Environmental Pollution 242 (2018) 922–930

Contents lists available at ScienceDirect

Environmental Pollution

journal homepage: www.elsevier.com/locate/envpol

Leaf magnetic properties as a method for predicting heavy metal concentrations in $PM_{2.5}$ using support vector machine: A case study in Nanjing, China☆

Xiang'zi Leng [a], Xin Qian [a, b], Meng Yang [b, c], Cheng Wang [a], Huiming Li [a, *], Jinhua Wang [a]

[a] State Key Laboratory of Pollution Control and Resources Reuse, School of the Environment, Nanjing University, Nanjing 210023, China
[b] Jiangsu Collaborative Innovation Center of Atmospheric Environment and Equipment Technology (CICAEET), Jiangsu Key Laboratory of Atmospheric Environment Monitoring and Pollution Control (AEMPC), School of Environmental Sciences and Engineering, Nanjing University of Information Science and Technology, Nanjing 210044, China
[c] Jiangsu Key Laboratory of Atmospheric Environment Monitoring and Pollution Control, School of Environmental Science and Engineering, Nanjing University of Information Science & Technology, Nanjing 210044, China

ARTICLE INFO

Article history:
Received 23 October 2017
Received in revised form
2 July 2018
Accepted 2 July 2018
Available online 6 July 2018

Keywords:
Particulate matter
Metal elements
Environmental magnetism
Support vector machine (SVM)
Prediction models

ABSTRACT

The aim of this study was to establish a method for predicting heavy metal concentrations in $PM_{2.5}$ (particulate matter with a diameter of less than 2.5 μm) using support vector machine (SVM) models combined with magnetic properties of leaves. In this study, $PM_{2.5}$ samples and the leaves of three common evergreen tree species were collected simultaneously during four different seasons in Nanjing, China. A SVM algorithm was used to establish models for the prediction of airborne heavy metal concentrations based on leaf magnetic properties, with or without meteorological factors and pollutant concentrations as input variables. Results showed that the annual average $PM_{2.5}$ concentration was 58.47 μg/m^3. $PM_{2.5}$ concentrations, leaf magnetic properties, and nearly all airborne heavy metals had higher concentrations in winter than in spring, summer, or fall. Ferrimagnetic minerals preponderant in dust-loaded leaves were sampled from the three tree species. Models using magnetic properties of leaves from *Ligustrum lucidum Ait* and *Osmanthus fragrans Lour* yielded better prediction effects than those based on the leaves of *Cedar deodara G. Don*, showing relatively higher correlation coefficient (R) values and lower errors in both training and test stages. Fe and Pb concentrations were well-simulated by the prediction models, with R values > 0.7 in both training and test stages. By contrast, the concentrations of V, Co, Sb, Tl, and Zn were relatively poor-simulated, with most R values < 0.7 in both training and test stages. Predictions for the main urban areas of Nanjing showed that the highest heavy metal concentrations occurred near industrial and traffic pollution sources. Our results provide a cost-effective approach for the prediction of airborne heavy metal concentrations based on the biomagnetic monitoring of tree leaves.

1. Introduction

Atmospheric particulate matter (PM) pollution has been becoming one of the most serious environmental challenges affecting humankind in recent decades. Atmospheric heavy metals can induce many adverse health effects (Kampa and Castanas, 2008; Moshammer and Wallner, 2011; Dunea et al., 2016) and impose a long-term burden on biogeochemical cycling in ecosystems (Iii et al., 2002; Dominici et al., 2006; Chen et al., 2013). It is important to determine the pollution levels of atmospheric heavy metals in urban areas for the development of risk mitigation strategies. However, traditional methods based on chemical analysis are time-consuming and expensive, which calls for approaches that are more convenient.

Based on the relationship between elemental content and magnetic properties, the magnetic approach has been suggested as a simple and cost-effective tool to evaluate heavy metal pollution in various environments, including sediments (Zhang et al., 2011), soils (Magiera et al., 2015), street dust (Zhu et al., 2013), tree leaves

☆ This paper has been recommended for acceptance by Dr. Joerg Rinklebe.
* Corresponding author.
E-mail address: valen222@126.com (H. Li).

https://doi.org/10.1016/j.envpol.2018.07.007
0269-7491/© 2018 Elsevier Ltd. All rights reserved.

(Norouzi et al., 2015), and airborne PM (Revuelta et al., 2014). Atmospheric PM is usually deposited on leaf surfaces and trapped in leaf waxes (Dzierzanowski et al., 2011). PM deposited on leaf surfaces typically consist of larger particles (Dzierzanowski et al., 2011), whereas waxes typically trap finer PM (Sgrigna et al., 2015). There are various particle accumulation efficiencies of leaf biomass varies depending on the characteristics of the leaves (wax layer properties, microsurface roughness, and presence of trichomes, among others) (Dzierzanowski et al., 2011; Sæbø et al., 2012), pollutant emissions (distance from source) (Kardel et al., 2011), and meteorological factors (Gautam et al., 2005). Compared with PM filters or street dust, dust-loaded tree leaves that may accumulate PM over many days are easier to collect and can be used to reflect long-term changes in atmospheric particle-bound heavy metals (Matzka and Maher, 1999; Rai, 2013). Indeed, the magnetic properties of tree leaves have been used to evaluate a wide range of atmospheric pollutants, including PM (Mitchell and Maher, 2009; Hofman et al., 2014a), polycyclic aromatic hydrocarbons (Jordanova et al., 2010), and nitrogen dioxide (NO_2) (Brackx et al., 2016). However, the use of leaf magnetic properties as a basis for models that are predicting atmospheric heavy metal levels in PM filters has not been reported, yet.

Prediction models for airborne-pollutant concentrations are either deterministic or statistical. Deterministic models are usually too difficult to be carried out due to the lack of photochemical reaction processes and input data of different types (Díaz-Robles et al., 2008; Koo et al., 2015). By contrast, statistical models are usually simpler with easier access to necessary input variables. Previous reports, have demonstrated that statistical models are commonly used in many ways, such as multiple linear regression (G. S. Kapur et al., 2001), artificial neural networks (Voukantsis et al., 2011; Luna et al., 2014), support vector machine (SVM) (Zhao et al., 2006; Goodarzi et al., 2012; Revuelta et al., 2014), clusterwise regression (Poggi and Portier, 2011), inverse analysis (Koo et al., 2015), and a complementary ensemble empirical mode decomposition (CEEMD)-support vector regression (SVR)-grey wolf optimizer (GWO) hybrid model (Niu et al., 2016). Notably, SVM, developed in the machine learning community, not only takes into account the approximation error but also avoids the problems of overfitting and dimension disaster (Byvatov et al., 2003). It is built on the structural risk minimization (SRM) principle, which seeks to minimize an upper bound of generalization error rather than minimize training error (Revuelta et al., 2014). Thus, the SRM principle has been proved to be very effective for solving nonlinear regression problems (Zhao et al., 2006). SVM-based approaches have been successfully used to identify a range of atmospheric pollutants, including PM (Wang et al., 2015a), sulfur dioxide (SO_2) (Osowski and Garanty, 2007), NO_2 (Osowski and Garanty, 2007), carbonic monoxide (CO) (Yeganeh et al., 2012), ozone (O_3) (Ortiz-García et al., 2010), and airborne heavy metals (Leng et al., 2017; Li et al., 2017).

Evergreen tree leaves are common across many urban areas and are easily collected at high spatial resolutions (Hofman et al., 2017). Therefore, the usage of evergreen tree leaves for the prediction of airborne elemental concentrations may provide a practical alternative to the complex PM sampling methods and chemical analyses of airborne heavy metals. In this study, $PM_{2.5}$ samples and the leaves of three evergreen tree species were collected simultaneously during four different seasons in Nanjing, a typical metropolitan city of China. The magnetic properties of the leaves were characterized while the heavy metal concentrations in $PM_{2.5}$ were measured. The primary aim of this case study was to establish a method for predicting heavy metal concentrations in $PM_{2.5}$ using SVM models with meteorological factors, pollutant concentrations and leaf magnetic properties as inputs, and thus exploiting the demonstrated relationship between heavy metal concentrations and the magnetic properties of leaves. In addition, atmospheric heavy metal concentrations in urban areas of Nanjing were predicted at an optimal spatial resolution based on well-established models.

2. Materials and methods

2.1. Sampling

Nanjing (32°03′N, 118°46′E), the capital of Jiangsu province, is an important industrial production center and the main transportation hub along the Yangtze River Delta in China with a population of more than 8.2 million in 2016. It belongs to the north subtropical monsoon climate zone. $PM_{2.5}$ samples were collected on Whatman quartz microfiber using medium-volume $PM_{2.5}$ samplers (model XY-2200, Qingdao Xuyu Environmental Co., LTD, China) with a flow rate of 100 L/min from the Xianlin campus of Nanjing University, located in the northern suburbs of Nanjing, approximately 5.6 km away from the Nanjing Jinling oil refinery (Fig. S1). Before and after sampling, filters were conditioned in desiccators for 48 h at 25 °C and 40% relative humidity, and weighed using a microbalance (Mettler-Toledo, Greifensee, Switzerland). The $PM_{2.5}$ mass collected on the filters were obtained by subtracting the pre-sampling weights from the post-sampling weights. $PM_{2.5}$ was sampled from: December 4, 2015 to February 28, 2016 (winter), March 2 to May 28, 2016 (spring), June 2 to August 31, 2016 (summer), and September 4 to November 30, 2016 (autumn). A total of 84 $PM_{2.5}$ samples were collected. Continuous sampling lasted 32 h with 8 h (7:00–15:00) per day. Meteorological data (wind speed: 03002-L20, R.M Young, USA; temperature and relative humidity: HMP155A, Fineland; pressure: CS106, USA) and atmospheric gaseous pollutant concentrations (NO_2: 17i, Themo-Fisher, USA; SO_2: 450i, Themo-Fisher, USA; O_3: 49i, Themo-Fisher, USA) were recorded simultaneously at an automatic air quality monitoring station deployed near the $PM_{2.5}$ sampling site.

Synchronous leaf sampling of three species (*Osmanthus fragrans Lour*, *Ligustrum lucidum Ait*, and *Cedar deodara G. Don*) was performed every 4 day at a site 0.3–0.6 km away from the $PM_{2.5}$ sampling site. A total of 252 leaf samples (84 samples × 3 tree species) were collected. The trees that were selected for sampling remained green and healthy throughout the year. They were approximately 3- to 4-year-old and had a height of 2–3 m, readily allowing leaf collection. For each species, two to three trees were selected and three to four leaves were collected from different sides of the tree at a height approximately 1.5–2.0 m above the ground using ceramic scissors to avoid the contamination of the samples with metals. Thus, a total of 6–12 leaves were obtained per sampling day for each kind of tree. The collected leaves were immediately placed in polyethylene bags, labeled, and kept in a refrigerator at 4 °C. They were totally dried in an oven (DGG-9070B, Shanghai Senxin Experimental Instrument Co., LTD, Shanghai, China) for 48 h at 55–60 °C before analysis. Then they were weighed using a microbalance.

For predicting the elemental concentrations in $PM_{2.5}$ by using trained SVM models, leaf collection was conducted in a sample area made up of 25 uniform 5 km × 5 km grids representing the main urban areas of Nanjing (Fig. S2). Overall, the sampling grid for this case study covered an area of about 625 km^2. Two residential areas were selected for most of the grids with three residential areas selected for a small part of the grids when there was no residential area in other nearby grids (Fig. S2). *O. fragrans Lour* was selected based on its spatial distribution simulation and conditions favoring heavy metal concentrations. This tree is one of the most widely distributed trees in Nanjing. It has broadleaf evergreen leaves and is