SCIENTIFIC REPORTS
nature research

OPEN

Biomagnetic monitoring combined with support vector machine: a new opportunity for predicting particle-bound-heavy metals

Qian'ying Dai[1], Mengfan Zhou[1], Huiming Li[2] ✉, Xin Qian[1,3] ✉, Meng Yang[3,4] & Fengying Li[3,4]

Biomagnetic monitoring includes fast and simple methods to estimate airborne heavy metals. Leaves of *Osmanthus fragrans Lour* and *Ligustrum lucidum Ait* were collected simultaneously with PM_{10} from a mega-city of China during one year. Magnetic properties of leaves and metal concentrations in PM_{10} were analyzed. Metal concentrations were estimated using leaf magnetic properties and meteorological factors as input variables in support vector machine (SVM) models. The mean concentrations of many metals were highest in winter and lowest in summer. Hazard index for potentially toxic metals was 5.77, a level considered unsafe. The combined carcinogenic risk was higher than precautionary value (10^{-4}). Ferrimagnetic minerals were dominant magnetic minerals in leaves. Principal component analysis indicated iron & steel industry and soil dust were the common sources for many metals and magnetic minerals on leaves. However, the poor simulation results obtained with multiple linear regression confirmed strong nonlinear relationships between metal concentrations and leaf magnetic properties. SVM models including leaf magnetic variables as inputs yielded better simulation results for all elements. Simulations were promising for Ti, Cd and Zn, whereas relatively poor for Ni. Our study demonstrates the feasibility of prediction of airborne heavy metals based on biomagnetic monitoring of tree leaves.

Atmospheric particulate matter (PM) pollution is a serious global environmental problem and it has been closely linked to a wide range of adverse health outcomes[1–4]. Besides the size and concentration of PM, there is growing evidence that its chemical composition, especially its trace metal components, is crucial to its toxicity[5–7]. The metal particles adsorbed to PM can be transported over large distances and deposited in soils and water bodies or on the leaves of plants via wet and dry deposition, thereby posing environmental risks[8,9]. Consequently, there is an urgent need to investigate the pollution levels caused by particle-bound heavy metals in urban areas.

Tree leaves have large surface areas and thus serve as efficient passive collectors of atmospheric dust[10]. In general, coarse PM is usually deposited on leaf surfaces whereas finer PM is trapped in leaf waxes[11]. With their wide distribution in urban areas and ease of sampling, tree leaves have been used to investigate the spatial and temporal patterns of atmospheric pollutants, including PM[12,13], heavy metals[10,14,15], polycyclic aromatic hydrocarbons[16,17], and nitrogen oxides[18,19]. Geochemical methods for determinations of atmospheric heavy metals involve the collection and processing of PM filter samples. Alternative methods include atomic absorption spectrometry, inductively coupled plasma atomic emission spectroscopy and ICP mass spectrometry, but they are laborious and time-consuming and therefore not suited to large-scale pollution monitoring. However, the magnetic particles adsorbed on tree leaves are a good proxy to measure pollution caused atmospheric heavy metals. Several reports have focused on the relationships between leaf magnetic properties and metal concentrations in leaf samples[10,20–23] or in deposited atmospheric dust[23,24]. However, their methods cannot be used to directly analyze actual

[1]State Key Laboratory of Pollution Control and Resources Reuse, School of the Environment, Nanjing University, Nanjing, 210023, China. [2]School of Environment, Nanjing Normal University, Nanjing, 210023, China. [3]Jiangsu Collaborative Innovation Center of Atmospheric Environment and Equipment Technology (CICAEET), Nanjing University of Information Science & Technology, Nanjing, 210044, China. [4]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. ✉e-mail: valen222@126.com; xqian@nju.edu.cn

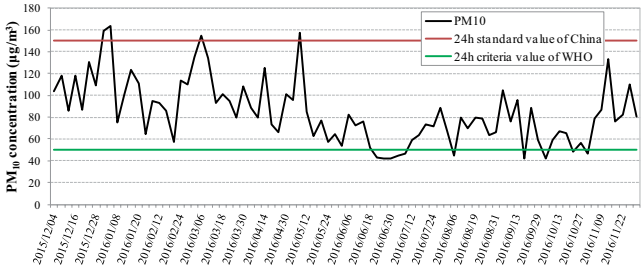

Figure 1. Trend in the 32-hour averaged PM_{10} concentrations ($\mu g/m^3$) during the considered sampling period.

pollution caused by particle-bound heavy metals in the atmosphere, because metal concentrations are normally measured using PM of a certain particle size and collected on pumped-air filters. By contrast, few studies have evaluated the statistical relationships between leaf magnetic properties and heavy metals trapped on pumped-air filters[14].

The levels of atmospheric pollutants can be predicted using deterministic and statistical models. Deterministic models usually require information about pollution sources and emission quantity, as well as sufficient knowledge about the physical processes and chemical reactions among pollutants[25,26]. Statistical models are simpler and better-suited to identifying the dependencies underlying pollutant concentrations and their potential predictors; as such, they often have a higher accuracy[27,28]. Among the statistical approaches used to forecast atmospheric pollutant levels are multiple linear regression[29,30], grey model[31], clusterwise regression[32], random forest partition model[33], artificial neural networks[34,35], support vector machine[36,37] and hybrid models[38,39]. Among those methods, the support vector machine (SVM) algorithm, which is based on the structural risk minimization principle, has increasingly been applied to solve non-linear regression problems, because it takes into account the error approximation in the data and generalization of the models. Previous studies used SVM approaches to predict a series of atmospheric pollutants, including NO_2[40], CO[41] (GarcíaNieto *et al.*, 2013), O_3[42] (Ortiz-García *et al.*, 2010), PM (Cheng *et al.*, 2019) and particle-bound heavy metals[43], based on emission information and meteorological data. However, the potential of statistical models combined with leaf magnetic properties to predict atmospheric heavy metals has yet to be fully explored.

In this study, we examined the relationship between leaf magnetic characteristics and heavy metal concentrations in atmospheric PM_{10} from a large metropolitan city in China. Based on this relationship, we developed a method that uses leaf magnetic properties and meteorological factors as input variables in non-linear statistical models to accurately predict atmospheric particle-bound heavy metals.

Results and Discussion

PM_{10} concentrations. As shown in Fig. 1, approximately 88% of the daily PM_{10} concentrations were higher than the 24-h guideline value of 50 $\mu g/m^3$ proposed by the World Health Organization (WHO), whereas <5% of the daily PM_{10} concentrations were above the 24-h Chinese National Ambient Air Quality Standard (NAAQS) limit of 150 $\mu g/m^3$. The annual mean PM_{10} concentration in the study area was 84 $\mu g/m^3$ (range: 42–164 $\mu g/m^3$), which was slightly higher than the annual limit of 70 $\mu g/m^3$ set by the NAAQS and much higher than the annual guideline value of 20 $\mu g/m^3$ proposed by the WHO. The mean PM_{10} concentration changed seasonally, decreasing in the order of 105 $\mu g/m^3$ (range: 58–164 $\mu g/m^3$) in winter, 97 $\mu g/m^3$ (range: 57–157 $\mu g/m^3$) in spring, 75 $\mu g/m^3$ in autumn (range: 42–133 $\mu g/m^3$), and 64 $\mu g/m^3$ in summer (range: 42–88 $\mu g/m^3$). The temporal trends of the meteorological parameters and atmospheric pollutants during the sampling period are shown in Supplementary Figs. S1 and S2. The concentrations of atmospheric pollutants including PM_{10}, $PM_{2.5}$, SO_2, NO_2 and CO were higher in winter mainly because of the emissions of domestic heating systems and the unfavorable meteorological conditions, such as low wind speed and low temperature, which enhance the accumulation of air pollutants[44]. The lower concentrations of atmospheric pollutants during summer were related with the high temperature, abundant rain and the relatively strong diffusion capacity[45].

Concentrations of particle-bound-heavy metals. The mean concentrations of the particulate-bound elements are summarized in Table 1. In general, Fe and Zn were the most abundant heavy metals in PM_{10}, whereas Co and Cd were present at lower concentrations. The mean concentrations of Cr, Cu, Fe, Mn, Pb and Zn were highest in winter, those of Cd, Ni and V were highest in autumn, and those of As, Co and Ti were highest in spring. The mean concentrations of most of the measured elements (except As and Co) were lowest in summer.

In assessments of heavy metal contaminations, the enrichment factor (EF), calculated by normalizing a tested element against a conservative reference element, is commonly used to distinguish between anthropogenic influences and natural background levels[46,47]. The calculation and classification of the EF as applied in this study can be found in the Supplementary Information. As shown in Fig. 2, there were no obvious seasonal difference in the EFs of the different elements. The mean EFs of Co, Fe, Mn, and V were <10, indicative of a minimal enrichment of these metals and their having originated mainly from crustal sources. Cr was moderately enriched ($10 < EF < 100$), whereas As, Cd, Cu, Ni, Pb and Zn ($EF > 100$) were anomalously enriched. Thus, all of these elements were likely to have derived from anthropogenic sources, such as steel smelting, fly ash from coal burning, vehicle emissions, waste incineration, and contaminated soil[48].

	Spring	Summer	Autumn	Winter	Average
As	8.54 ± 4.43	8.00 ± 3.87	6.43 ± 2.49	7.22 ± 3.73	7.57 ± 3.73
Cd	1.73 ± 0.77	1.52 ± 0.63	3.51 ± 1.04	2.17 ± 1.08	2.20 ± 1.17
Co	0.58 ± 0.23	0.49 ± 0.24	0.40 ± 0.20	0.48 ± 0.28	0.48 ± 0.25
Cr	38.4 ± 17.7	29.4 ± 10.8	39.0 ± 13.1	49.9 ± 14.1	38.8 ± 15.7
Cu	39.3 ± 18.4	29.9 ± 8.29	35.8 ± 16.5	46.8 ± 12.4	37.7 ± 15.4
Fe	1043 ± 453	689 ± 225	786 ± 420	1064 ± 284	891 ± 382
Mn	52.3 ± 26.1	35.9 ± 9.59	56.7 ± 16.3	62.0 ± 19.0	51.1 ± 20.8
Ni	54.8 ± 18.9	52.8 ± 21.2	82.1 ± 27.4	68.1 ± 23.9	63.8 ± 24.8
Pb	38.8 ± 15.5	30.9 ± 9.56	45.0 ± 14.0	62.0 ± 20.5	43.7 ± 18.8
Ti	52.1 ± 25.1	37.5 ± 16.4	41.7 ± 17.2	51.4 ± 16.4	45.5 ± 19.8
V	4.85 ± 1.83	4.55 ± 1.41	6.56 ± 2.77	4.74 ± 1.28	5.15 ± 2.14
Zn	296 ± 135	247 ± 125	355 ± 111	399 ± 102	321 ± 131

Table 1. Heavy metal concentrations in PM_{10} samples during the four seasons (ng/m^3).

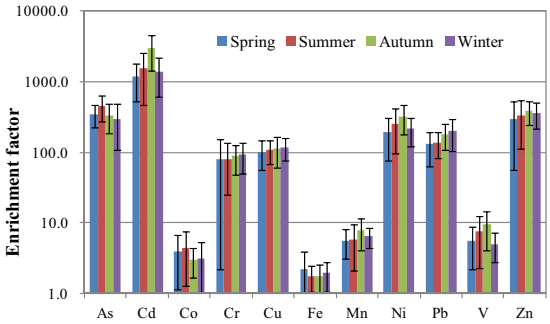

Figure 2. Enrichment factors of heavy metals in PM_{10} during the four seasons.

Health impacts based on guideline values and risk assessment. Comparisons of the elemental concentrations in PM_{10} with the limits imposed by the NAAQS (GB3095–2012) and WHO are shown in Supplementary Fig. S3. The Mn, Pb and V concentrations in all PM_{10} samples were far below the NAAQS (GB3095–2012) and WHO limits. The Cd concentration was lower than the NAAQS (GB3095–2012) and WHO limits with exception of three PM_{10} samples collected in autumn, in which the concentration was slightly higher. By contrast, the As concentration in 64.3% and 53.6% of the PM_{10} samples exceeded the NAAQS limit of 6 ng/m^3 and 6.6 ng/m^3, respectively. The Ni concentration in nearly all of the PM_{10} samples was above the WHO limit of 25 ng/m^3.

The health risk caused by exposure to the analyzed airborne metals via inhalation is shown in Supplementary Table S1. Among the studied metals, Ni, Pb and As had higher EC values. The HQ values for the inhalation of As, Cd, Co, Cr, Mn and V were below the safe limit (1) for both children and adults. Ni had the highest HQ value (3.83) in these two subpopulations. The hazard index (HI) for these metals was 5.77, which was above the safe limit (1), indicating accumulative noncarcinogenic risks for adults and children. For carcinogens, the acceptable risk range is between 1×10^{-6} and 1×10^{-4} according to the US EPA's risk management policy. The carcinogenic risks of As, Cd, Co, Ni, and Pb inhaled from PM_{10} were less than the precautionary value (10^{-4}), both for children and adults, but the carcinogenic risks of Cr for the two subpopulations were higher. The combined carcinogenic risk was 2.39×10^{-4} for children and 9.55×10^{-4} for adults. Both values were higher than the precautionary value. Thus, for every one million children and one million adults living in the local environment, approximately 3 children and 10 adults are at risk of developing cancer during their lifetime due to exposure to toxic metals via PM_{10} inhalation.

Leaf magnetic properties. Both χ_{LF} and SIRM generally reflect the quantities of magnetic, and especially ferromagnetic (e.g., magnetite) minerals in a sample, but they are also influenced by the presence of paramagnetic and diamagnetic minerals[49]. χ_{ARM} is particularly sensitive to single-domain ferrimagnetic grains[50]. As shown in Table 2, in the leaves of both tree species χ_{LF} and SIRM decreased seasonally, in the order of winter > spring > autumn > summer. However, there were differences in the seasonal pattern of χ_{ARM}, which decreased in the order of winter > autumn > spring> summer for *Osmanthus fragrans Lour*, and in the order summer > winter > autumn > spring for *Ligustrum lucidum Ait*. As shown in Fig. 3, the SIRM and χLF values correlated linearly, suggesting that ferrimagnetic minerals were the dominant magnetic minerals in the leaf samples[48]. The lack of a significant correlation between χ_{LF} and χ_{ARM} (Fig. 3) indicated that single-domain grains did not dominate these ferrimagnetic minerals. The ratios of χ_{ARM} to χ_{LF}, χ_{ARM} to SIRM, and SIRM to χ_{LF} are used to estimate mineral magnetic grain-size variations, with increasing ratios indicating decreasing grain size[48,50,51]. For both tree species,

Property	*Osmanthus fragrans Lour*					*Ligustrum lucidum Ait*				
	Spring	Summer	Autumn	Winter	Average	Spring	Summer	Autumn	Winter	Average
χ_{LF} (10^{-8} m^3/kg)	2.63 ± 0.93	0.88 ± 0.31	1.63 ± 0.88	3.27 ± 0.82	2.07 ± 1.19	2.36 ± 0.74	1.46 ± 0.46	1.57 ± 0.63	3.08 ± 0.80	2.10 ± 0.93
χ_{ARM} (10^{-8} m^3/kg)	1.41 ± 1.49	1.06 ± 0.35	2.11 ± 1.36	2.13 ± 1.86	1.65 ± 1.41	0.68 ± 0.31	1.64 ± 1.69	1.55 ± 0.85	1.56 ± 1.84	1.36 ± 1.37
SIRM (10^{-6} Am^2/kg)	289 ± 108	184 ± 32.1	248 ± 51.8	390 ± 54.1	274 ± 100	304 ± 60.5	275 ± 41.0	333 ± 38.4	431 ± 46.7	333 ± 75.0
χ_{ARM}/χ_{LF}	0.59 ± 0.67	1.51 ± 1.41	1.82 ± 2.00	0.66 ± 0.56	1.15 ± 1.38	0.30 ± 0.12	1.20 ± 0.96	1.11 ± 0.69	0.49 ± 0.45	0.78 ± 0.75
χ_{ARM}/SIRM (10^{-5}m/A)	5.24 ± 4.84	5.83 ± 1.92	9.12 ± 7.04	5.33 ± 4.32	6.35 ± 4.97	2.28 ± 1.16	6.16 ± 6.78	4.75 ± 2.79	3.48 ± 3.63	4.22 ± 4.42
SIRM/χ_{LF} (10^2A/m)	111 ± 29.1	240 ± 124	179 ± 63.0	127 ± 35.8	166 ± 90.3	136 ± 30.5	209 ± 81.8	247 ± 102	149 ± 43.1	186 ± 82.5

Table 2. Magnetic properties of the collected leaf samples of *Osmanthus fragrans Lour* and *Ligustrum lucidum Ait*.

	Method	Input variables
Model I	SVM	PM_{10} + meteorological factors
Model II	MLR	PM_{10} + meteorological factors + magnetic parameters of *Osmanthus fragrans Lour*
Model III	MLR	PM_{10} + meteorological factors + magnetic parameters of *Ligustrum lucidum Ait*
Model IV	SVM	PM_{10} + meteorological factors + magnetic parameters of *Osmanthus fragrans Lour*
Model V	SVM	PM_{10} + meteorological factors + magnetic parameters of *Ligustrum lucidum Ait*

Table 3. The method and input variables of the five developed models.

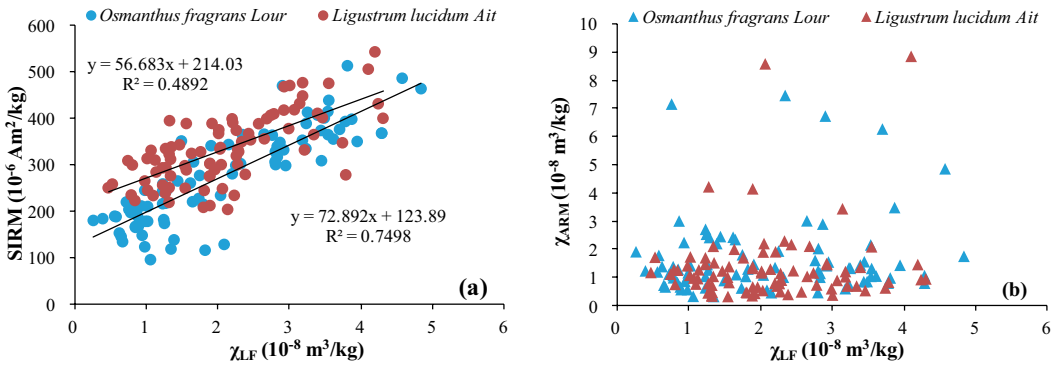

Figure 3. Scatter plots of (**a**) χ_{LF} vs. SIRM and (**b**) χ_{LF} vs. χ_{ARM} in the leaves of *Osmanthus fragrans Lour* and *Ligustrum lucidum Ait*.

all three ratios were lowest in spring and higher in autumn and summer, which suggested leaf accumulation of a larger number of finer grains in the latter two seasons.

The SIRM values and the ratio of SIRM to χ_{LF} were significantly higher in *Ligustrum lucidum Ait* than in *Osmanthus fragrans Lour*, whereas the ratios of χ_{ARM} to χ_{LF} and χ_{ARM} to SIRM were significantly lower in *Ligustrum lucidum Ait* than in *Osmanthus fragrans Lour*. The annual mean values of χ_{LF} and SIRM for *Osmanthus fragrans Lour* were $2.07 \pm 1.19 \times 10^{-8}$ m^3/kg and $274 \pm 100 \times 10^{-6}$ Am^2/kg, respectively, and for *Ligustrum lucidum Ait* $2.10 \pm 0.93 \times 10^{-8}$ m^3/kg and $333 \pm 75.0 \times 10^{-6}$ Am^2/kg, respectively (Table 3). According to a review of Hofman *et al.*[52] and a previous study by our group[14], the SIRM values of the two tree species were in the middle range while the χ_{LF} values were lower than published values. Several different factors determine leaf magnetic properties. Particle accumulation by leaves is influenced by species-specific characteristics of the trees, such as phenology, growth status, leaf area density and leaf characteristics, e.g., wax layer properties, surface roughness and trichomes presence[53]. However, sampling height, the leaf exposure period, PM source distance and strength, as well as meteorological conditions, e.g., wind, rain, drought and seasonal dynamics, also play important roles[52]. Further work is needed to reveal the underlying relationships among meteorological conditions, airborne heavy metals from various sources and leaf properties.

Principal component analysis (PCA). The relationships among heavy metals, PM_{10}, meteorological factors and magnetic parameters were analyzed in a PCA (Supplementary Tables S2 and S3). For the leaf magnetic parameters of *Osmanthus fragrans Lour*, five factors, accounting for 76.041% of the total variance, were obtained. The first factor, accounting for 35.272% of the total variance, was dominated by Cr, Cu, Fe, Mn, Pb, Ti, Zn, PM_{10}, temperature, pressure, χ_{LF} and SIRM, which indicated that the metal sources were the iron and steel industry and soil dust. Factor 2, accounting for 15.619% of the total variance, was dominated by Cd, Ni, V, wind speed, χ_{ARM}, χ_{ARM}/χ_{LF}, χ_{ARM}/SIRM, and SIRM/χ_{LF}, indicating industrial activities that resulted in the release of magnetic

minerals of a certain grain size. Factor 3 explained 11.765% of the total variance and was dominated by As and Co. Arsenic is a typical element associated with coal combustion[54], whereas Co, with EF < 10, is indicative of crustal source. Therefore, this factor may reflect mixed sources of coal combustion and natural process. Factor 4, accounting for 8.998% of the total variance, was dominated by relative humidity, whereas factor 5, dominated by Co and Cu and representing 4.386% of the total variance, indicated traffic activities and road dust as the major sources[55].

When the leaf magnetic parameters of *Ligustrum lucidum Ait* were included in the PCA, four similar factors, accounting for 71.397% of the total variance, were obtained, with the first, second, third, and fourth components explaining 33.612%, 16.626%, 12.013%, 9.147% of the variance, respectively. The four factors were dominated by Cr, Cu, Fe, Mn, Pb, Ti, Zn, PM_{10}, temperature, relative humidity, pressure, χ_{LF} and SIRM (component 1); Cd, Ni, V, wind speed, χ_{ARM}/χ_{LF} and SIRM/χ_{LF} (component 2); As and Co (component 3) and χ_{ARM} and χ_{ARM}/SIRM (component 4).

Simulation results and implications. The linkage of atmospheric heavy metals and magnetic particles is based on the fact that heavy metals such as Zn, Cd, Pb and Cr can be incorporated into the particle structure during combustion processes and/or by subsequent surface adsorption[56,57]. Magnetic properties can thus act as an effective proxy for airborne heavy metals. The influence of leaf magnetic properties on heavy metal accumulation was examined in this study by predicting metal concentrations with and without leaf magnetic variables while including PM_{10} concentrations and meteorological factors (Table 3). The R, MAE and RMSE results are listed in Supplementary Tables S4–S7. The predicted vs. observed concentration and the residuals were plotted for Pb and are shown in Fig. 4.

When only the PM_{10} concentration and meteorological factors served as SVM inputs, the training R value of all the studied elements was between 0.565 and 0.819, and the test R value between 0.528 and 0.816. The training R and test R values of Cu were the lowest (0.565 and 0.528, respectively), and those of Ti the highest (0.819 and 0.816, respectively). The training R and test R values of Co, Cr, Fe, Mn and Ni were between 0.6 and 0.7, and those of As, Cd, Pb, V and Zn between 0.7 and 0.8.

However, the simulation results of the stepwise MLR were not satisfactory, even when leaf magnetic variables were included as input variables (Supplementary Table S5). For *Osmanthus fragrans Lour*, the training R values of the metals were between 0.587 (Mn) and 0.780 (Pb); for the test R, the values were <0.6, except in the case of Zn with test R of 0.693. For *Ligustrum lucidum Ait*, the training R values of all of metals were between 0.494 (Mn) and 0.681 (Pb); for the test R, the values were <0.6, with the exceptions of Cd and Co. These results obtained using a linear approach demonstrated the strong nonlinear relationships between metal concentrations and the input variables, which is consistent with our previous findings[30,43].

When the leaf magnetic variables of *Osmanthus fragrans Lour* were included in the SVM model, the training R and test R values of all the elements were in the range of 0.693–0.918 and 0.667–0.903, respectively. The training and test R values of Cd, Cu, Ti and Zn were >0.8, with the highest values being those of Ti (0.918 and 0.903, respectively). Both the training and the test R values of As, Co, Cr, Fe, Mn, Pb and V were between 0.7 and 0.8 whereas Ni had the lowest values (0.693 and 0.667, respectively). The addition of the leaf magnetic variables of *Ligustrum lucidum Ait* into the SVM model yielded training R and test R values for all metals in the range of 0.661–0.875 and 0.630 to 0.859, respectively. The training and test R values of Cd, Ti, V and Zn were >0.8, with Ti again having the highest value (0.875 and 0.859, respectively). Both the training and the test R values of As, Co, Cu, Fe, Mn and Pb were between 0.7 and 0.8, but Cr and Ni had lower training and test R values (0.6–0.7). Thus, when the leaf magnetic variables were added as inputs, the training and the test R values of all the elements increased to varying degrees. The lower MAE and RMSE values of most of the metals (except Cr and Pb in model V) obtained in the training stage demonstrated the improved accuracy of model IV and model V when the leaf magnetic properties were included. In the test stage, the MAE and RMSE values of As, Cu, Mn, Pb, Ti, V and Zn of model IV, as well as As, Cd, Fe, Pb, Ti and V of model V were lower than the corresponding values of model I.

The improvement in the models achieved by including leaf magnetic properties as inputs was quantified by calculating the improvement rates (IRs) of model IV and model V for the training R and test R of each metal. The IR was calculated as follows[43]:

$$IR = (R_{\text{model IV/V}} - R_{\text{model I}})/R_{\text{model I}} \quad (1)$$

As shown in Fig. 5, in the training stage, the IRs of Co, Cu and Mn were better whereas those of As, Ni and Pb were relatively poor, both for model IV and model V. In the test stage, the IR values of Cu, Co and Cr of model IV were higher, as were those of Cu, Co, Zn and Fe of model V (both models: >0.15). By contrast, the IR values of Ni and V of model IV and those of Mn, Ni, Cr and Pb of model V were lower (both models: <0.05). In general, Cu and Co, both of which were attributed in the PCA to traffic activities and road dust, improved by the inclusion of the leaf magnetic properties of *Osmanthus fragrans Lour* and *Ligustrum lucidum Ait*. In our previous report[43], the IRs of Co, Cu, Fe, Mn and Zn were also better with the inclusion of magnetic variables of $PM_{2.5}$ when simulating the mass-related concentrations of heavy metals in $PM_{2.5}$ by using SVM models.

Model IV, in which the leaf magnetic properties of *Osmanthus fragrans Lour* were included as inputs, showed better simulation effects for As, Cd, Cr, Cu, Mn, Ni, Pb and Ti. For these metals, the training and test R values of this model were higher than those of model V (*Ligustrum lucidum Ait*), whereas the simulation of Co, Fe, V and Zn was better in model V. This result was mainly related to the morphological and physiological characteristics of the two tree species (e.g., leaf area density, wax layer properties, surface roughness)[58] but the slight differences in the ambient environments (e.g., soil dust, road traffic, buildings) were also likely to have played a role[59].

In general, models IV and V performed best for Ti, Cd and Zn, as evidenced by training R and test R values > 0.8. For Ni, however, the performances of these models were relatively poor, based on training R and test R

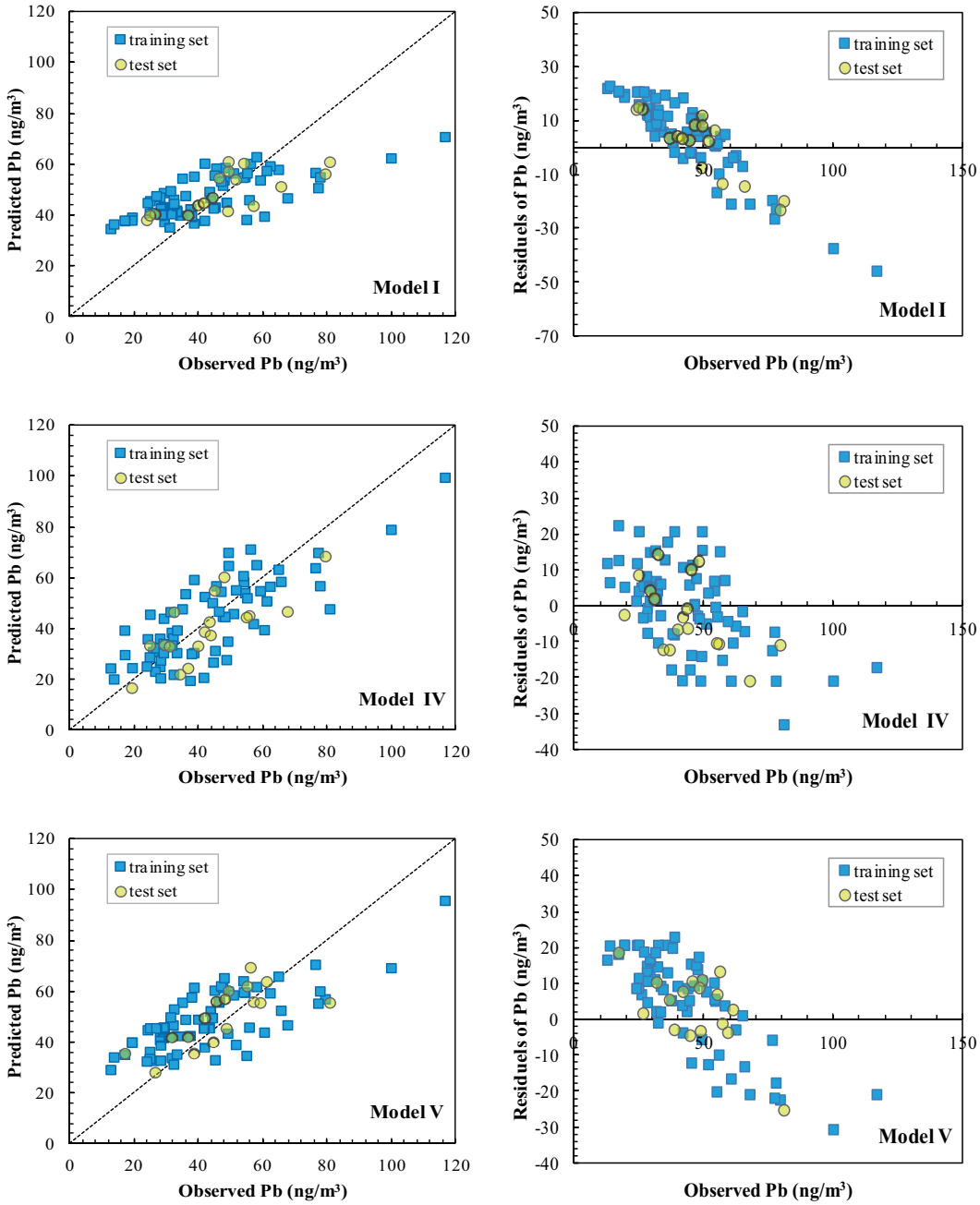

Figure 4. Predicted vs. observed concentrations and residuals plots of Pb for the training and test stages as described by models I, IV and V.

values < 0.7. Ti was identified as a crustal element, whereas Ni, with the highest noncarcinogenic health risk and originating from mixed industrial activities, had a relatively poor simulation and very little improvement afforded by the inclusion of leaf magnetic properties. This finding is consistent with previous reports of a more reliable linkage between heavy metal concentrations and magnetic parameters in environments with similar and/or "single" source contributions, whereas multiple sources of heterogeneous chemical and magnetic particle can complicate determinations of the relationships between atmospheric heavy metals and magnetic parameters[52,60]. Source-specific magnetic fingerprints and their associations with atmospheric heavy metals remain to be elucidated in further research.

Conclusions

The linkage between heavy metals in PM_{10} and the leaves of *Osmanthus fragrans Lour* and *Ligustrum lucidum AitLigustrum lucidum Aitc* were studied using SVM models. The annual mean PM_{10} concentration was 84 μg/m^3 (range: 42–164 μg/m^3). The elements As, Cd, Cu, Ni, Pb and Zn were anomalously enriched and Cr was moderately enriched whereas Co, Fe, Mn, and V were mainly from crustal sources. Ni had the highest noncarcinogenic

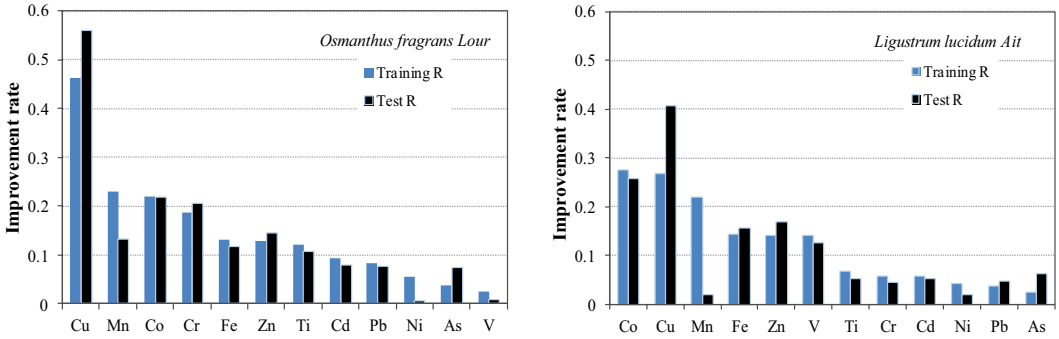

Figure 5. Improvement rate obtained by comparison of models IV (*Osmanthus fragrans Lour*) and I, and models V (*Ligustrum lucidum Ait*) and I.

risk, and Cr the highest carcinogenic risks. The combined noncarcinogenic and carcinogenic risks posed by inhalation exposure to airborne heavy metals were both above the safe limit or precautionary level. The χ_{LF} and SIRM values of the leaves of both tree species decreased in the order of winter > spring > autumn > summer, and the χ_{ARM} values in the order of winter > autumn > spring > summer. The dominant magnetic minerals in the leaf samples were ferrimagnetic minerals. PCA revealed that the heavy metals in PM_{10} have common sources with the magnetic minerals in leaf samples.

A subset of PM_{10} concentrations, meteorological factors and leaf magnetic properties were then used as input variables to simulate heavy metals concentrations. The poor simulation results obtained by MLR evidenced the nonlinear relationships between the airborne metal concentrations and the input variables. The inclusion of leaf magnetic variables improved the simulation results for all of the studied elements, with the largest improvements in Cu and Co and the lowest improvement in Ni. SVM models with leaf magnetic variables of the two tree species as inputs performed better for Ti, Cd and Zn but relatively poorly for Ni. Our study thus demonstrates that the concentrations of most airborne toxic heavy metals can be estimated using a simple and efficient biomagnetic diagnostic method.

Methods

Sampling. Nanjing (118°46′E, 32°03′N), the second largest city in the Yangtze River Delta region of China, is an important industrial production area and the main transportation hub in southeastern China. It has a north subtropical monsoon climate with a mean annual temperature of 16 °C and a mean annual precipitation of 1106 mm. PM_{10} samples were collected on Whatman quartz microfibers using medium-volume PM 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 (Supplementary Fig. S4), located in the northern suburbs of Nanjing and near the city's northern industrial districts. Continuous sampling of PM_{10} lasted 32 h with 8 h (7:00 am-15:00 pm) per day was conducted, 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). To ensure the samples were representative, we avoided sampling during rainy or windy weather. A total of 84 PM_{10} samples were collected. Meteorological data were recorded synchronously at an automatic air quality monitoring station located near the study site. Before and after sampling, the filters were conditioned for 48 h in a desiccator at 25 °C and 40% relative humidity, and then weighed to determine PM_{10} mass.

Osmanthus fragrans Lour and *Ligustrum lucidum Ait*, two species of evergreen trees widely distributed in Nanjing, were selected for leaf sampling because the hairs on their leaf surfaces facilitate the adsorption of atmospheric particles. Leaf samples of the two trees were collected every 4 day during the same duration as PM sampling to keep consistence with the 32 h-sampled PM filters, from a site ~400 m away from the PM sampling site (Fig. S4). Specifically, each leaf sample was collected on the forth day of PM sampling to ensure the necessary accumulation of PM on leaves. The 4-day sampling duration both of PM_{10} and tree leaf was chosen by considering the effects of meteorological conditions and the daily variation of PM_{10} concentration from Nanjing. The distance between the two tree species was <100 m. For each tree species, two healthy trees next to each other were selected from which four leaves were collected from each one using ceramic scissors. The trees used for sampling in this study were all 3- to 4-year-old with a height of 2–3 m. The leaves were obtained from different sides of the tree and at a height from the ground of 1.5–2.0 m. The eight leaves were pooled to obtain one leaf sample, with 84 leaf samples collected in total for each tree species. The leaf samples were immediately placed in polyethylene bags and kept in a refrigerator at 4 °C. All leaf samples were totally dried in an oven at 55–60 °C before the analysis.

Magnetic measurements. Low frequency (0.875 kHz; χ_{LF}) was measured on about 2 g dried leaf sample using a KLY-3S kappa bridge (Agico, Czech Republic). Isothermal Remanent Magnetization (IRM) was induced in a field of 1000 mT (SIRM) using a Molspin pulse magnetizer. The anhysteretic remanent magnetization (ARM) was realized using a DTECH AF demagnetizer (Molspin, UK), delivering 0.04 mT of direct current (DC) and a peak alternating field of 100 mT. Measurements are then expressed as susceptibility of ARM (χ_{ARM}) by dividing the remanence by the DC bias field. For the purpose of quantitation, all magnetic properties of each leaf sample were normalized on mass-specific basis.

Analysis of heavy metal concentrations. Metal elements were released from the PM_{10} samples by digestion with a mixture of HNO_3, HCl and HF. The concentrations of Fe and Zn were determined using inductively coupled plasma optical emission spectrometry (Perkin Elmer SCIEX, Optima 5300 DV, Norway). The concentrations of As, Cd, Co, Cr, Cu, Mn, Ni, Pb, Ti, V and Zn were determined by using inductively coupled plasma mass spectrometry (Perkin Elmer SCIEX, Elan 9000, Norway). Four blank filters were also digested and measured for metal concentrations simultaneously. Then the concentration of one element was corrected by subtracting its average concentration of blank filters. SRM 1649a (urban particulate matter) was used for quality assurance and control with the recovery of all the studied elements between 88% and 109%.

Health risk assessment. The carcinogenic and noncarcinogenic risks posed by potentially toxic metals through the direct inhalation of PM_{10} were calculated using the human health risk assessment models of the US Environmental Protection Agency[61,62]. The models include exposure assessment and risk characterization. Sensitivity was determined for children and adults. The inhalation exposure concentration (EC), hazard quotient (HQ) of the noncarcinogenic risk, and the carcinogenic risk (CR) were calculated as described in the Supporting Information. The hazard index (HI) is equal to the sum of the HQ and was used to assess the overall potential of noncarcinogenic effects.

Simulation models. Although initially developed for classification problems[63], SVM models have been extended to solve nonlinear regression estimations by the introduction of an ε-insensitive loss function. Detailed information on SVM theory is provided in several publications[36,37,64] and is thus described only briefly in the following:

Firstly, a kernel function is used to map the input variables to a high-dimensional feature space, after which the SVM approximates a set of data with a linear function:

$$y = f(x, \omega) = \sum_{i=1}^{m} \omega_i \varnothing(x_i) + b \tag{2}$$

where $\varnothing(x_i)$ is the features of the input variables after their kernel transformation, and ω_i and b are the coefficients estimated by minimizing the regularized risk function. After kernel transformation, the data are linearly separable in the new feature space. In this study, the Gaussian radial basis function kernel was applied:

$$k(x_i, x_j) = \exp(-\gamma \cdot ||x_i - x_j||^2) \tag{3}$$

where γ is the parameter of the kernel, and xi and xj are two independent variables.

The coefficients ω_i and b are estimated by minimizing the regularized risk function:

$$R = C\frac{1}{N}\sum_{i=1}^{N} L_\varepsilon(y_i, f(x_i, \omega)) + \frac{1}{2}||\omega||^2 \tag{4}$$

where the term $C\frac{1}{N}\sum_{i=1}^{N} L_\varepsilon(y_i, f(x_i, \omega))$ is the empirical error (risk), measured using the ε-insensitive loss function:

$$L_\varepsilon(y_i, f(x_i, \omega)) = \begin{cases} 0, & if\ |y - f(x, \omega)| \le \varepsilon \\ |y - f(x, \omega)| - \varepsilon, & \text{otherwise} \end{cases} \tag{5}$$

where ε is a prescribed parameter called the regularized term and is defined as the approximation accuracy of the training data points. The loss function ignores errors when their value is less than that of ε. The term $1/2||\omega||^2$ is the regularization term, which serves as a measure of function flatness. The value of the regularized constant C determines the trade-off between empirical error and the regularization term. Finally, the dual problem of Eq. (4) is often resolved by the introduction of the Lagrange multiplier method:

$$f(x) = \sum_{i=1}^{N} (\alpha_i - \alpha_i^*) K(x, x_i) \tag{6}$$

where α_i and α_i^* are the introduced Lagrange multipliers.

In this study, MATLAB R2013a and libsvm-3.21 were used to build the SVM models. The data were randomly partitioned into two sets: 80% for training and 20% for testing. The maximum and minimum concentrations of one target element observed during the sampling period were retained in the training set to develop a reliable model. A subset of PM_{10} concentrations, meteorological factors (temperature, relative humidity, pressure, and wind speed), with or without leaf magnetic properties (χ_{LF}, χ_{ARM}, SIRM, χ_{ARM}/χ_{LF}, χ_{ARM}/SIRM and SIRM/χ_{LF}), were used as the input variables. Among the successful models, the best model was selected based on the higher correlation coefficient of the observed versus predicted output and on fewer errors in the training and test stages.

Models using a stepwise multiple linear regression (MLR) were also established to simulate metal concentrations with the same independent variables used in the SVM models, by applying SPSS 23.0. As shown in Table 3, five models were developed according to the statistical methods and input variables.

Evaluation of model performance. The correlation coefficient (R) of the observed vs. predicted concentration of each heavy metal was used to measure the fit performance of each model. The mean absolute

error (MAE) and root mean squared error (MSE), which provide a global estimate of the difference between the observed and predicted outputs, were used to measure residual errors. In general, a higher R combined with a lower MAE and RMSE was considered to indicate better modeling of a metal element. R, MAE and RMSE were calculated as described in the Supplementary Information.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Received: 17 February 2020; Accepted: 5 May 2020;
Published online: 25 May 2020

References

1. Brook, R. D. *et al*. Particulate matter air pollution and cardiovascular disease: an update to the scientific statement from the American Heart Association. *Circulation.* **121**, 2331–2378 (2010).
2. Calderón-Garcidueñas, L. *et al*. Prefrontal white matter pathology in air pollution exposed Mexico city young urbanites and their potential impact on neurovascular unit dysfunction and the development of Alzheimer's disease. *Environ. Res.* **146**, 404–417 (2016).
3. Choi, J. *et al*. Harmful impact of air pollution on severe acute exacerbation of chronic obstructive pulmonary disease: particulate matter is hazardous. *Int. J. Chron Obstr. Pulm. Dis.* **13**, 1053–1059 (2018).
4. Jia, X. F. *et al*. Effects of fine particulate on heart rate variability in Beijing: a panel study of healthy elderly subjects. *Int. Arch. Occup. Environ. Health*. **85**, 97–107 (2012).
5. Bollati, V. *et al*. Exposure to metal-rich particulate matter modifies the expression of candidate microRNAs in peripheral blood leukocytes. *Environ. Health Perspect.* **118**, 763–768 (2010).
6. Uski, O. *et al*. Effect of fuel zinc content on toxicological responses of particula bte matter from pellet combustion *in vitro*. *Sci. Total Environ.* **511**, 331–340 (2015).
7. Yuan, Y. *et al*. In vitrotoxicity evaluation of heavy metals in urban air particulate matteron human lung epithelial cells. *Sci. Total Environ.* **678**, 301–308 (2019).
8. McConnell, J. R. & Edwards, R. Coal burning leaves toxic heavy metal legacy in the Arctic. *P. Natl. Acad. Sci.* **105**, 12140–12144 (2008).
9. Tagliani, S. M. *et al*. Content, mineral allocation and leaching behavior of heavy metals in urban $PM_{2.5}$. *Atmos. Environ.* **153**, 47–60 (2017).
10. Norouzi, S., Khademi, H., Cano, A. F. & Acosta, J. A. Biomagnetic monitoring of heavy metals contamination in deposited atmospheric dust, a case study from Isfahan, Iran. *J. Environ. Manage.* **173**, 55–64 (2016).
11. Dzierzanowski, K., Popek, R., Gawronska, H., Saebø, A. & Gawronski, S. W. Deposition of particulate matter of different size fractions on leaf surfaces and in waxes of urban forest species. *Int. J. Phytoremediation.* **13**, 1037–1046 (2011).
12. Hofman, J., Wuyts, K., Wittenberghe, S. V. & Samson, R. On the temporal variation of leaf magnetic parameters: Seasonal accumulation of leaf-deposited and leaf-encapsulated particles of a roadside tree crown. *Sci. Total Environ.* **493**, 766–772 (2014).
13. Kardel, F., Wuyts, K., Maher, B. A. & Samson, R. Intra-urban spatial variation of magnetic particles: monitoring via leaf saturation isothermal remanent magnetisation (SIRM). *Atmos. Environ.* **55**, 111–120 (2012).
14. Leng, X. Z. *et al*. 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. *Environ. Pollut.* **242**, 922–930 (2018).
15. Mahera, B. A., Mooreb, C. & Matzkac, J. Spatial variation in vehicle-derived metal pollution identified by magnetic and elemental analysis of roadside tree leaves. *Atmos. Environ.* **42**, 364–373 (2008).
16. Jordanova, D. *et al*. Magnetic signature of different vegetation species in polluted environment. *Stud. Geophys. Geod.* **54**, 417–442 (2010).
17. Lehndorff, E. & Schwark, L. Biomonitoring of air quality in the Cologne Conurbation using pine needles as a passive sampler—Part II: polycyclic aromatic hydrocarbons (PAH). *Atmos. Environ.* **38**, 3793–3808 (2004).
18. Hofman, J. *et al*. Increasing the spatial resolution of air quality assessments in urban areas: A comparison of biomagnetic monitoring and urban scale modeling. *Atmos. Environ.* **92**, 130–140 (2014).
19. McIntosh, G., Gómez-Paccard, M. & Osete, M. L. The magnetic properties of particles deposited on Platanus x hispanica leaves in Madrid, Spain, and their temporal and spatial variations. *Sci. Total Environ.* **382**, 135–146 (2007).
20. Cao, L. W. *et al*. Magnetic response to air pollution recorded by soil and dust-loaded leaves in a changing industrial environment. *Atmos. Environ.* **119**, 304–313 (2015).
21. Castanheiro, A. *et al*. Leaf accumulation of atmospheric dust: Biomagnetic, morphological and elemental evaluation using SEM, ED-XRF and HR-ICP-MS. *Atmos. Environ.* **221**, 117082 (2020).
22. Kardel, F., Wuyts, K., De Wael, K. & Samson, R. Biomonitoring of atmospheric particulate pollution via chemical composition and magnetic properties of roadside tree leaves. *Environ. Sci. Pollu. Res.* **25**, 25994–26004 (2018).
23. Castanheiro, A., Samson, R. & DeWael, K. Magnetic- and particle-based techniques to investigatemetal deposition on urban green. *Sci. Total Environ.* **571**, 594–602 (2016).
24. Wuyts, K. *et al*. A new opportunity for biomagnetic monitoring of particulate pollution in an urban environment using tree branches. *Atmos. Environ.* **190**, 177–187 (2018).
25. Baklanov, A. *et al*. Towards improving the simulation of meteorological fields in urban areas through updated/advanced surface fluxes description. *Atmos. Chem. Phys.* **8**, 523–543 (2008).
26. Wu, Q. L. & Lin, H. X. A novel optimal-hybrid model for daily air quality index prediction considering air pollutant factors. *Sci. Total Environ.* **683**, 808–821 (2019).
27. Feng, R. *et al*. Recurrent Neural Network and random forest for analysis and accurate forecast of atmospheric pollutants: A case study in Hangzhou, China. *Journal. Clean. Product.* **231**, 1005–1015 (2019).
28. Nieto, P. G., Combarro, E. F., Del Coz Díaz, J. J. & Montañés, E. A SVM-based regression model to study the air quality at local scale in Oviedo urban area (Northern Spain): a case study. *Appl. Math. Comput.* **219**, 8923–8937 (2013).
29. Li, C., Hsu, N. C. & Tsay, S. A study on the potential applications of satellite data in air quality monitoring and forecasting. *Atmos. Environ.* **45**, 3663–3675 (2011).
30. Leng, X. Z. *et al*. Prediction of size-fractionated airborne particle-bound metals using MLR, BP-ANN and SVM analyses. *Chemosphere.* **180**, 513–522 (2017).
31. Li, J.J., Gong, D.P. & Liu, X.R. Prediction and analysis of air pollutants concentrations in Wuwei City of Gansu Province based on GM(1,1). Model. *Environ. Sci. Manage*. (In Chinese) **(37)**, 65–71 (2012).
32. Poggi, J. M. & Portier, B. PM_{10} forecasting using clusterwise regression. *Atmos. Environ.* **45**, 7005–7014 (2011).
33. Kamińska, J. A. A random forest partition model for predicting NO_2 concentrations from traffic flow and meteorological conditions. *Sci. Total Environ.* **651**, 475–483 (2019).

34. Luna, A. S. *et al*. Prediction of ozone concentration in tropospheric levels using artificial neural networks and support vector machine at Rio de Janeiro, Brazil. *Atmos. Environ.* **98**, 98–104 (2014).
35. Liu, Z. J. *et al*. Exploring the potential relationship between indoor air quality and the concentration of airborne culturable fungi: a combined experimental and neural network modeling study. *Environ. Sci. Pollut. Res.* **25**, 3510–3517 (2018).
36. Suárez Sánchez, A., García Nieto, P. J., Riesgo Fernández, P., del Coz Díaz, J. J. & Iglesias-Rodríguez, F. J. Application of an SVM-based regression model to the air quality study at local scale in the Avilés urban area (Spain). *Math. Comput. Model.* **54**, 1453–1466 (2011).
37. Suleimana, A., Tightb, M. R. & Quinnb, A. D. Applying machine learning methods in managing urban concentrations of traffic-related particulate matter (PM_{10} and $PM_{2.5}$). *Atmos. Pollut. Res.* **10**, 134–144 (2019).
38. Polat, K. A novel data preprocessing method to estimate the air pollution (SO_2): Neighbor-based feature scaling (NBFS). *Neural. Comput. Appl.* **21**, 1987–1994 (2012).
39. Liu, H. *et al*. An intelligent hybrid model for air pollutant concentrations forecasting: Case of Beijing in China. *Sustain. Cities. Soc.* **47**, 101471 (2019).
40. Lu, W. Z. & Wang, W. J. Potential assessment of the "support vector machine"method in forecasting ambient air pollutant trends. *Chemosphere.* **59**, 693–701 (2005).
41. García Nieto, P. J., Combarro, E. F., del Coz Díaz, J. J. & Montañés, E. A SVM-based regression model to study the air quality at local scale in Oviedo urban area (Northern Spain): A case study. *Appl. Mathe. Compute.* **219**, 8923–8937 (2013).
42. Ortiz-García, E. G., Salcedo-Sanz, S., Pérez-Bellido, A. M., Portilla-Figueras, J. A. & Prieto, L. Prediction of hourly O_3 concentrations using support vector regression algorithms. *Atmos. Environ.* **44**, 4481–4488 (2010).
43. Li, H. M. *et al*. Magnetic properties as a proxy for predicting fine-particle-bound heavy metals in a support vector machine approach. *Environ. Sci. Tech.* **51**, 6927–6935 (2017).
44. Kang, H. Q. *et al*. Analysis of a longlasting haze episode in Nanjing, China. *Atmos. Res.* **120-121**, 78–87 (2014).
45. Li, H. M. *et al*. Fractionation of airborne particulate-bound elements in haze-fog episode and associated health risks in a megacity of southeast China. *Environ. Pollut.* **208**, 655–662 (2016).
46. Gao, J. J. *et al*. Seasonal and spatial variation of trace elements in multi-size airborne particulate matters of Beijing, China: Mass concentration, enrichment characteristics, source apportionment, chemical speciation and bioavailability. *Atmos. Environ.* **99**, 257–265 (2014).
47. Schleicher, N. J. *et al*. Temporal variability of trace metal mobility of urban particulate matter from Beijing-A contribution to health impact assessments of aerosols. *Atmos. Environ.* **45**, 7248–7265 (2011).
48. Oldfield, F. Environmental magnetism – a personal perspective. *Quater. Sci. Rev.* **10**, 73–85 (1991).
49. Thompson, R. & Oldfield, F. *Environ. Magnetism.* (Allen and Unwin, London, 1986).
50. Maher, B. A. Magnetic properties of some synthetic sub-micron magnetites. *Geophys. J. Int.* **94**, 83–96 (1988).
51. Banerjee, S. K. J, K. & J. M. A rapid method for magnetic granulometry with applications to environmental studies. *Geophys. Res. Lett.* **8**, 333–336 (1981).
52. Hofman, J. *et al*. Biomagnetic monitoring of atmospheric pollution: a review of magnetic signatures from biological sensors. *Environ. Sci. Technol.* **51**, 6648–6664 (2017).
53. Castanheiro, A. *et al*. Leaf accumulation of atmospheric dust: Biomagnetic, morphological and elemental evaluation using SEM, ED-XRF and HR-ICP-MS. *Atmos. Environ.* **221**, 117082 (2020).
54. Mooibroek, D., Schaap, M., Weijers, E. P. & Hoogerbrugge, R. Source apportionment and spatial variability of $PM_{2.5}$ using measurements at five sites in the Netherlands. *Atmos. Environ.* **45**, 4180–4191 (2011).
55. Duan, J. C. & Tan, J. H. Atmospheric heavy metals and Arsenic in China: Situation, sources and control policies. *Atmos. Environ.* **74**, 93–101 (2013).
56. Lu, S. G., Yu, X. L. & Chen, Y. Y. Magnetic properties, microstructure and mineralogical phases of technogenic magnetic particles (TMPs) in urban soils: Their source identification and environmental implications. *Sci. Total Environ.* **543**, 239–247 (2016).
57. Moreno, E., Sagnotti, L., Dinarès-Turell, J., Winkler, A. & Cascella, A. Biomonitoring of traffic air pollution in Rome using magnetic properties of tree leaves. *Atmos. Environ.* **37**, 2967–2977 (2003).
58. Chen, L., Liu, C., Zhang, L., Zou, R. & Zhang, Z. Variation in tree species ability to capture and retain airborne fine particulate matter ($PM_{2.5}$). *Sci. Rep.* **7**, 3206 (2017).
59. Janhäll, S. Review on urban vegetation and particle air pollution–deposition and dispersion. *Atmos. Environ.* **105**, 130–137 (2015).
60. Li, H. M., Qian, X., Wei, H. T. & Zhang, R. B. Magnetic properties as proxy for the evaluation of heavy metal contamination in urban street dusts of Nanjing, Southeast China. *Geophys. J. Int.* **199**, 1354–1366 (2014).
61. US EPA (U.S. Environmental Protection Agency). Risk Assessment Guidance for Super Fund Volume I Human Health Evaluation Manual (Part a). EPA/540/1-89/002, http://www.epa.gov/swerrims/riskassessment/ragsa/index.htm (1989).
62. US EPA (U.S. Environmental Protection Agency).Risk Assessment Guidance for Superfund (RAGS), Volume I Human Health Evaluation Manual (Part F, Supplemental Guidance for Inhalation Risk Assessment). EPA-540-R-070e002, OSWER9285.7-82, http://www.epa.gov/swerrims/riskassessment/ragsf/index.htm (2009).
63. Cortes, C. & Vapnik, V. Support-Vector Networks. *Mach. Learn.* **20**, 273–297 (1995).
64. Juhos, I., Makra, L. & Balazs, T. Forecasting of traffic origin NO and NO_2 concentrations by Support Vector Machines and neural networks using Principal Component. *Analysis. Simul. Model. Pract. Th.* **16**, 1488–1502 (2008).

Acknowledgements

This work was supported by the National Natural Science Foundation of China (grant no. 41771533), Natural Science Foundation of Jiangsu Province, China (grant no. BK20171339) and Open Fund of State Key Laboratory of Pollution Control and Resources Reuse (grant no. PCRRF19026).

Author contributions

H.M.L. and X.Q. contributed to the overall concept design of this work. Q.Y.D. and M.F.Z. performed the experiments and wrote the manuscript. M.Y. and F.Y.L. participated in sampling and data analysis. All authors reviewed the paper.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41598-020-65677-8.

Correspondence and requests for materials should be addressed to H.L. or X.Q.