

SCIENTIFIC REPORTS



OPEN

Temporal Patterns in Fine Particulate Matter Time Series in Beijing: A Calendar View

Jianzheng Liu^{1,2}, Jie Li¹ & Weifeng Li^{1,2}

Received: 26 January 2016

Accepted: 04 August 2016

Published: 26 August 2016

Extremely high fine particulate matter (PM_{2.5}) concentration has become synonymous to Beijing, the capital of China, posing critical challenges to its sustainable development and leading to major public health concerns. In order to formulate mitigation measures and policies, knowledge on PM_{2.5} variation patterns should be obtained. While previous studies are limited either because of availability of data, or because of problematic a priori assumptions that PM_{2.5} concentration follows subjective seasonal, monthly, or weekly patterns, our study aims to reveal the data on a daily basis through visualization rather than imposing subjective periodic patterns upon the data. To achieve this, we conduct two time-series cluster analyses on full-year PM_{2.5} data in Beijing in 2014, and provide an innovative calendar visualization of PM_{2.5} measurements throughout the year. Insights from the analysis on temporal variation of PM_{2.5} concentration show that there are three diurnal patterns and no weekly patterns; seasonal patterns exist but they do not follow a strict temporal division. These findings advance current understanding on temporal patterns in PM_{2.5} data and offer a different perspective which can help with policy formulation on PM_{2.5} mitigation.

Beijing, the capital of China where more than 20 million people reside, would probably never have considered it would gain the title “Capital of Smog” that was used for London 60 years ago. But now the title seems to fit Beijing appropriately. Clearly, air pollution not only undermines the reputation of Beijing as a historic world-renowned city, but more importantly it poses citizens and the government with a critical challenge for the sustainable development of urbanization that involve major public health concerns. Of all the most common detrimental air pollutants, fine particulate matter (PM_{2.5}) is believed to be the most serious pollutant due to its harmful health impact on the cardiovascular, respiratory, and pulmonary functionality in humans^{1,2}. There are increasing evidences that PM_{2.5} is associated with population mortality^{3,4}, cardiovascular and respiratory diseases mortality⁵, and has adverse impacts on growth of new-borns⁶, and even on mental health and can cause anxiety⁷.

Actions need to be taken to mitigate PM_{2.5} problems in Beijing as well as in other cities of China. For this, we should first measure PM_{2.5} and analyse it to determine inherent variation patterns. Until now, a handle of research efforts have been made to this end^{8–16}. While these studies report interesting results about PM_{2.5}, they are defected in several ways. Some research are limited because of data availability such as a limited number of PM_{2.5} monitoring stations^{8,11}, or that data is only available for a limited period^{12–14}. These studies fail to provide a sufficient overview of PM_{2.5} concentration patterns across the city of Beijing and through a full year. There have been studies that analysed PM_{2.5} measurements data in a full year across Beijing that were provided by the newly launched air pollution monitoring network since later 2012^{9,10,15,16}, but these studies, when analysing the temporal variation of PM_{2.5} concentration, used an a priori assumption that PM_{2.5} concentration follows seasonal, monthly, or weekly patterns. The reasoning in these studies is that since the PM_{2.5} concentration probably follows seasonal, monthly, or weekly patterns, the analysis framework could be based on an imposed seasonal, monthly, or weekly profile analysis. We argue that the variation of PM_{2.5} concentration may vary on different time scales other than these predefined scales, and studies using these predefined time scales are likely to provide incomplete information, and therefore miss important insights.

Our study, instead of making arbitrary assumptions on weekly, monthly, and seasonal patterns, prefers to reveal the data. Using a full year of PM_{2.5} ground-level measurements from January 2014 to December 2014 in Beijing, our study conducted two time-series clustering analyses for all the daily PM_{2.5} measurements. In this way,

¹Department of Urban Planning and Design, Faculty of Architecture, The University of Hong Kong, Hong Kong.

²Shenzhen Institute of Research and Innovation, The University of Hong Kong, Shenzhen, China. Correspondence and requests for materials should be addressed to W.L. (email: wfli@hku.hk)

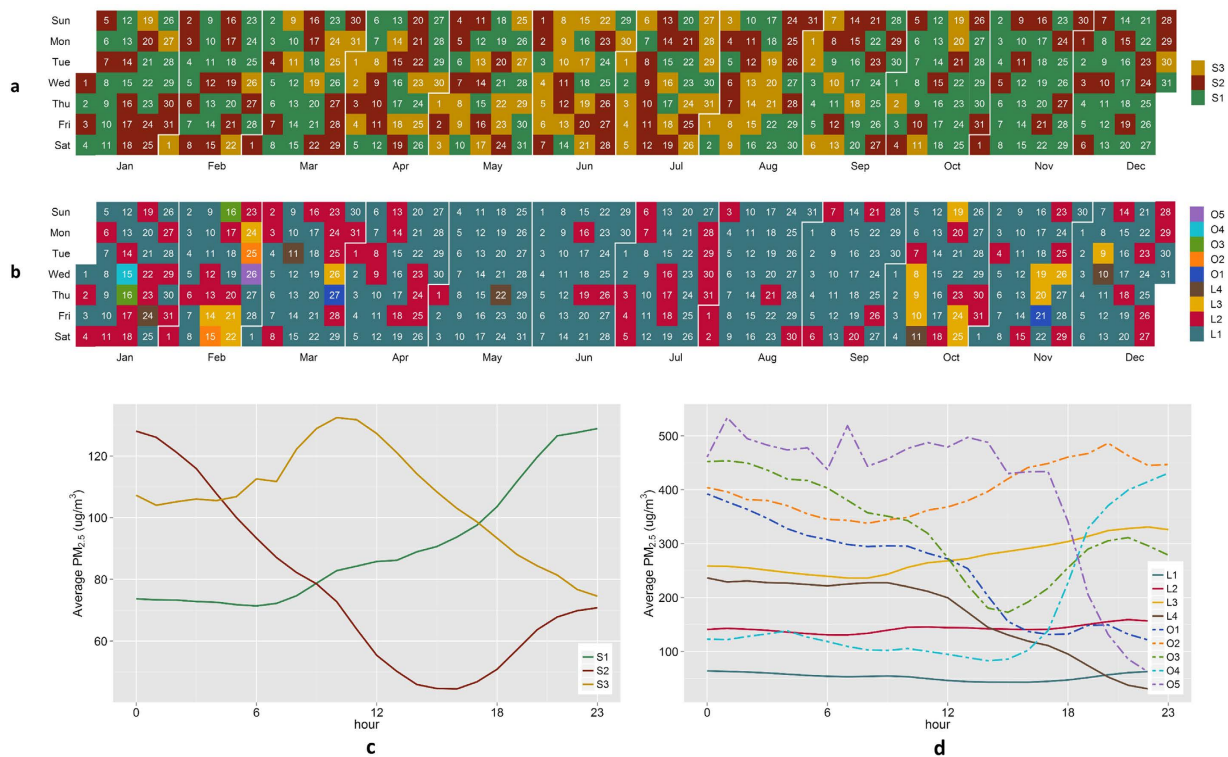


Figure 1. Calendar views of $PM_{2.5}$ concentration clusters in Beijing in the year 2014. (a) shows $PM_{2.5}$ time-series cluster result based on correlation distance, and the letter S denotes “shape”; (b) shows the cluster result based on Euclidean distance, L denotes “level” and O refers to “outlier”. (c) shows the averaged $PM_{2.5}$ trend for clusters based on correlation distance, and (d) shows the averaged $PM_{2.5}$ for clusters based on Euclidean distance. Note that the colours and labels are matched for each cluster for consistency, and the lines for O1, O2, O3, O4, and O5 are set to dash for clear presentation.

our study offers an innovative calendar visualization of $PM_{2.5}$ concentration on a daily basis over the year of 2014, which yields important insights on temporal variation patterns of $PM_{2.5}$ concentration.

The contribution of our study is two-fold. First, our study presents an innovative and straightforward calendar visualization of daily $PM_{2.5}$ time-series in Beijing in the year of 2014. This technique provides a very useful tool to visualize and understand the data and can be applied to examine temporal patterns of other air pollutants. Second, the insights generated from the two calendar plots advance our understanding of Beijing’s $PM_{2.5}$ concentration. Compared to previous studies on Beijing’s $PM_{2.5}$ concentration, our study offers a different perspective and brings in insights on $PM_{2.5}$ concentration that are more complete and convincing.

Results and Discussion

Figure 1 shows two calendar views of the cluster analyses using the correlation distance and Euclidean distance, and two corresponding trend curves of averaged $PM_{2.5}$ concentrations. We obtain three clusters for the analysis based on correlation distance, and each of them has 162, 117, and 86 time-series (days). We named these as S1, S2, and S3, as shown in Fig. 1a,c. For the cluster analysis based on Euclidean distance, nine clusters are formed, each consisting of 255, 82, 15, 5, 2, 2, 2, 1, and 1 time-series. They are named as L1, L2, L3, L4, O1, O2, O3, O4, and O5 (Fig. 1b,d). Those clusters with less than three time-series, namely O1, O2, O3, O4, and O5, are considered as “outliers” that either have extremely high $PM_{2.5}$ concentration or exhibit odd variation patterns. We will discuss these “outliers” later.

Interpretation on calendar visualization. The calendar plot based on correlation distance (Fig. 1a) and the corresponding curve (Fig. 1c) shows the cluster result based on shape differences among the 365 $PM_{2.5}$ time-series. The result shows that there are three distinct variation patterns for the $PM_{2.5}$ time-series. An increasing pattern from 0 AM to 11 PM in a day is most likely to be observed from January to March and from September to December (S1 in Fig. 1a,c). For these days that show an increasing $PM_{2.5}$ concentration pattern, the maximum $PM_{2.5}$ concentration of the day usually occurs at night. The decreasing pattern can be observed in all months throughout the year (S2 in Fig. 1a,c) and this pattern attains its minimum value in the afternoon. The third pattern with a shape like an inverted V often take place from April to August (S3 in Fig. 1a,c) and the $PM_{2.5}$ concentrations during these days usually peaks at noon. These results show that the diurnal patterns of $PM_{2.5}$ vary from day to day through the year, and $PM_{2.5}$ concentration in the daytime could be higher than at night in many days, which complement previous studies concluding that diurnal variation of $PM_{2.5}$ change by seasons and $PM_{2.5}$ concentration at night is higher than that in the daytime^{10,16}.

Our findings are consistent with a previous research which identified a ‘sawtooth cycle’ of $PM_{2.5}$ variation¹⁷. During a ‘sawtooth cycle’, the $PM_{2.5}$ concentration first rises over a few days, which corresponds to the increasing pattern in our study (S1 in Fig. 1a,c), and then falls, which matches the decreasing pattern in our study (S2 in Fig. 1a,c)¹⁷. One possible interpretation is that the increasing and decreasing patterns (S1 and S2 in Fig. 1a,c) are largely formulated by the passage of cold front. When the cold front arrives, high-speed wind associated with the cold front blows the pollution away and thus the $PM_{2.5}$ concentration is decreasing. But when the cold front moves on, cold air underlies the warm air as the cold air is denser and heavier, which leads to temperature inversion. The temperature inversion traps $PM_{2.5}$ pollution near the surface and makes the $PM_{2.5}$ concentration increasing.

Human activities such as heating and combustion, as well as weather conditions including wind, boundary layer height, etc. are closely linked to the variation of $PM_{2.5}$ concentration^{18,19}. As we can see in Fig. 1a,c, not all variation patterns in $PM_{2.5}$ concentration match the daily cycle of human activities such as transportation that usually peaks in the morning and afternoon during a full day. The third pattern (S3 in Fig. 1a,c) is the closest one that possibly matches the daily cycle of human activities but this pattern usually happens from April to August. This finding suggests that the effect of human activities on variations of $PM_{2.5}$ concentration may vary at different time periods. We speculate that from January to March and September to December, weather conditions including cold front, wind, boundary layer height, etc., may be the major factors determining variations in $PM_{2.5}$ concentration. However, from April to August, the weather conditions (e.g., cold front) weaken and human activities thus might have stronger impact on $PM_{2.5}$ variation.

The cluster result based on differences in $PM_{2.5}$ concentration levels can be found in the calendar plot based on Euclidean distance (Fig. 1b) and the corresponding curve (Fig. 1d). We can see that a majority of days in the year have an averaged $PM_{2.5}$ concentration of around $50\mu\text{g}/\text{m}^3$ (L1 in Fig. 1b,d), a figure far beyond the WHO ($25\mu\text{g}/\text{m}^3$) and USA air quality standards ($15\mu\text{g}/\text{m}^3$). The calendar plot also indicates that high averaged $PM_{2.5}$ concentration around $150\mu\text{g}/\text{m}^3$ (L2 in Fig. 1b,d) are likely to occur in every month throughout the year. Also, extremely high $PM_{2.5}$ concentration above $250\mu\text{g}/\text{m}^3$ (L3, O1, O2, O3, O4, and O5 in Fig. 1b,d) can be usually observed in January, February, March, October, November, and December. This finding is consistent with previous studies concluding that $PM_{2.5}$ concentration is generally the highest during winter and lowest during summer^{15,16}.

Outliers. A few “outliers” (O1, O2, O3, O4, and O5 in Fig. 1b,d) can be found in Fig. 1b. For example, two notable “outliers” O4 and O5 on January 15 and February 26, 2014, respectively, show quite drastic variations across the day. As we can see, extremely high $PM_{2.5}$ concentrations (O5 has a maximum $PM_{2.5}$ concentration of $534\mu\text{g}/\text{m}^3$) are observed on the two days and the two incidents were reported by the Guardian²⁰, Time magazine²¹, and Financial Times²².

One event of particular interest is the Asia-Pacific Economic Cooperation (APEC) Summit on 10 and 11 November 2014 in Beijing. It is reported that in order to maintain a blue sky in Beijing during the APEC Summit, coordinated efforts were taken by the governments of Beijing and six surrounding provinces before the summit²³. Measures included impositions on road traffic and plant operations. The two calendar visualization plots in our study indicate that $PM_{2.5}$ concentration was very high in mid-October before the summit. For example, on October 19, 24, and 25, the $PM_{2.5}$ concentration was over $150\mu\text{g}/\text{m}^3$. After the emission control measures were enforced, the $PM_{2.5}$ concentration was greatly reduced on November 1. However, on November 4, a sharp increase in $PM_{2.5}$ concentration was observed, which was around $150\mu\text{g}/\text{m}^3$. Fortunately, a significant reduction occurred on November 5 and $PM_{2.5}$ concentration returned to lower level afterwards by November 15, four days after the summit. These interpretations from the two calendar plots can also be obtained from local observations, but here we would like to note that the two calendar visualizations in our study offer a much more straightforward understanding of the whole picture of $PM_{2.5}$ variations over time than using other tools.

Seasonal and weekly patterns? As we can see from the two cluster results, both shape and level variation do exhibit a rough seasonal pattern but the pattern do not follow strict seasonal divisions. As Fig. 1a shows, S1 pattern usually occurs in around winter seasons (from January to March and from September to December) and S3 patterns often happens around summer times (from April to August). Figure 1b shows a rough seasonal pattern too. Days in L3 cluster usually occur near winter (in February, March, October, November and December but not January) although days in L1 and L2 clusters can be found in any month throughout the full year which doesn’t exhibit very clear seasonal pattern. There may exist significant differences in $PM_{2.5}$ concentration levels between different seasons^{10,15,16}, however we argue that the arbitrary seasonal division of variation in $PM_{2.5}$ concentration may result in information loss and conceal potentially important insights. The calendar visualization used in our study, however, provides an informative and straightforward way to look into variation patterns of air pollutants.

Several studies reported that there existed weekly patterns in $PM_{2.5}$ concentration in Beijing^{9,13}. And their findings are not consistent with each other. One study stated that the lowest concentrations occurred in Mondays while the highest concentrations appeared from Thursdays to Saturday⁹; another study concluded that $PM_{2.5}$ concentrations on weekdays were lower than that on weekends¹³. Our findings, however, do not observe these reported weekly patterns. Figure 1b shows that among all 52 weeks in 2014, higher $PM_{2.5}$ concentrations in weekdays than those in weekends are observed in at least 18 weeks. For example from March 24 to 30, the lowest $PM_{2.5}$ concentrations were observed on weekends while the highest were on weekdays (Fig. 1b). Further calculations show that about half of all weeks in 2014 have higher averaged concentrations on weekdays, and the lowest $PM_{2.5}$ concentrations are observed on Mondays of only 13 weeks and in only 30 weeks the highest concentrations appeared from Thursday to Saturday. Our results do not support the reported weekly patterns.

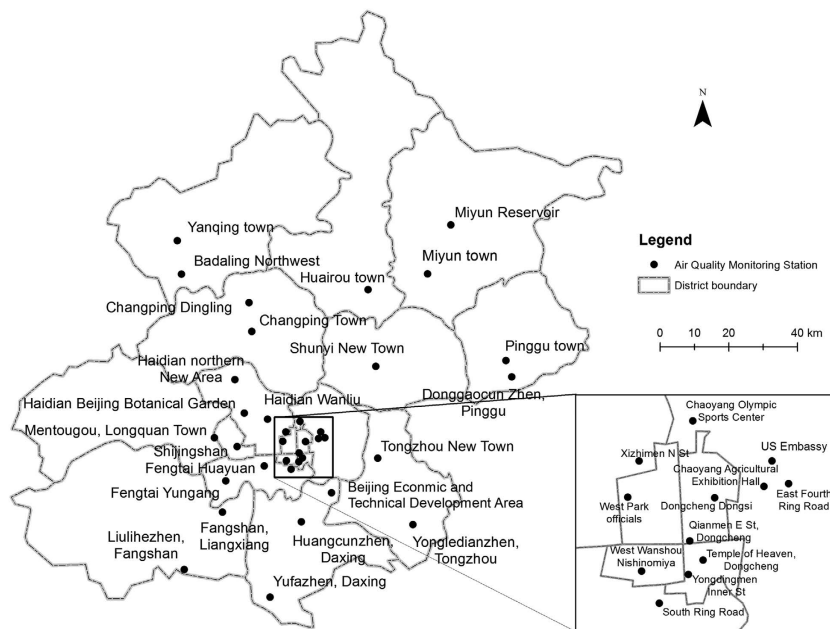


Figure 2. Air Quality Monitoring Stations in Beijing. The map is generated by the authors using ArcGIS 10.2.2 (www.esri.com).

We did not observe any explicit and universal weekly variation pattern after visual inspection over the two calendar plots (Fig. 1a,b) and further calculations. This finding suggests that the weekly cycle of human activities may not play a key role in determining variations in $PM_{2.5}$ concentration. Our finding complements and improves previous studies that report weekly patterns in $PM_{2.5}$ concentration in Beijing^{9,13}.

Future research. As we know, $PM_{2.5}$ pollution can be measured in terms of optical properties and chemical compositions in addition to the mass concentration^{24–27}. With the help of the calendar visualization technique used in this study, these informative properties and other air pollutants such as NO_2 , SO_2 , and O_3 can help provide a better understanding of the air pollution problem.

Data and Methods

Data. The $PM_{2.5}$ measurement data in Beijing used in this study were originally obtained from the official hourly air quality reporting platform (<http://zx.bjmemc.com.cn/>) run by Beijing Environment Protection Agency. This platform is part of the national air quality monitoring network initiated in late 2012. The data is rich, reporting hourly concentrations of six pollutants: particulate matter with aerodynamic diameter no greater than 2.5 microns ($PM_{2.5}$), particulate matter with aerodynamic diameter less than 10 microns (PM_{10}), and sulphur dioxide (SO_2), nitrogen dioxide (NO_2), ozone (O_3), and carbon monoxide (CO) in 35 stations across Beijing (Fig. 2). However, the data is not easily accessible because the online reporting system only reports the air quality of the day and does not show historical data and is unavailable to the public. Fortunately, third parties created by civic efforts such as PM25.in, AQISTUDY.cn, and EPMAP.org have been crawling this data since late 2013.

Our study uses one-year air quality monitoring data from 1 January 2014 to 31 December 2014 from AQISTUDY.cn, EPMAP.org, and the US Embassy Beijing Air Quality Monitor (Fig. 2). We noticed that there are missing hourly measurements in all the three data sources. Therefore, we combined them to get complete $PM_{2.5}$ measurement data covering 24 hours of all the 365 days in 2014. The US Embassy Beijing Air Quality Monitor is operated by the US Department of State. The US Department of State requires that the following disclaimer by included in any publication that uses these data: “State Air observational data are not fully verified or validated; these data are subject to change, error, and correction. The data and information are in no way official”.

A comprehensive data quality check on the raw data is conducted to reduce the impact of problematic data points, including duplicated data records, missing measurements with a placeholder, implausible zeros, etc. After the data quality check, the hourly $PM_{2.5}$ measurement data for all 35 stations are then aggregated into one averaged $PM_{2.5}$ concentration per hour for cluster analysis as explained below.

Method. Since we have 24 hourly $PM_{2.5}$ measurements for each day, it implies we have 365 time-series objects with 24 data points each to analyse. We would like to aggregate together time-series objects with similar variation patterns of $PM_{2.5}$ concentration and separate those with dissimilar patterns into different groups. Thus, we employ time-series clustering technique to mine the data.

In general, there are two essential components in a clustering analysis: clustering algorithm and distance measure²⁸. Clustering algorithm controls the procedures on how similar objects are clustered, while distance measures are used to establish the resemblance between two objects. There are several algorithm and distance measures available in the field of cluster analysis but our study employed the most straightforward and suitable

clustering method and metrics. Specifically, we use average-linkage agglomerative hierarchical clustering as the clustering method because this method generates repeatable and consistent results and does not require the number of clusters to be specified as compared with K-means²⁹, and it is usually able to obtain more robust cluster results than other hierarchical clustering methods³⁰.

Distance measures were selected based on two basic features of the PM_{2.5} time-series data: level and shape. Level refers to the quantity of PM_{2.5} concentration, and the Euclidean distance is used to identify the level difference between PM_{2.5} time-series. Shape refers to trends in PM_{2.5} concentration variation with respect to time, and we use Pearson's correlation-based distance to capture the shape difference between PM_{2.5} time-series. We derived a generalized correlation-based dissimilarity function from this study³¹ by making the coefficient α and power β adjustable (equation (1)).

$$D(x, y) = (\alpha(1 - \rho))^\beta \quad (1)$$

where the correlation coefficient $\rho = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}$, with $\alpha, \beta > 0$

This dissimilarity function satisfies all the requirements for dissimilarity measure: the non-negativity, symmetry, and identity^{32,33}. When both α and β are set to 1, this dissimilarity function becomes the classic Pearson's correlation coefficient distance that has been used in several studies³⁴. In our study, however, we deliberately set α and β to 0.5 and 0.25, respectively, in order to attain a desirable robust cluster result.

We employ the cophenetic correlation coefficient to examine the validity and robustness of the cluster analysis. Cophenetic correlation coefficient is a measure of how faithfully the hierarchical cluster results represent the dissimilarity among observations³⁵. It is defined as the linear correlation coefficient between the original pairwise dissimilarities and the cophenetic dissimilarities obtained from the dendrogram. The value of this coefficient varies between 0 and 1. A higher cophenetic correlation coefficient indicates a better cluster solution, and a value of 0.8 or higher is usually regarded as a successful cluster application³⁶.

It turns out that the cophenetic correlation coefficients for Euclidean-distance-based and correlation-distance-based cluster analyses are 0.86 and 0.81, respectively, suggesting that both cluster results are robust and valid.

We used Python version 2.7.5 to process and analyse the data, and R version 3.2.2 to draw the calendar plots.

References

1. Yuan, Y., Liu, S., Castro, R. & Pan, X. PM2.5 monitoring and mitigation in the cities of China. *Environmental science & technology* **46**, 3627–3628, doi: 10.1021/es300984j (2012).
2. Han, L., Zhou, W. & Li, W. Increasing impact of urban fine particles (PM2.5) on areas surrounding Chinese cities. *Scientific reports* **5**, 12467, doi: 10.1038/srep12467 (2015).
3. Wong, C. M. *et al.* Satellite-Based Estimates of Long-Term Exposure to Fine Particles and Association with Mortality in Elderly Hong Kong Residents. *Environ Health Perspect* **123**, 1167–1172, doi: 10.1289/ehp.1408264 (2015).
4. Samet, J. M., Dominici, F., Currier, I. & Zeger, S. L. Fine particulate air pollution and mortality in 20 US cities, 1987–1994. *New England journal of medicine* **343**, 1742–1749 (2000).
5. Dominici, F. *et al.* Fine particulate air pollution and hospital admission for cardiovascular and respiratory diseases. *JAMA: the journal of the American Medical Association* **295**, 1127–1134 (2006).
6. Jedrychowski, W. *et al.* Gender differences in fetal growth of newborns exposed prenatally to airborne fine particulate matter. *Environ. Res.* **109**, 447–456 (2009).
7. Power, M. C. *et al.* The relation between past exposure to fine particulate air pollution and prevalent anxiety: observational cohort study. *BMJ* **350**, h1111, doi: 10.1136/bmj.h1111 (2015).
8. Streets, D. G. *et al.* Air quality during the 2008 Beijing Olympic Games. *Atmospheric Environment* **41**, 480–492, doi: 10.1016/j.atmosenv.2006.08.046 (2007).
9. Chen, W., Tang, H. & Zhao, H. Diurnal, weekly and monthly spatial variations of air pollutants and air quality of Beijing. *Atmospheric Environment* **119**, 21–34 (2015).
10. Huang, F. *et al.* PM2.5 Spatiotemporal Variations and the Relationship with Meteorological Factors during 2013–2014 in Beijing, China. *PLoS One* **10**, e0141642 (2015).
11. Zheng, S., Pozzer, A., Cao, C. & Lelieveld, J. Long-term (2001–2012) concentrations of fine particulate matter (PM2.5) and the impact on human health in Beijing, China. *Atmospheric Chemistry and Physics* **15**, 5715–5725 (2015).
12. Chai, F. *et al.* Spatial and temporal variation of particulate matter and gaseous pollutants in 26 cities in China. *Journal of Environmental Sciences* **26**, 75–82, doi: 10.1016/S1001-0742(13)60383-6 (2014).
13. Hu, J., Wang, Y., Ying, Q. & Zhang, H. Spatial and temporal variability of PM2.5 and PM10 over the North China plain and the Yangtze River delta, China. *Atmospheric Environment* **95**, 598–609, doi: 10.1016/j.atmosenv.2014.07.019 (2014).
14. Rohde, R. A. & Muller, R. A. Air Pollution in China: Mapping of Concentrations and Sources. *PLoS One* **10**, e0135749, doi: 10.1371/journal.pone.0135749 (2015).
15. Wang, Y., Ying, Q., Hu, J. & Zhang, H. Spatial and temporal variations of six criteria air pollutants in 31 provincial capital cities in China during 2013–2014. *Environ. Int.* **73**, 413–422, doi: 10.1016/j.envint.2014.08.016 (2014).
16. Zhang, Y.-L. & Cao, F. Fine particulate matter (PM2.5) in China at a city level. *Scientific Reports* **5**, 14884, doi: 10.1038/srep14884 (2015).
17. Jia, Y., Rahn, K. A., He, K., Wen, T. & Wang, Y. A novel technique for quantifying the regional component of urban aerosol solely from its sawtooth cycles. *Journal of Geophysical Research: Atmospheres (1984–2012)* **113**, doi: 10.1029/2008JD010389 (2008).
18. Wang, J., Hu, Z., Chen, Y., Chen, Z. & Xu, S. Contamination characteristics and possible sources of PM10 and PM2.5 in different functional areas of Shanghai, China. *Atmospheric Environment* **68**, 221–229 (2013).
19. Zhao, X. *et al.* Seasonal and diurnal variations of ambient PM2.5 concentration in urban and rural environments in Beijing. *Atmospheric Environment* **43**, 2893–2900, doi: 10.1016/j.atmosenv.2009.03.009 (2009).
20. Gardner, D. *Beijing's smog is increasingly toxic for China's politicians.* (2014) Available at: <http://www.theguardian.com/environment/2015/jan/20/beijings-smog-increasingly-toxic-chinas-politicians> (Accessed: 9 December 2015).
21. Beech, H. *China's Smog Is So Bad They're Now Calling It a 'Nuclear Winter'.* (2014) Available at: <http://time.com/9802/beijing-air-pollution-nuclear-winter/> (Accessed: 9 December 2015).
22. Hornby, L. *China pollution: Trouble in the air.* (2014) Available at: <http://on.ft.com/1bMwV4d> (Accessed: 9 December 2015).

23. Zheng, S. R. *Beijing Smog: The Day After 'APEC Blue'*. (2014) Available at: <http://thediplomat.com/2014/11/beijing-smog-the-day-after-apec-blue/> (Accessed: December 9 2015).
24. Wang, X., Huang, J., Zhang, R., Chen, B. & Bi, J. Surface measurements of aerosol properties over northwest China during ARM China 2008 deployment. *Journal of Geophysical Research: Atmospheres* **115**, doi: 10.1029/2009JD013467 (2010).
25. Zhang, R. *et al.* Chemical characterization and source apportionment of PM_{2.5} in Beijing: seasonal perspective. *Atmos. Chem. Phys.* **13**, 7053–7074, doi: 10.5194/acp-13-7053-2013 (2013).
26. Bi, J., Huang, J., Hu, Z., Holben, B. & Guo, Z. Investigating the aerosol optical and radiative characteristics of heavy haze episodes in Beijing during January of 2013. *Journal of Geophysical Research: Atmospheres* **119**, 9884–9900 (2014).
27. Huang, R.-J. *et al.* High secondary aerosol contribution to particulate pollution during haze events in China. *Nature* **514**, 218–222, doi: 10.1038/nature13774 (2014).
28. Wang, X., Smith, K. & Hyndman, R. Characteristic-based clustering for time series data. *Data mining and knowledge Discovery* **13**, 335–364 (2006).
29. Aghabozorgi, S., Seyed Shirkhorshidi, A. & Ying Wah, T. Time-series clustering – A decade review. *Information Systems* **53**, 16–38, doi: 10.1016/j.is.2015.04.007 (2015).
30. Rodrigues, F. M. & Diniz-Filho, J. A. F. Hierarchical structure of genetic distances: effects of matrix size, spatial distribution and correlation structure among gene frequencies. *Genetics and Molecular Biology* **21**, 233–240 (1998).
31. Golay, X. *et al.* A new correlation-based fuzzy logic clustering algorithm for fMRI. *Magnetic Resonance in Medicine* **40**, 249–260 (1998).
32. Veltkamp, R. C. Shape matching: Similarity measures and algorithms. Paper presented at *SMI 2001 International Conference on Shape Modeling and Applications*, Genova. IEEE. doi: 10.1109/SMA.2001.923389 (2001, May 7).
33. Wang, L., Yang, C. & Feng, J. On learning with dissimilarity functions. Paper presented at *Proceedings of the 24th international conference on Machine learning*, Corvallis, Oregon, USA. New York, NY, USA: ACM. doi: 10.1145/1273496.1273621 (2007, June 20).
34. Jaskowiak, P. A., Campello, R. J. & Costa, I. G. On the selection of appropriate distances for gene expression data clustering. *BMC bioinformatics* **15**, doi: 10.1186/1471-2105-15-S2-S2 (2014).
35. Sokal, R. R. & Rohlf, F. J. The comparison of dendrograms by objective methods. *Taxon* **11**, 33–40 (1962).
36. Romesburg, C. *Cluster analysis for researchers* (Lulu Press, 2004).

Acknowledgements

This research was supported by National Natural Science Foundation of China (Grant No. 41471370).

Author Contributions

J. Liu designed the study, performed data analysis and wrote the manuscript. J. Li participated in data analysis. W.L. participated in data analysis, reviewed and approved the manuscript.

Additional Information

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Liu, J. *et al.* Temporal Patterns in Fine Particulate Matter Time Series in Beijing: A Calendar View. *Sci. Rep.* **6**, 32221; doi: 10.1038/srep32221 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2016