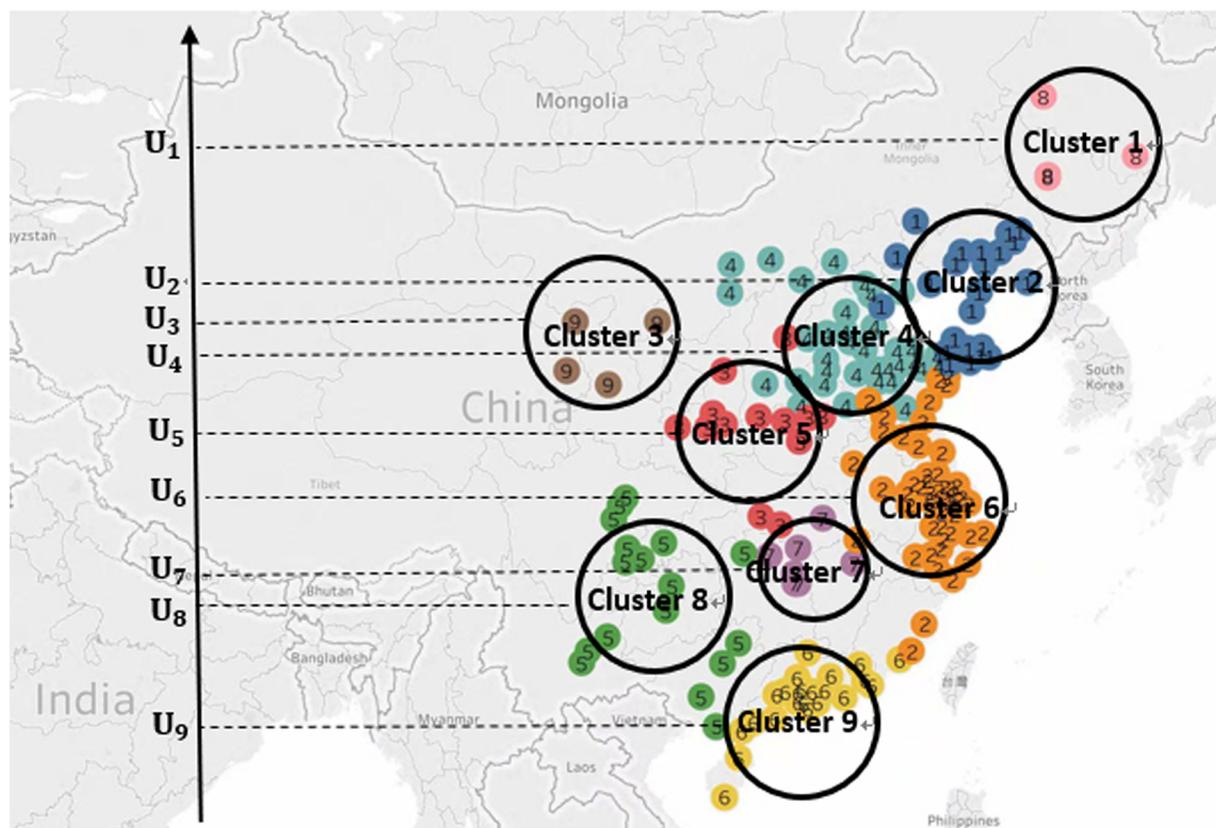


Fig 2. Nine clusters obtained by M_8 -motif spectral clustering algorithm [20]. Tableau Public 10.3 (<https://public.tableau.com/>) was used to create the map.

<https://doi.org/10.1371/journal.pone.0197666.g002>

$PM_{2.5}$ transport in China, ordering cluster U_i by geographical characteristics and meteorological characteristics are the best intuitive method. Wind is a main factor influencing air transport [2, 25, 26] and China is with a monsoon. Specifically, in about September the winter monsoon first reaches northern China and soon moves southly to the Yangtze River and the South China Sea. Summer monsoon prevails to southern China in early May; Then it prevails to the southwestern region in late May or early June; Mid-summer monsoon suddenly prevails into the Yangtze River and Huaihe River Basin in June, mid-summer monsoon again moves to North China and Northeast China in July. Therefore the monsoon of China leads to the fact that most of its wind directions are south-to-north or north-to-south. Therefore, It is meaningful to project the 9 city clusters located from north to south in the map of China onto y-axis of the Cartesian coordinates, and name these city clusters as U_1 to U_9 as their geographical location from north to south as in Fig 2.

These clusters are naturally embedded into the two dimensional Euclidean space. As a result, the PDE model, which describes $PM_{2.5}$ transport between the clusters, would be two-dimensional PDEs. For simplicity, in this paper, we project the clusters into y-axis and focus on one dimensional PDE. We will study two-dimensional PDE models in a future work.

Linear PDE model

Differential equation can describe lots of dynamic systems in fields of physics, mathematical biology, social science and many other fields [27–30], particularly the “information” diffusion

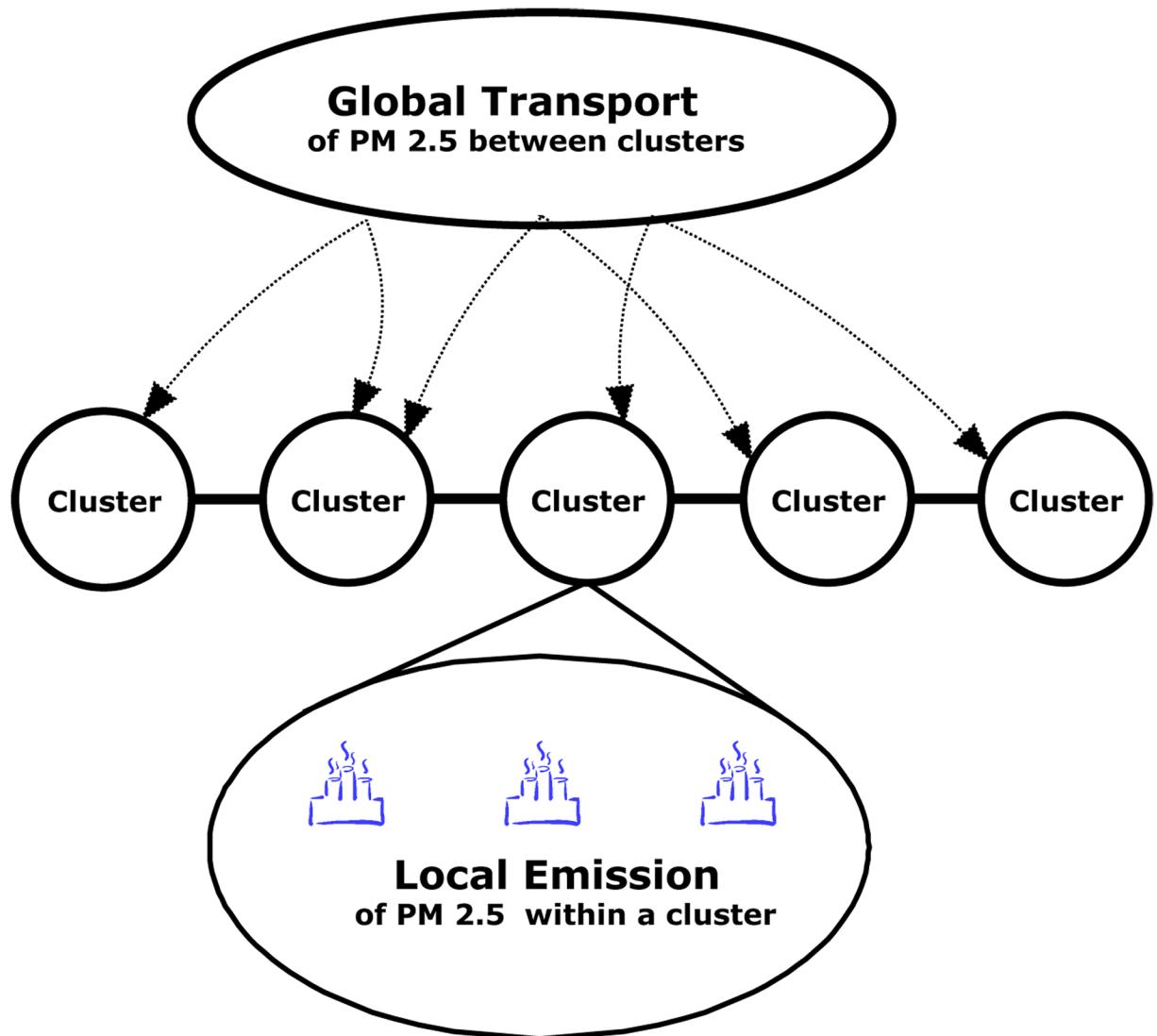


Fig 3. $PM_{2.5}$ transport processes.

<https://doi.org/10.1371/journal.pone.0197666.g003>

problems. In epidemic models, it describes the spread of infectious diseases among a population [29, 31] to help people understand the disease's spreading patterns, such as trends and ratio of people infected. Epidemic models have inspired many unsolved mathematical questions and re-energized novel mathematical researches [27]. In recent century, epidemic models also have been adopted to describe and predict computer virus infections [32, 33]. Recently, some works [17–19] use differential equations to describe information propagation such as news and rumors on social network.

Ordinary differential equation(ODE) and partial differential equation(PDE) are two basic and main classes of models to describe dynamic models. The former is more suitable for studying global patterns, such as trends of people getting infected by epidemics, but cannot describe the connections among individuals, such as who infects whom. However, besides time variable t , partial differential equation has spatial factors involved. Therefore, spatial

dimension provides a quantitative way to describe the local interactions between individuals of the system into global views.

PDE models obey the energy conservation law [34] that is the rate at which a quantity changes in a given domain must equal the rate at which it flows to the boundary plus the rate at which it is created or destroyed in the domain. In different fields, conversation law can describe different scenarios [17–19]. In the following, we propose a linear PDE-based model to abstractly translate the $PM_{2.5}$ transport in two processes: local emission and global transport (in Fig 3).

To quantify the conversation law for $PM_{2.5}$ concentration, let $q(y, t)$ denote the average $PM_{2.5}$ concentration of the cities in the cluster at distance y during time t . The value of $q(y, t)$ depends on two processes. First, the cities in U_y can produce $PM_{2.5}$ by themselves and we call this process as local emission. Secondly, $PM_{2.5}$ in U_x where $x \neq y$ can flow into U_y by wind, namely regional transport [2] and we call this regional transport as global transport.

Following is the description of the PDE model

$$\frac{\partial q(x, t)}{\partial t} = \frac{\partial}{\partial x} \left(d(x) \frac{\partial q(x, t)}{\partial x} \right) + r(t)q(x, t)h(x),$$

$$q(x, 1) = \phi(x), l < x < L,$$

$$\frac{\partial q}{\partial x}(l, t) = \frac{\partial q}{\partial x}(L, t) = 0, t > 1,$$

where

- $\frac{\partial}{\partial x} \left(d(x) \frac{\partial q(x, t)}{\partial x} \right)$ is the global transport, which describes the regional transport of $PM_{2.5}$ between different clusters. This mathematical expression is widely used in spatial biology and epidemiology [30], where it describes the spatial spread of infectious disease.
- $-d(x) \frac{\partial q(x, t)}{\partial x}$ measures the amount of the quantity crossing the section at x and time t and this results from a principle analogous to Fick law [34].
- $d(x)$ is a positive function of x and it describes the $PM_{2.5}$ transport ability of the cluster at location x . Actually, different regions in China have different $PM_{2.5}$ transport abilities, therefore using a positive piece-wise function to describe $d(x)$ will be better, which will be studied in our future work. To be simple, we use $d(x) = e^{-b_0 x}$ in this paper, here b_0 is a parameter which needs to be determined by the real data.
- $r(t)q(x, t)h(x)$ represents the local emission and it denotes the rate at which $PM_{2.5}$ is produced or dissipated within the cluster at location x and time t . The amount of reproduction or dissipation is proportional to both the existing amount of $PM_{2.5}$ and the amount of various varying resources. This mathematical expression has been used to describe and predict various population dynamics such as the growth of bacteria and tumor over time [34].
- $h(x)$ represents the spatial heterogeneity of $PM_{2.5}$ in the cluster at location x and it depicts the $PM_{2.5}$ emissions from various sources such as road dust, vehicle exhaust, biomass burning, industrial emissions of this city cluster. In this paper, different clusters represent different regions in China. Different levels of economy, population of different regions lead to essentially different $PM_{2.5}$ emissions [2]. $h(x)$ is constructed by piecewise first-order Lagrange Interpolation with the interpolated points are $x = 1, 2, \dots, 9$, where each cluster is located, therefore nine references $b_i = h(x_i)$, $i = 1, 2, \dots, 9$ need to be determined.

- $r(t)$ is the growth rate with time t in local process. It depicts $PM_{2.5}$ dissipation with the external changing factors such as wind or certain other atmospheric conditions [2, 25, 26]. Without reproduction, $PM_{2.5}$ will dissipate with wind, therefore the form $r(t)$ can be expressed as $r(t) = \alpha + e^{-\beta(t-\delta)}$, where α , β and δ are constant values. This form of $r(t)$ has also been used in describing information decay in online social network [17]. More plausible expression of $r(t)$ will be discussion in our future work.
- $\frac{\partial q}{\partial x}(l, t) = \frac{\partial q}{\partial x}(L, t) = 0, t > 1$, is the Neumann boundary condition [34]. To be simple, here we suppose there is no flux of $PM_{2.5}$ across the boundaries at $x = l, L$.
- $q(x, 1) = \phi(x)$ is an initial $PM_{2.5}$ concentration function, which can be constructed from the history data of $PM_{2.5}$ by the cubic splines interpolation.

The basic mathematical properties, such as existence, uniqueness and positivity of the solution of the models can be established from the standard theorems for parabolic PDEs [35]. In this paper, the PDE is validated with real data from China National Environmental Monitoring Center.

Accuracy definition

We evaluate the performance of the proposed linear diffusive model by comparing the concentration level calculated by the model with the actual observations. y_t denotes the actual value, and \hat{y}_t is the predicted value. Then the prediction performance is evaluated from two different technical indicators.

- MAID (mean absolute increment accuracy), which is proposed in this paper based on the practical significance, defined as

$$\text{MAID} = \frac{1}{n} \sum_{t=1}^n \text{AID}(t),$$

where

$$\text{AID}(t) = 1 - \frac{|y_t - \hat{y}_t|}{5},$$

n is the number of sample points in test data set and AID evaluates the absolute accuracy at each sample point. There are totally five $PM_{2.5}$ concentration levels from level one to level five and AID describes the absolute accuracy in the view of level length. For example, if $|y_t - \hat{y}_t| = 0.6$, then the prediction accuracy of this sample point is $1 - \frac{0.6}{5} = 0.88$.

- MRA(mean relative accuracy), which is defined as

$$\text{MRA} = \frac{1}{n} \sum_{t=1}^n \left(1 - \frac{|y_t - \hat{y}_t|}{y_t} \right),$$

where n is the number of the sample points in the test data set.

Results

After we finish the work of collecting the real $PM_{2.5}$ concentration data, we conclude the prediction process as follows: Firstly, we normalize the data for decreasing the experimental errors and calculate the average $PM_{2.5}$ daily level of each cluster. Then, we use the training data set of three days to predict the $PM_{2.5}$ concentration for the following day, i.e., we use days 1-3, 2-4,

3-5, . . . as training data, and predict the following day 4, 5, 6, . . . correspondingly and record the prediction accuracy for all 9 regions at the day 4, 5, 6, . . . Specifically, we give the detailed prediction process for day 4 as an instance. For the training data set of day 1, 2, 3, we interpolate the discrete data of day 1 in constructing the initial function $\phi(x)$ of the PDE model; Next, we use the data of day 2 and day 3 to train the parameters of the PDE model by fminsearch function in matlab; Finally, we use this PDE model with a new initial function $\phi(x)$ trained by the data of day 3 to predict the $PM_{2.5}$ concentration of day 4.

Fig 4 illustrates the prediction results of regions 1-9 at each day from January 4, 2016 to October 26, 2016. Clearly to see that there is a consistent trend between predicted data (represented by red lines) and real observations (presented by green lines). Specifically, Figs 5 and 6 show the prediction accuracy of each city region at different time t over 300 days from January 4, 2016 to October 26, 2016. Based on the accuracy definition MAID, in **Fig 5** most of plus marks exist above the horizontal line of 0.9, which means that most of days have a 90% higher prediction value for each cluster and in **Table 1** the mean accuracy values of the 9 clusters are all higher than 91%. In **Fig 6**, the plus marks are not so gathered like that in **Fig 5**, that is because the accuracy definition of MRA is a rule which measures the accuracy of $PM_{2.5}$ concentration value while MAID is a rule which measures the accuracy of $PM_{2.5}$ concentration

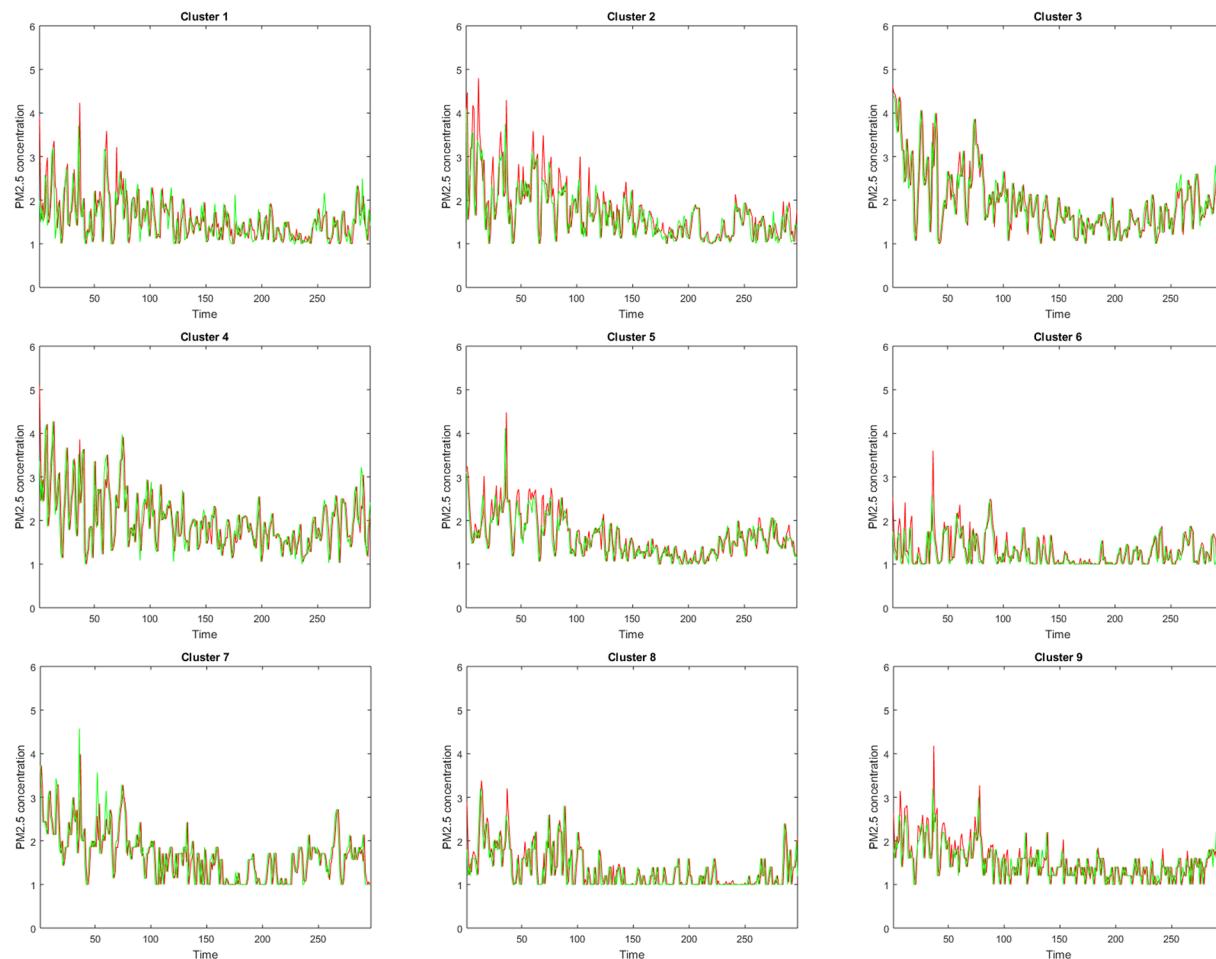


Fig 4. The daily $PM_{2.5}$ concentration level from January 4, 2016 to October 26, 2016. Red line represents the prediction value and green line stands for the real data.

<https://doi.org/10.1371/journal.pone.0197666.g004>

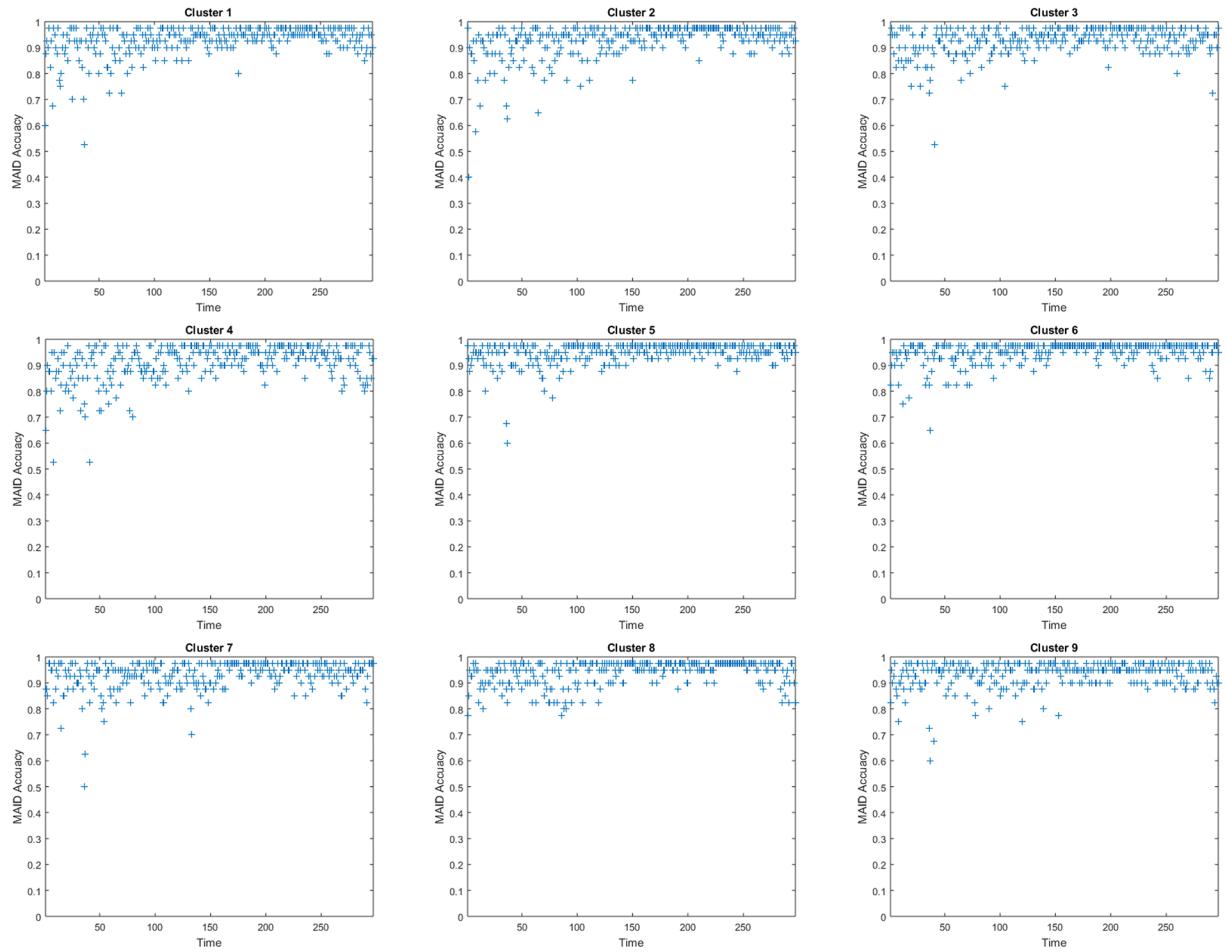


Fig 5. The accuracy of each cluster (city region) from January 4, 2016 to October 26, 2016, based on the accuracy definition of MAID.

<https://doi.org/10.1371/journal.pone.0197666.g005>

value range. In Fig 6, we can still find that most of plus marks exist above the horizontal line of 0.8, which means that most of days have a 80% higher prediction value for each cluster and in Table 1 the mean accuracy values of the 9 clusters are all higher than 80%. Besides, we also find that in Table 1 the average prediction accuracy of the PDE model over all city regions and over all days is 93% and 83% based on the two different accuracy definition respectively.

Specifically, Fig 7 lists the predicted results and the graphs of $h(x)$ and $r(t)$ on April 27, 2016 as an instance. The parameters in $h(x)$ and $r(t)$ on April 27, 2016 are listed in Table 2. The results of other days are omitted due to space limit.

Discussion

In this paper, we propose a PDE-based model to characterize $PM_{2.5}$ transport in 189 main cities of China in order to make short-term prediction of $PM_{2.5}$ concentration. This is an innovative try to use mathematical methods to describe air pollution problem. This work aims to help Chinese government or health organization to forecast the $PM_{2.5}$ concentration one or two days in advance and raise an alarm for specific regions at these moments when there is a high probability that the $PM_{2.5}$ concentration is high.

However, the prediction performance and the PDE model may be improved further in the following respects. Firstly, We can embed the problem into a high dimension space R^2 by

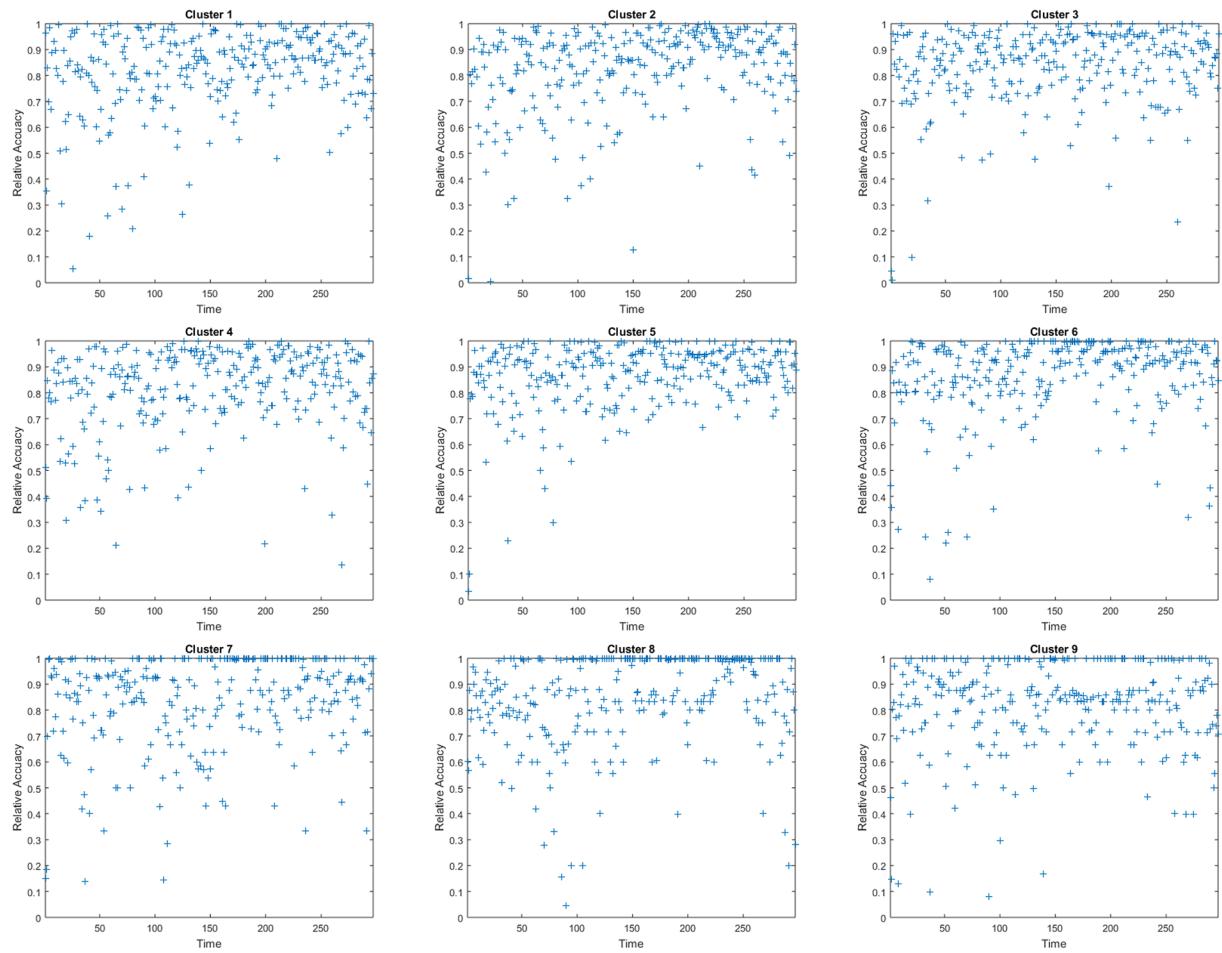


Fig 6. The accuracy of each cluster (city region) from January 4, 2016 to October 26, 2016, based on the accuracy definition of MRA.

<https://doi.org/10.1371/journal.pone.0197666.g006>

Table 1. The average accuracy of different regions and the average accuracy over all regions based on different accuracy definition.

	MAID	MRA		MAID	MRA
Region 1	0.926178451	0.806965014	Region 6	0.948400673	0.858288077
Region 2	0.926262626	0.821313941	Region 7	0.928367003	0.834495023
Region 3	0.924410774	0.823766934	Region 8	0.939478114	0.848018478
Region 4	0.911363636	0.802610969	Region 9	0.930808081	0.809905853
Region 5	0.946212121	0.865824998	Regions 1-9	0.93127572	0.830132143

<https://doi.org/10.1371/journal.pone.0197666.t001>

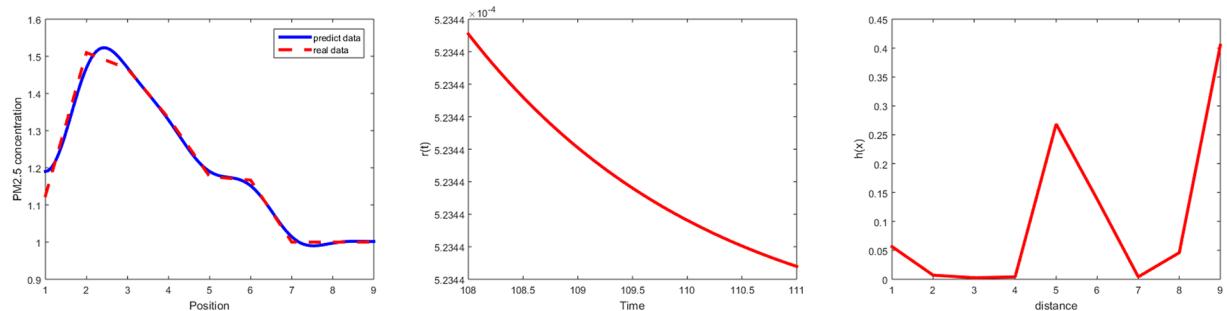


Fig 7. The predicted results of each cluster, $h(x)$ and $r(t)$, of the PDE model on April 27, 2016.

<https://doi.org/10.1371/journal.pone.0197666.g007>

Table 2. The values of these parameters in $h(x)$ and $r(t)$ of the PDE model on April 27, 2016.

Parameter	Value	Parameter	Value	Parameter	Value
b_1	0.031296298	b_5	0.125161344	b_9	0.211384796
b_2	0.430962739	b_6	0.096457084	α	0.000000025
b_3	0.009605521	b_7	0.000040027	β	0.001943744
b_4	0.000081749	b_8	0.054862183	δ	7.245630868

<https://doi.org/10.1371/journal.pone.0197666.t002>

considering a few other variables that may affect the prediction performance. To better describe $PM_{2.5}$ transport, robin boundary condition or others will also be discussed in our future work. A piecewise positive function about the diffusion term $b(x)$ will better describe the reality and we will perfect it work in the following work. Secondly, in this paper, we choose the increment between each cluster to be 1, it can be adjusted according to the specific cases. Certainly, a better increment is based on a more better distance metric of the system. Lastly, we could propose an surveillance system based on the prediction outcomes of this study for practical use.

Author Contributions

Conceptualization: Haiyan Wang, Shuhua Chang.

Data curation: Yufang Wang.

Formal analysis: Yufang Wang, Haiyan Wang.

Funding acquisition: Haiyan Wang, Shuhua Chang.

Methodology: Yufang Wang.

Project administration: Yufang Wang, Haiyan Wang, Shuhua Chang.

Software: Yufang Wang, Adrian Avram.

Supervision: Shuhua Chang.

Validation: Yufang Wang, Shuhua Chang.

Visualization: Yufang Wang.

Writing – original draft: Yufang Wang.

Writing – review & editing: Yufang Wang, Haiyan Wang, Shuhua Chang.

References

1. Lu F, Xu DQ, Cheng YB, Do SX, Guo C, Jiang X, et al. Systematic review and meta-analysis of the adverse health effects of ambient $PM_{2.5}$ and PM_{10} pollution in the Chinese population. Environmental research. 2015; 136:196–204. <https://doi.org/10.1016/j.envres.2014.06.029> PMID: 25460637
2. Chen D, Liu X, Lang J, Zhou Y, Wei L, Wang X, et al. Estimating the contribution of regional transport to $PM_{2.5}$ air pollution in a rural area on the North China Plain. Science of The Total Environment. 2017; 583:280–291. <https://doi.org/10.1016/j.scitotenv.2017.01.066> PMID: 28119005
3. Zheng G, Duan F, Su H, Ma Y, Cheng Y, Zheng B, et al. Exploring the severe winter haze in Beijing: the impact of synoptic weather, regional transport and heterogeneous reactions. Atmospheric Chemistry and Physics. 2015; 15(6):2969–2983. <https://doi.org/10.5194/acp-15-2969-2015>
4. Wang H, Qiao L, Lou S, Zhou M, Ding A, Huang H, et al. Chemical composition of $PM_{2.5}$ and meteorological impact among three years in urban Shanghai, China. Journal of Cleaner Production. 2016; 112:1302–1311. <https://doi.org/10.1016/j.jclepro.2015.04.099>
5. Zhang YL, Cao F. Fine particulate matter ($PM_{2.5}$) in China at a city level. Scientific Reports. 2015; 5.

6. Chuang MT, Zhang Y, Kang D. Application of WRF/Chem-MADRID for real-time air quality forecasting over the Southeastern United States. *Atmospheric Environment*. 2011; 45(34):6241–6250. <https://doi.org/10.1016/j.atmosenv.2011.06.071>
7. Yahya K, Zhang Y, Vukovich JM. Real-time air quality forecasting over the southeastern United States using WRF/Chem-MADRID: Multiple-year assessment and sensitivity studies. *Atmospheric Environment*. 2014; 92:318–338. <https://doi.org/10.1016/j.atmosenv.2014.04.024>
8. Li C, Hsu NC, Tsay SC. A study on the potential applications of satellite data in air quality monitoring and forecasting. *Atmospheric Environment*. 2011; 45(22):3663–3675. <https://doi.org/10.1016/j.atmosenv.2011.04.032>
9. Benas N, Beloconi A, Chrysoulakis N. Estimation of urban PM10 concentration, based on MODIS and MERIS/AATSR synergistic observations. *Atmospheric environment*. 2013; 79:448–454. <https://doi.org/10.1016/j.atmosenv.2013.07.012>
10. Mao X, Shen T, Feng X. Prediction of hourly ground-level PM 2.5 concentrations 3 days in advance using neural networks with satellite data in eastern China. *Atmospheric Pollution Research*. 2017;. <https://doi.org/10.1016/j.apr.2017.04.002>
11. Emili E, Popp C, Petitta M, Riffler M, Wunderle S, Zebisch M. PM 10 remote sensing from geostationary SEVIRI and polar-orbiting MODIS sensors over the complex terrain of the European Alpine region. *Remote sensing of environment*. 2010; 114(11):2485–2499. <https://doi.org/10.1016/j.rse.2010.05.024>
12. Tian J, Chen D. A semi-empirical model for predicting hourly ground-level fine particulate matter (PM 2.5) concentration in southern Ontario from satellite remote sensing and ground-based meteorological measurements. *Remote Sensing of Environment*. 2010; 114(2):221–229. <https://doi.org/10.1016/j.rse.2009.09.011>
13. Ma Z, Hu X, Huang L, Bi J, Liu Y. Estimating ground-level PM2.5 in China using satellite remote sensing. *Environmental science & technology*. 2014; 48(13):7436–7444. <https://doi.org/10.1021/es5009399>
14. Diekmann O, Heesterbeek JAP. Mathematical epidemiology of infectious diseases: model building, analysis and interpretation. vol. 5. John Wiley & Sons; 2000.
15. Cantrell RS, Cosner C. Spatial ecology via reaction-diffusion equations. John Wiley & Sons; 2004.
16. Temam R. Navier-stokes equations. vol. 2. North-Holland Amsterdam; 1984.
17. Wang F, Wang H, Xu K. Diffusive logistic model towards predicting information diffusion in online social networks. In: Distributed Computing Systems Workshops (ICDCSW), 2012 32nd International Conference on. IEEE; 2012. p. 133–139.
18. Wang F, Wang H, Xu K, Wu J, Jia X. Characterizing information diffusion in online social networks with linear diffusive model. In: Distributed Computing Systems (ICDCS), 2013 IEEE 33rd International Conference On. IEEE; 2013. p. 307–316.
19. Wang F, Wang H, Xu K, Raymond R, Chon J, Fuller S, et al. Regional level influenza study with geo-tagged twitter data. *Journal of medical systems*. 2016; 40(8):1–8.
20. Wang Y, Wang H, Chang S, Liu M. Higher-order Network Analysis of Fine Particulate Matter (PM 2.5) Transport in China at City Level. *Scientific Reports*. 2017; 7(1):13236. <https://doi.org/10.1038/s41598-017-13614-7> PMID: 29038572
21. Yan S, Wu G. Network Analysis of Fine Particulate Matter (PM2.5) Emissions in China. *Scientific Reports*. 2016; 6.
22. Benson AR, Gleich DF, Leskovec J. Higher-order organization of complex networks. *Science*. 2016; 353(6295):163–166. <https://doi.org/10.1126/science.aad9029> PMID: 27387949
23. Kwon KH, Xu WW, Wang H, Chon J. Spatiotemporal Diffusion Modeling of Global Mobilization in Social Media: The Case of 2011 Egyptian Revolution. *International Journal of Communication*. 2016; 10:19.
24. Kwon KH, Wang H, Raymond R, Xu WW. A spatiotemporal model of twitter information diffusion: An example of egyptian revolution 2011. In: Proceedings of the 2015 International Conference on Social Media & Society. ACM; 2015. p. 4.
25. Nguyen MV, Park GH, Lee BK. Correlation analysis of size-resolved airborne particulate matter with classified meteorological conditions. *Meteorology and Atmospheric Physics*. 2017; 129(1):35–46. <https://doi.org/10.1007/s00703-016-0456-y>
26. Tian G, Qiao Z, Xu X. Characteristics of particulate matter (PM 10) and its relationship with meteorological factors during 2001–2012 in Beijing. *Environmental pollution*. 2014; 192:266–274. <https://doi.org/10.1016/j.envpol.2014.04.036> PMID: 24857048
27. Brauer F, Castillo-Chavez C, Castillo-Chavez C. Mathematical models in population biology and epidemiology. vol. 40. Springer; 2001.
28. Teschl G. Ordinary differential equations and dynamical systems. vol. 140. American Mathematical Society Providence; 2012.

29. Logan JD. Applied partial differential equations. Springer; 2014.
30. Brauer F. Compartmental models in epidemiology. In: Mathematical epidemiology. Springer; 2008. p. 19–79.
31. Logan JD. Transport modeling in hydrogeochemical systems. vol. 15. Springer Science & Business Media; 2013.
32. Khanh NH, Huy NB. Stability analysis of a computer virus propagation model with antidote in vulnerable system. *Acta Mathematica Scientia*. 2016; 36(1):49–61. [https://doi.org/10.1016/S0252-9602\(15\)30077-1](https://doi.org/10.1016/S0252-9602(15)30077-1)
33. Piqueira JRC, Araujo VO. A modified epidemiological model for computer viruses. *Applied Mathematics and Computation*. 2009; 213(2):355–360. <https://doi.org/10.1016/j.amc.2009.03.023>
34. Nátr L, Murray JD. Mathematical Biology. I. An Introduction. *Photosynthetica*. 2002; 40(3):414–414. <https://doi.org/10.1023/A:1022616217603>
35. Friedman A. Partial differential equations of parabolic type. Courier Dover Publications; 2008.