

RESEARCH

Open Access



Hazard ranking of wastewater sources in a highly polluted river in northern Taiwan by using positive matrix factorization with metal elements

Pei-Yuan Hsieh¹, Huan-Chun Lin¹, Gen-Shuh Wang¹, Yuan-Jeng Hsu², Yi-Ju Chen², Tzu-Hui Wang², Ren-De Wang², Chun-Yu Kuo², Di-Wen Wang², Ho-Tang Liao¹ and Chang-Fu Wu^{1*} 

Abstract

Improving water quality is a critical issue worldwide. However, the general parameters (i.e., temperature, pH, turbidity, total solids, fecal coliform, dissolved oxygen, biochemical oxygen demand, phosphates, and nitrates) used in water quality index estimations are unable to identify pollution from industrial wastewater. This study investigated pollution sources at a river pollution hotspot by using the positive matrix factorization (PMF) model. A two-phase sampling collection along a highly polluted river in northern Taiwan was designed. The sampling spots were distributed along the river in Phase I to monitor the spatial variation of river pollutants. A pollution hotspot was determined based on two indices, namely the summed concentrations of metal elements and a metal index (MI). In Phase II, the river water samples were collected from the hotspot twice daily over 30 consecutive days to monitor the temporal variation of river pollutants. Source profiles of metal elements were obtained during the monitoring period. The Phase II samples were then factorized using the PMF model. Factor profiles retrieved from the PMF model were further assigned to industrial categories through Pearson correlation coefficients and hierarchical classification. The results indicated that the main pollution source was bare printed circuit boards (BPCB), which contributed up to 92% of the copper in the pollution hotspot. In terms of MI apportionment of 11 metals related to health effects, BPCB contributed 91% of the MI in high pollution events. Overall, the MI apportionment provides linkages between pollution level and human health. This is an evidence for policymakers that the regulation of the effluents of BPCB is an effective means to controlling copper concentrations and thus improving water quality in the study area.

Keywords: River water pollution, Industrial wastewater, Source profile, Source apportionment

1 Introduction

On September 25, 2015, the 193 member states of the United Nations adopted the 2030 Agenda for its 17 Sustainable Development Goals (SDGs) [1]. The agenda was expected to guide the actions of international

communities over the next 15 years (2016–2030). One of the 17 SDGs involves improving water quality by reducing river water pollution. The causes of river water pollution include industrial activity, domestic sewage, livestock wastewater, and agricultural wastewater.

The water quality index (WQI) is commonly used to evaluate general water quality [2–4]. Multiple parameters are involved in the estimation of WQI scores, and weighted coefficients are assigned to each parameter based on statistical surveys. The general parameters are

*Correspondence: changfu@ntu.edu.tw

¹ Institute of Environmental and Occupational Health Sciences, College of Public Health, National Taiwan University, Taipei 10055, Taiwan
Full list of author information is available at the end of the article



© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

temperature, pH, turbidity, total solids, fecal coliform, dissolved oxygen, biochemical oxygen demand, phosphates, and nitrates [5, 6]. However, an index using these parameters generally cannot be used to measure pollution from industrial wastewater.

River water influenced by discharge from an industrial district usually has a complex composition including metal elements, water-soluble ions, and volatile organic compounds. Stable and non-degradable metal elements could act as suitable tracers for monitoring pollution discharged by industries [7, 8]. A metal index (MI) based on maximum allowable concentration (MAC) has been introduced to evaluate the quality of water resources [9, 10]. Usually, the MAC of an element is set by regulatory agencies to protect human health and ecological environment, which makes the MI a valuable indicator.

To decrease the health hazards of polluted water and improve water quality, identifying pollution sources is critical. Various statistical techniques have been applied for source apportionment. Principle component analysis (PCA) has been widely used in the source apportionment of water pollution [11–15]. PCA yields linear combinations of original features and can transform large datasets into a set of factors with reduced dimensions. However, PCA can only provide qualitative information, and it is susceptible to outliers. Furthermore, the results of PCA does not have physical meaning because the values of factors decomposed using PCA can be negative. Positive matrix factorization (PMF) is recognized as an effective scientific tool and as an improved factor analysis tool for source apportionment [16]. Many studies have successfully applied PMF to the source apportionment of ambient particulate matter [17, 18]. In recent years, PMF has also been used to investigate pollution sources in aqueous environments (e.g., [19–23]). PMF has the following advantages: (1) it includes the integration of non-negative constraints; (2) uncertainties in the data are introduced into the model; (3) it does not need to be orthogonal, which makes the formed factors close to real profile of pollution sources; (4) it quantifies source contributions; and (5) the factors are not excessively influenced by outliers [24–26]. However, in most of these studies, source apportionment was estimated using spatial samplings, which neglected temporal variation in the source contributions. Furthermore, the number of metal species was limited and may be insufficient to distinguish complex industrial sources.

The objective of this study was to use metal species with a receptor model to estimate potential source profiles and their contributions in a river pollution hotspot. The Ta-Liao-Keng River was selected as an example to demonstrate the feasibility of the approach. To the best of our knowledge, this is the first study adopting sequential

time series data of river water samples for PMF modeling with factor profiles identified by comparison with profiles obtained from industrial effluents. Potential sources of MI and their contributions were apportioned to clarify the relationship between pollution levels and human health.

2 Material and methods

2.1 Study area and river water sample collection

The Ta-Liao-Keng River is a tributary of the Dahan River, which is a major river in Taiwan (Fig. 1). The length of the Ta-Liao-Keng River is 12.3 km, and its catchment area is approximately 29.4 km². The river has four branches and flows eastward from Taoyuan City to New Taipei City (A fishbone diagram is provided in Fig. S1 of Supplementary Materials). The main industrial area is located at a downstream section of the river and is especially concentrated on Tandigou, a tributary (S4 in Fig. 1). According to a report published by the local government in 2017, the Ta-Liao-Keng River was the main contributor of copper (Cu) to the Dahan River. Thus, further investigation of the source pollution, especially for Cu, in the Ta-Liao-Keng River is imperative.

For this study, a sampling scheme with two phases was designed. In the first phase (Phase I, end of April 2019), four sampling sites along the Ta-Liao-Keng River were selected to locate a hotspot for further investigation. The river water quality was monitored from downstream to upstream. Four additional sampling sites at four branches served as background sites. A total of eight river water samples were collected during this phase. During the sampling period, mild weather (with at most light rainfall) minimized the potential contributions of suspended solids from sediments to metal elements in the river.

Two pollution indices, the sum of elemental metal concentrations and MI, were applied to investigate the pollution hotspot. The MI was defined as

$$MI = \sum_i \frac{c_i}{(MAC_i)} \quad (1)$$

where C_i is the concentration of one of 11 controlled elements, i.e., silver (Ag), arsenic (As), cadmium (Cd), Cu, hexavalent chromium (Cr⁶⁺), mercury (Hg), manganese (Mn), nickel (Ni), lead (Pb), selenium (Se), and zinc (Zn). Table 1 presents the MAC values of these 11 elements, which are enforced by Taiwan Environmental Protection Administration (TEPA), for surface water. Because trivalent chromium and Cr⁶⁺ can convert back and forth, a measurement of total chromium can avoid missing one of these. Instead of using Cr⁶⁺, total chromium was measured and used to estimate the MI. Pollution is a cause for concern when the MI is greater than one.

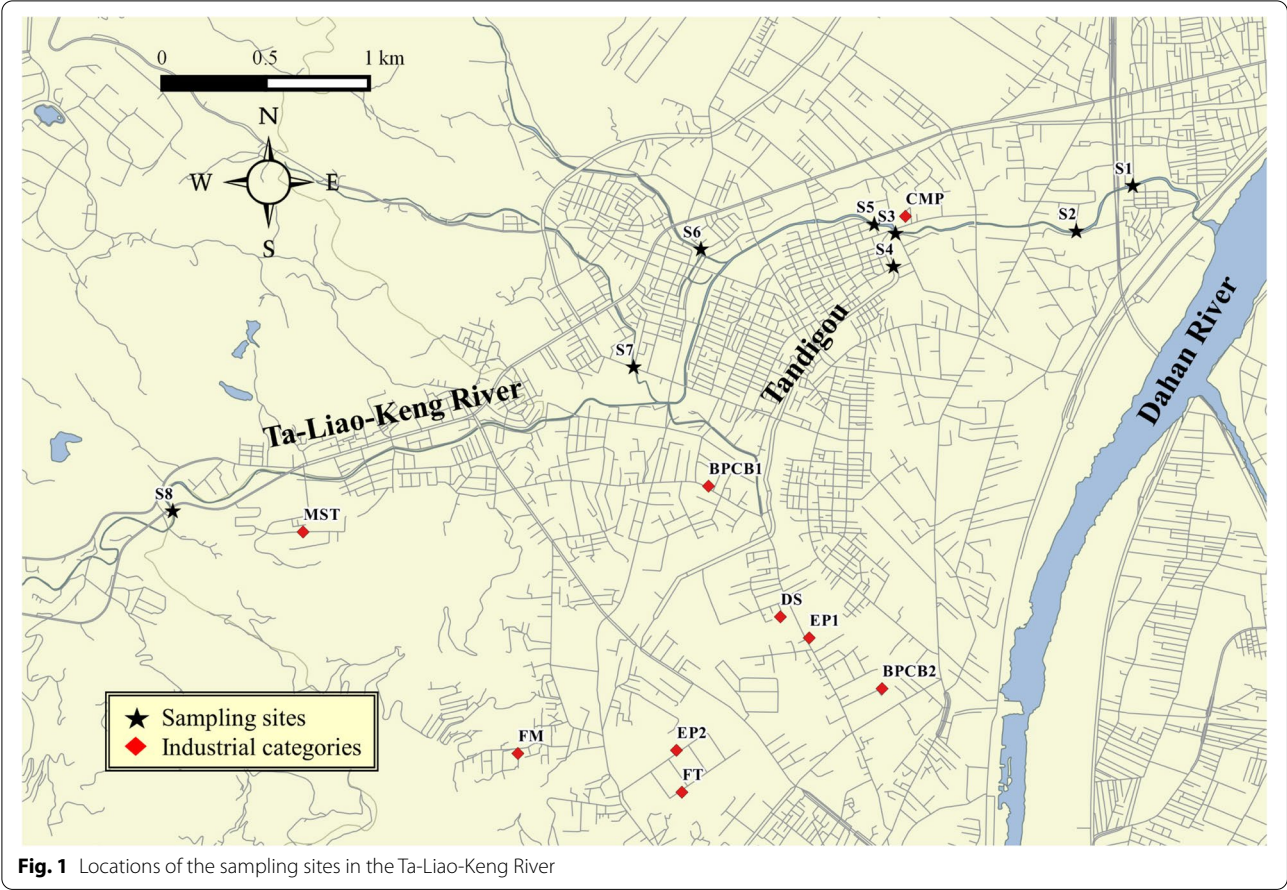


Table 1 Water quality standards of TEPA for heavy metal content in surface water

Element	Ag	As	Cd	Cr ⁶⁺	Cu	Hg	Mn	Ni	Pb	Se	Zn
MAC ^a (µg L ⁻¹)	50	50	5	50	30	1	50	100	10	10	500

^a MAC Maximum allowable concentration

In the second phase (Phase II, July 2019), samples were collected twice daily over 30 consecutive days ($N=60$) at the hotspot identified in Phase I. The collected samples were used to evaluate temporal variation in river water pollution and for PMF modeling.

2.2 Sampling of industrial wastewater

To investigate potential pollution sources and assign them to industrial categories, representative source profiles were needed. A total of 20 wastewater samples from 9 effluent discharges were collected as source profiles. Furthermore, because the river could be contaminated by untreated wastewaters which were discharged by factories unintentionally, three manufacturing

units were selected for sampling based on previous violations. Most industrial wastewater samples were collected twice during the sampling time to reduce uncertainty in the source profiles (Table S1). Overall, samples were collected for six industrial categories: chemical materials and products (CMP), bare printed circuit boards (BPCB), electroplating products (EP), food manufacturing (FM), finishing of textiles (FT), and metal surface treatment (MST). To further investigate the association between human activities and the river pollution, domestic sewage (DS) was collected as a potential source. The wastewater samples were collected and preserved in accordance with the Standard Method [27].

2.3 Chemical analysis

To quantify metal elements and trace elements in river water and wastewater samples, inductively coupled plasma–mass spectrometry was used according to the Standard Method [28]. A total of 52 elements were analyzed and the method detection limit (MDL) of these elements ranged from 0.03 to 1.67 $\mu\text{g L}^{-1}$ (Table S2).

2.4 Source apportionment

The PMF 5.0 software released by U.S. Environmental Protection Agency was used for source apportionment in the second phase [29]. In PMF modeling, the mass balance equation is expressed as

$$X_{ij} = \sum_{k=1}^p g_{ik} f_{kj} + e_{ij} \quad (2)$$

where X_{ij} is the j th species concentration measured in the i th sample, g_{ik} is the contribution of the k th source pollution in the i th sample, f_{kj} is the source profile defined by the j th species concentration in the contribution of the k th source pollution, and e_{ij} is the residual.

The uncertainty of measurements in PMF is calculated using

$$u_{ij} = \sqrt{(0.5 \times \text{MDL}_{ij})^2 + (0.1 \times X_{ij})^2} \quad (3)$$

when the measured concentration is higher than the MDL [29]. When the measured concentration is less than the MDL, the concentration is replaced with half the MDL, and the uncertainty is estimated as follows:

$$u_{ij} = \frac{5}{6} \times \text{MDL}_{ij}. \quad (4)$$

In PMF, the objective function is defined using

$$Q(E) = \sum_{i=1}^n \sum_{j=1}^m \left(\frac{e_{ij}}{u_{ij}} \right)^2 \quad (5)$$

where n is the number of samples and m is the number of species. The optimal values of contributions (g_{ik}) and sources profiles (f_{kj}) can then be retrieved by minimizing $Q(E)$.

To determine the number of factors, PMF was run sequentially using a varied number of factors from 3 to 8. The scaled residual (r_{ij}) of each run is used to calculate the maximum individual column mean (IM),

$$\text{IM} = \max_{j=1 \dots m} \left(\frac{1}{n} \sum_{i=1}^n r_{ij} \right) \quad (6)$$

and the maximum individual column standard deviation (IS),

$$\text{IS} = \max_{j=1 \dots m} \left(\sqrt{\frac{1}{n-1} \sum_{i=1}^n (r_{ij} - \bar{r}_i)^2} \right) \quad (7)$$

The appropriate number of factors is that for which the derivatives of IM and IS seem to level off [30].

In PMF modeling, bootstrapping is performed to evaluate the stability of the solution and the uncertainties of the source contribution estimates. The reproducibility of bootstrapping on each run (three to eight factors) is used to determine the number of factors, as well. In this study, 100 was chosen for the number of bootstrap runs. The steps taken in this research is shown in Fig. S2.

3 Results and discussion

3.1 Spatial and temporal variations in river water measurements

The sum of elemental concentrations was low from the upstream reaches and gradually increased along the river in Phase I (Table S3, Sites S8, S5, S3, and S1). The two species with the highest concentrations were aluminum (Al) and iron (Fe); the sum of Al and Fe at S6 was 86% of the elemental concentrations. Al and Fe are the most abundant metals in the Earth's crust. To focus on industrial pollutants, Al and Fe were eliminated from the following estimations. The highest elemental concentration was observed at Tandigou (S4), a tributary of the Ta-Liao-Keng River. The concentration at S4 was almost six times greater than that of the background water (S8). The second highest value was observed in the other tributary, S2. The high concentrations in tributaries instead of in the main stream was probably due to inputs from domestic wastewater and industrial effluents in the downstream area. The MI at S4 was up to 42.8, 17 times higher than that at S8, which indicated high levels of pollution at S4. Based on the spatial analysis of the metal elements in Phase I, S4 was chosen as the pollution hotspot.

In Phase II, the pollution hotspot was monitored twice daily over 30 consecutive days beginning on July 1, 2019. Summary statistics for the measured elemental concentrations are provided in Table S2. The three elements with the highest mean concentrations (exclude Al and Fe) were Cu ($1246 \mu\text{g L}^{-1}$), Zn ($558 \mu\text{g L}^{-1}$), and Ni ($313 \mu\text{g L}^{-1}$). When applying the PMF model, data pretreatment is crucial. Species must be removed if more than 70% of the data points are below the detection limit. In this study, the following 13 species were excluded: dysprosium, erbium, europium, hafnium, holmium, iridium, platinum, ruthenium, samarium,

tellurium, thulium, ytterbium, and zirconium. Most of the excluded species are rare earth elements and precious metals. Furthermore, data quality is evaluated by signal-to-noise ratios in PMF modeling and inadequate species can be weighted down or eliminated during the optimization. We eliminated six elements with poor fit, namely gold, boron, cobalt, molybdenum, antimony, and tantalum. Finally, 31 of the 50 elements were included in the PMF model.

The highest total elemental concentration at S4 was observed on July 1 when the concentration was up to $13,465 \mu\text{g L}^{-1}$ (Fig. 2). The second highest total elemental concentration, $9100 \mu\text{g L}^{-1}$, was observed on July 15. In terms of the MI, the largest value, 258, was also observed on July 1 (Fig. 3). The second largest MI value, 170, was observed on July 4. A discrepancy was discovered when defining the pollution events using two indices. This is discussed in the following sections.

3.2 Characteristics of the metal elements in industrial wastewater

Most industrial wastewater was collected twice, and the samples for each factory were averaged to create a composite profile. The composite profiles of eight industrial and three manufacturing unit (with “m” in front of the category) discharges are illustrated in Fig. 4. CMP was represented by a battery factory, which emitted a high concentration of Pb. The profiles for BPCB were dominated by Cu, which comprised up to 82% of total elemental concentrations. The profiles of the EP were dominated by Ni, with proportions of up to 88%. FM was represented by a factory making frozen food, where discharge had high concentrations of strontium (Sr) and rubidium. The profile of FT was dominated by Sr and Mn. The MST was represented by a manufacturer of aluminum-related products, and Mn and Ni were dominant in its profile. The profile of DS was collected at an apartment complex

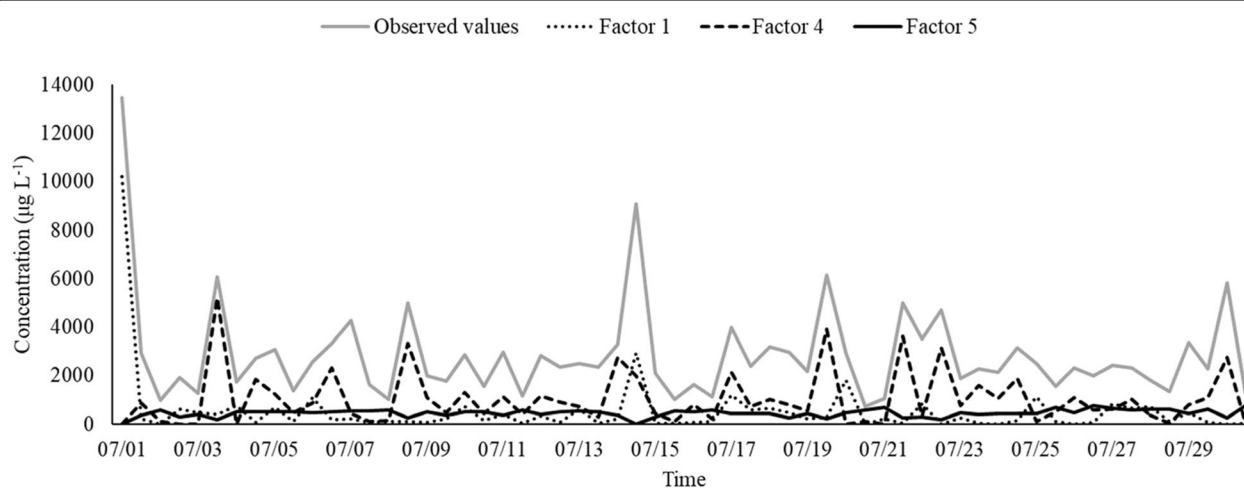


Fig. 2 Measured total metal elemental concentrations and the contributions of Factors 1, 4, and 5 over time

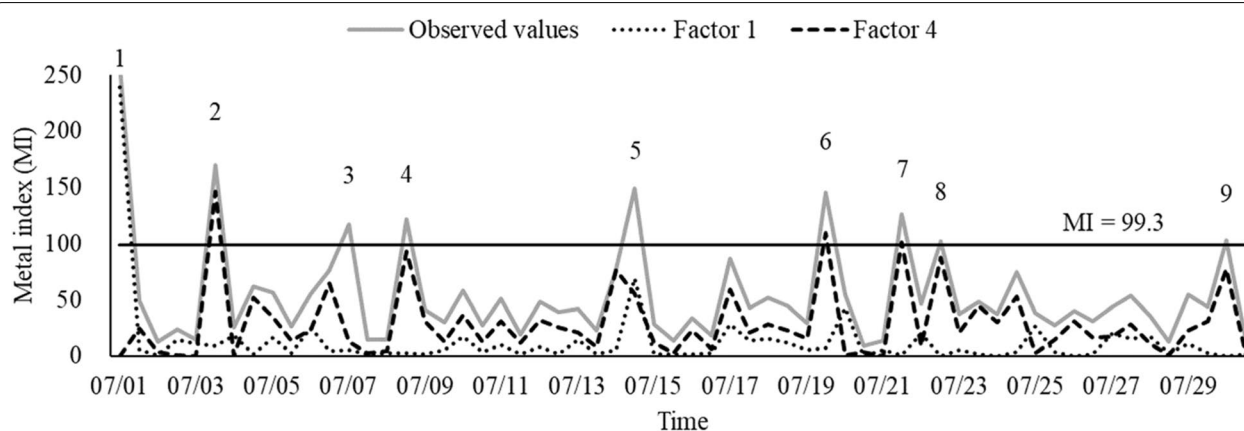


Fig. 3 Measured MI and contributions of Factors 1 and 4 estimated using PMF. The event number is displayed above the peak value

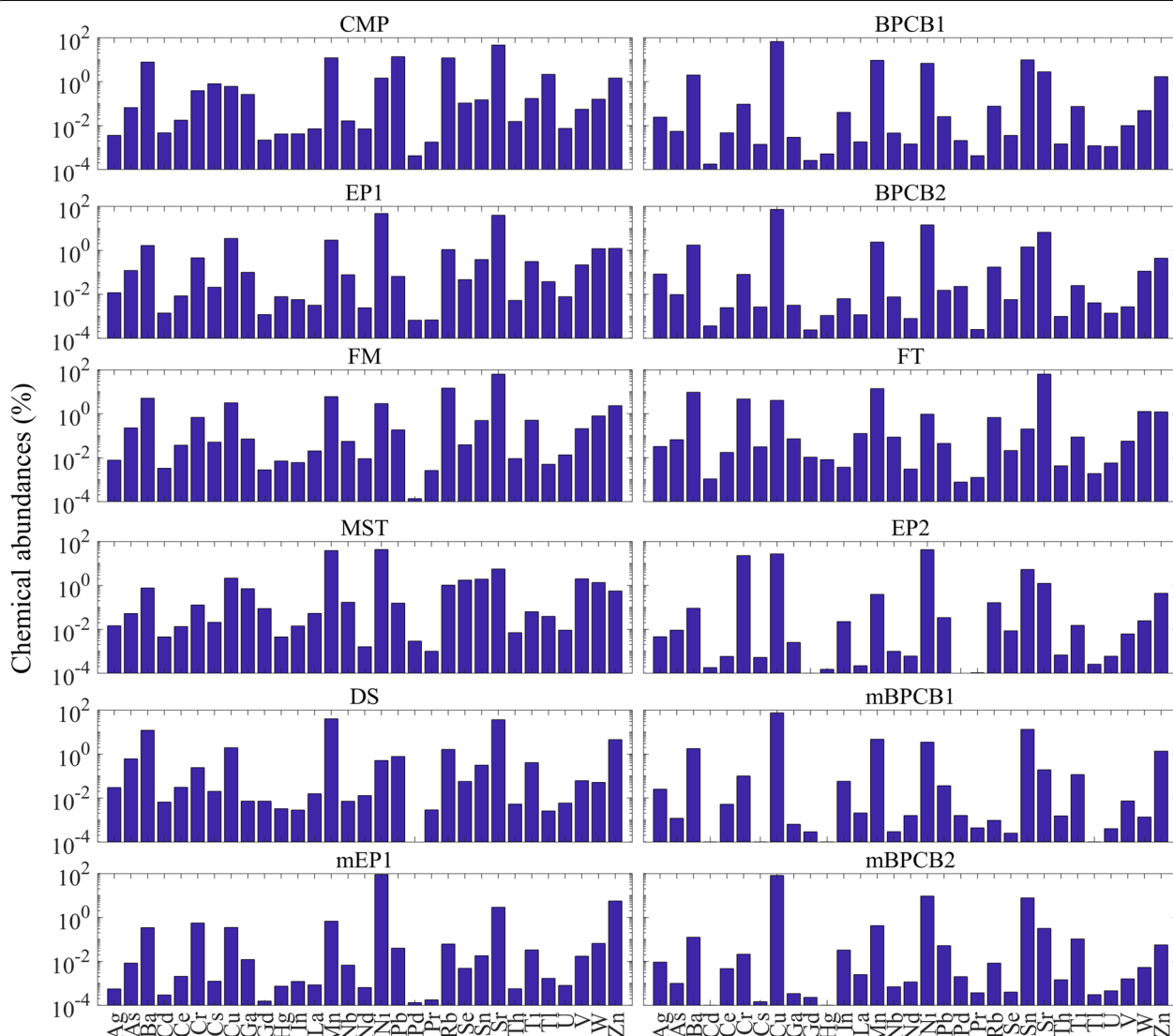


Fig. 4 Composite profiles of industrial effluents and manufacturing unit discharges

and was dominated by Mn and Sr. All compositions followed the effluent standards established by the TEPA.

3.3 Source apportionment

IM and IS analyses indicated that possible solutions were around six factors (Fig. S3). To identify the optimal solution, bootstrap analysis was introduced (Table S4). This revealed that five factors provided the most stable solution with the mapping rate $>80\%$ for each factor. The averaged source contributions of total elemental concentrations (the sum of 31 elements) and of Cu are depicted in Fig. 5. Factor 4 was the greatest contributor ($36 \pm 5\%$) to total elemental concentrations, followed by Factor 1 ($19 \pm 8\%$) and Factor 5 ($16 \pm 4\%$). For Cu, Factors 4

and 1 offered even larger contributions ($64 \pm 23\%$ and $28 \pm 18\%$, respectively). To identify the industrial categories of the three major contributors, Pearson correlations and hierarchical classifications were used. Table 2 lists the correlations of Factors 1, 4, 5, with all source profiles collected in this study. Factors 1 and 4 had strong correlations ($r \geq 0.96$) with BPCB1, BPCB2, and their manufacturing unit discharges. The correlation between Factor 1 and Factor 4 was as high as 0.99. This indicated that Factors 1 and 4 were in the same industrial category. Factor 5 was highly correlated with DS ($r = 0.95$). Figure 6 depicts a cluster dendrogram of the hierarchical classification of the factors and source profiles. Factor 1 was similar to BPCB1 and its manufacturing unit discharge, and Factor

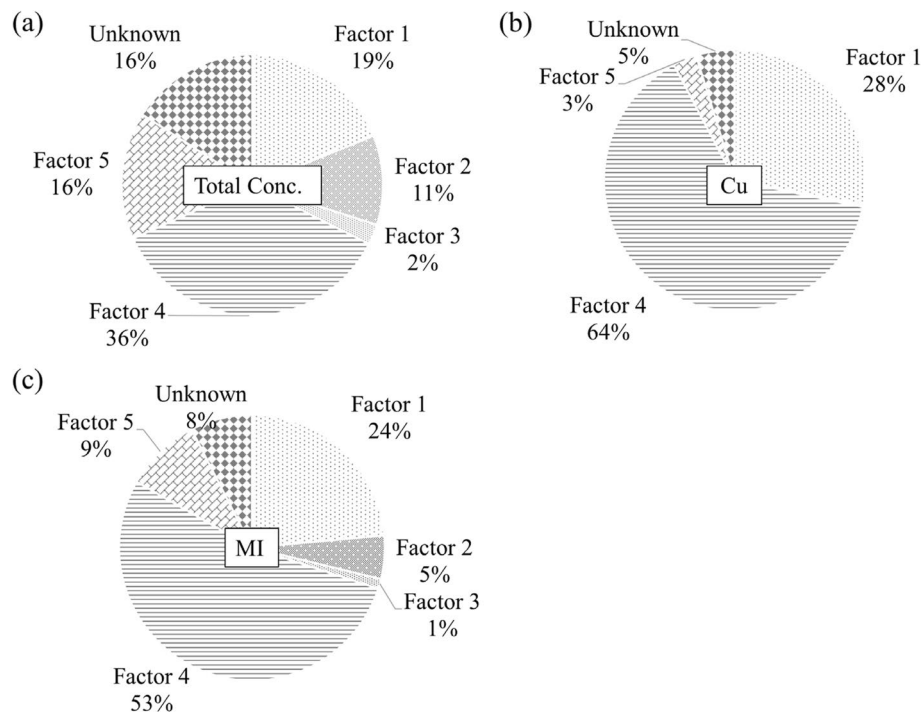


Fig. 5 Source contributions of five factors to (a) total elemental concentrations, (b) Cu, and (c) MI

Table 2 Correlations of Factors 1, 4, 5, with all source profiles

Source Pollution	Factor 1	Factor 4	Factor 5
CMP ^a	−0.04	−0.02	0.68
BPCP1	0.98	0.98	0.24
EP1	0.04	0.06	0.51
BPCP2	0.96	0.98	0.21
FM	0.01	0.03	0.63
FT	0.03	0.05	0.72
MST	0.03	0.06	0.63
EP2	0.49	0.49	0.14
DS	0.03	0.05	0.95
mBPCP1 ^b	0.98	0.98	0.16
mEP1	0.01	0.02	0.16
mBPCP2	0.98	0.98	0.13

^a CMP Chemical materials and products, BPCB Bare printed circuit boards, EP Electroplating products, FM Food manufacturing, FT Finishing of textiles, MST Metal surface treatment, DS Domestic sewage. Three manufacturing units collected from BPCP1, EP1 and BPCP2 are with "m" in front of the category

^b The category with m means the manufacturing unit

4 was similar to BPCB2 and its manufacturing unit discharge. Among all source profiles and factors, BPCB1 and BPCB2 were prominent in the hierarchical clustering. Moreover, Factor 5 was closely related to DS. BPCB1 and BPCB2 are Cu-dominated, and the profile of DS is

dominated by Mn and Sr. Overall, the profile matching based on the two methods was consistent, and the markers for the estimated factors corresponded to those in the measured profiles. Factors 1 and 4 were classified as BPCB1 and BPCB2, respectively, and Factor 5 was classified as DS.

The measured total elemental concentrations and the contributions of Factors 1, 4, and 5 over time are presented in Fig. 2. The contribution of Factor 4 closely matched certain high pollution events. On July 4, Factor 4 contributed 86% of total elemental concentrations and decreased in the afternoon. Factor 1 was the main contributor (76%) of the highest pollution event on July 1 even though it was not a persistent pollution source. By contrast, Factor 5 was a stable, persistent source of pollution, but its contribution was relatively small. Concentrations of Cu were greater than $3000 \mu\text{g L}^{-1}$ in many high pollution events (Fig. S4), and the ratio of Cu to total elemental concentration was $>50\%$. Factor 1 was the significant contributor of Cu on July 1. Factor 4 was the main source of Cu in high pollution events, except for the highest one. For Factor 1, Cu comprised 65% of the normalized profile (Fig. S5). Cu was also dominant in Factor 4 and represented up to 78% of total elemental concentrations. Instead of Cu, Mn and Sr were significant in Factor 5.

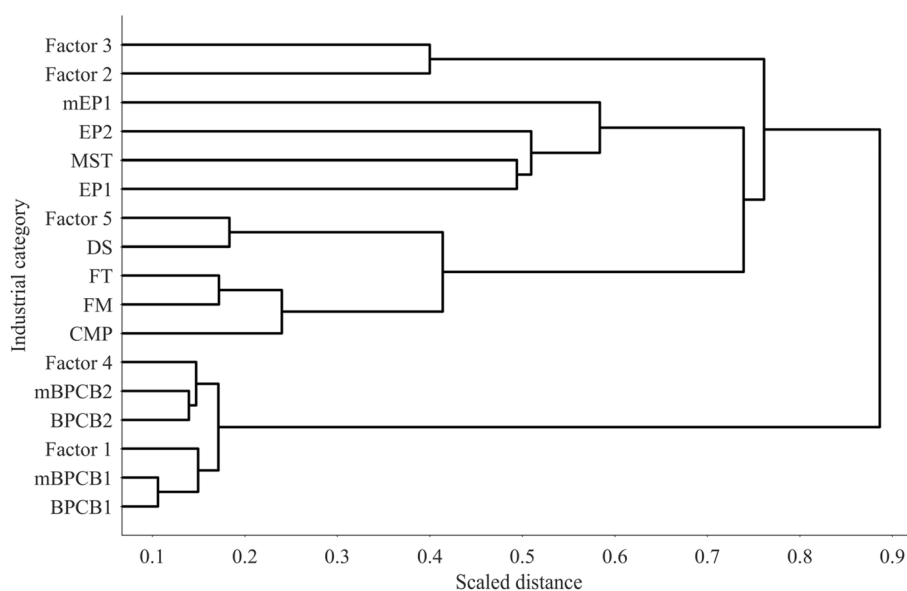


Fig. 6 Cluster dendrogram of hierarchical classification

3.4 MI apportionment

These analyses in the previous section have focused on source apportionments for mass concentrations. From the perspective of protecting human and ecological health, further quantifying the contributions of sources due to MI, which considered both the concentrations and MAC of the main toxic elements, is needed. The source-specific MI of the p th source (MI_p) was calculated using

$$MI_p = \sum x_{jp}^* MI_j \quad (8)$$

where x_{jp}^* is the mean concentration of species j in source p , based on the model results, and MI_j is the unit MI for species j .

The averaged MI was 53.9, with Cu contributing up to 77%. The other highly polluting species was Mn (10%), followed by Ni (6%). Figure 5c presents the MI apportionment results. Factor 4, which was classified as BPCB2, was the main MI contributor (53%) at S4. The second highest contributor (24%) was Factor 1, which was classified as BPCB1. The sum of MI contributions from BPCB1 and BPCB2 was up to 77%, indicating that BPCB was the industrial category with the greatest health risk at S4.

Figure 3 presents the time series of the measured MI scores and those of the two largest contributors. To compare the source contribution during pollution events and nonevents, high pollution events (HPEs) were defined as those with $MI > 99.3$ (i.e., mean MI plus one standard deviation), and other events classified as

general conditions (GCs). Nine HPEs were identified. The averaged MI of the HPEs was 143, whereas the averaged MI of the GCs was 38. The health risk of the HPEs was approximately 3.8 times greater than that of the GCs. For Events 2, 4, 6, 7, 8, and 9, BPCB2 was the main MI contributor; for Event 1, BPCB1 was the main contributor. By contrast, the pollutants of Event 5 were almost equally contributed by both BPCB1 and BPCB2. For Event 3, no obvious contributor was identified, which was also evidenced by the PMF model not fitting this event well. Overall, BPCB1 and BPCB2 were the main MI contributors to HPEs; the sum of their MI contributions was up to 91% (Fig. 7). During GCs, BPCB1 and BPCB2 were also the largest health risk contributors, and the sum of their contributions was 79%. DS contributed the second greatest health risk, with a 15% contribution. Although DS was not a dominant source pollution in HPEs, it was a stable source pollution that contributed 9% of the averaged MI.

3.5 Limitations

In this study, the MI was considerably higher than the reference value because of stringent MAC of Cu in Taiwan. The guideline for drinking water quality published by the World Health Organization [31] is $2000 \mu\text{g L}^{-1}$ for Cu, but the MAC of Cu is $30 \mu\text{g L}^{-1}$ in Taiwan. This makes estimates of the MI higher if Cu is the main pollutant. Nevertheless, establishing stringent a standard for Cu can protect aquatic biota [32, 33] and also human beings who are at the top of the food chain. Another limitation

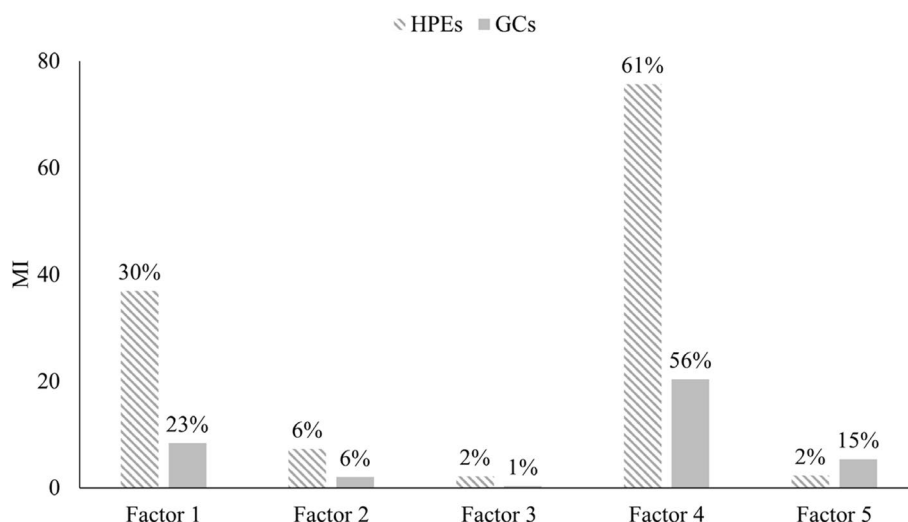


Fig. 7 MI apportionment in high pollution events (HPEs) and in general conditions (GCs)

is that the result estimated by the PMF model provided only information on source contributions by industrial categories and not the exact source locations. However, this modeling approach is valuable in narrowing down pollution sources to specific industrial categories.

4 Conclusions

This study investigated the source contributions at a river pollution hotspot by using PMF. This was the first study to apply sequential time series data of metal species in PMF to apportion sources and their contributions by industrial category. The results revealed that the main source was BPCB, and the sum of BPCB1 and BPCB2 contributions was up to 55% of the total elemental concentrations and 92% of copper in the study area. Because Cu element was the main pollutant in BPCB which was the main contributor, this leads to severe copper pollution in Tandigou. BPCB contributed 75% of total elemental concentrations in MI. For both elemental concentrations and the MI, which is related to human and ecological health, BPCB was the main contributor. In further investigation of the source pollution in HPEs, the MI contribution from BPCB was up to 91%. The results indicated that BPCB contributed toxic pollutants and that led to a high percentage of MI contribution. Although the waste water samples showed that the discharge for each factory monitored in this study followed the effluent standard, the overall influence on the Cu concentration in river water might still be enormous. Implementing comprehensive control strategies, such as tightening the effluent regulations, centralizing the water treatment system, and enhancing monitoring are warranted, especially for the BPCP industry.

In this study, we demonstrated the feasibility of using metal species with a receptor model to identify source pollution. The model results might provide useful information to policymakers for the effective management of river environments. For further investigation of river pollution, establishing a comprehensive source profile database is highly recommended for source identification.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s42834-022-00144-9>.

Additional file 1.

Acknowledgments

We thank our co-workers from the National Institute of Environmental Analysis, Environmental Protection Administration of Taiwan who greatly assisted and executed this project. We would also like to show our gratitude to the samplers and analysts for their hard work.

Authors' contributions

Peo-Yuan Hsieh: Conceptualization, Methodology, Software, Formal Analysis, Investigation, Writing – original draft, Visualization. Huan-Chun Lin: Investigation, Writing - review & editing. Gen-Shuh Wang: Conceptualization, Supervision, Project administration. Yuan-Jeng Hsu: Investigation, Data curation, Resources, Supervision, Project administration. Yi-Ju Chen: Resources, Data curation. Tzu-Hui Wang: Resources, Data curation. Ren-De Wang: Resources, Data curation. Chun-Yu Kuo: Resources, Data curation. Di-Wen Wang: Resources, Data curation. Ho-Tang Liao: Validation, Writing - review & editing. Chang-Fu Wu: Conceptualization, Supervision, Project administration, Writing - review & editing. The authors read and approved the final manuscript.

Funding

This work was funded by the Environmental Analysis Laboratory, Environmental Protection Administration of Taiwan (EPA-106-S3E4-02-05), and the Featured Areas Research Center Program within the framework of the Higher Education Sprout Project by the Ministry of Education of Taiwan (NTU-109 L9003).

Availability of data and materials

Restrictions apply to the availability of these data, which were used under license for this study. The data are available from the corresponding author with the permission of Environmental Analysis Laboratory, Environmental Protection Administration, Taiwan.

Declarations

Competing interests

The authors declare that they have no competing interests.

Author details

¹Institute of Environmental and Occupational Health Sciences, College of Public Health, National Taiwan University, Taipei 10055, Taiwan. ²Environmental Analysis Laboratory, Environmental Protection Administration, Taoyuan 32024, Taiwan.

Received: 13 January 2022 Accepted: 12 June 2022

Published online: 29 June 2022

References

- UNDESA. The Sustainable Development Goals Report 2017. New York: United Nations Department of Economic and Social Affairs; 2017.
- Ott WR. Water Quality Indices: A Survey of Indices Used in the United States. Washington, DC: US Environmental Protection Agency; 1978.
- Kumar D, Alappat BJ. NSF-water quality index: does it represent the experts' opinion? *Pract Period Hazard Toxic Radioact Waste Manage*. 2009;13:75–9.
- Horton RK. An index number system for rating water quality. *J Water Pollut Con F*. 1965;37:300–6.
- Brown RM, McClelland NI, Deininger RA, Tozer RG. A water quality index: do we dare? *Water Sewage Works*. 1970;117:339–43.
- Lumb A, Sharma TC, Bibault JF. A review of genesis and evolution of water quality index (WQI) and some future directions. *Water Qual Expos Hea*. 2011;3:11–24.
- Abdo MH, El-Nasharty SM. Physico-chemical evaluations and trace metals distribution in water-surficial sediment of Ismailia Canal, Egypt. *Nat Sci*. 2010;8:198–206.
- Ali MHH, Fishar MRA. Accumulation of trace metals in some benthic invertebrate and fish species relevant to their concentration in water and sediment of Lake Qarun, Egypt. *Egypt J Aquat Res*. 2005;31:289–301.
- Goher ME, Hassan AM, Abdel-Moniem IA, Fahmy AH, El-sayed SM. Evaluation of surface water quality and heavy metal indices of Ismailia Canal, Nile River, Egypt. *Egypt J Aquat Res*. 2014;40:225–33.
- Tamasi G, Cini R. Heavy metals in drinking waters from Mount Amiata (Tuscany, Italy). Possible risks from arsenic for public health in the Province of Siena. *Sci Total Environ*. 2004;327:41–51.
- Helena B, Pardo R, Vega M, Barrado E, Fernandez JM, Fernandez L. Temporal evolution of groundwater composition in an alluvial aquifer (Pisuerga River, Spain) by principal component analysis. *Water Res*. 2000;34:807–16.
- Kowalkowski T, Zbytniewski R, Szpejna J, Buszewski B. Application of chemometrics in river water classification. *Water Res*. 2006;40:744–52.
- Palma P, Alvarenga P, Palma VL, Fernandes RM, Soares AMVM, Barbosa IR. Assessment of anthropogenic sources of water pollution using multivariate statistical techniques: a case study of the Alqueva's reservoir, Portugal. *Environ Monit Assess*. 2010;165:539–52.
- Huang JL, Ho MH, Du PF. Assessment of temporal and spatial variation of coastal water quality and source identification along Macau peninsula. *Stoch Env Res Risk A*. 2011;25:353–61.
- Khuman SN, Chakraborty P, Cincinelli A, Snow D, Kumar B. Polycyclic aromatic hydrocarbons in surface waters and riverine sediments of the Hooghly and Brahmaputra Rivers in the Eastern and Northeastern India. *Sci Total Environ*. 2018;636:751–60.
- Paatero P, Tapper U. Positive matrix factorization: a nonnegative factor model with optimal utilization of error-estimates of data values. *Environmetrics*. 1994;5:111–26.
- Belis CA, Karagulian F, Larsen BR, Hopke PK. Critical review and meta-analysis of ambient particulate matter source apportionment using receptor models in Europe. *Atmos Environ*. 2013;69:94–108.
- Hopke PK. Review of receptor modeling methods for source apportionment. *J Air Waste Manage*. 2016;66:237–59.
- Soonthornnonda P, Christensen ER. Source apportionment of pollutants and flows of combined sewer wastewater. *Water Res*. 2008;42:1989–98.
- Li HY, Hopke PK, Liu XD, Du XM, Li FS. Application of positive matrix factorization to source apportionment of surface water quality of the Daliao River basin, northeast China. *Environ Monit Assess*. 2015;187:80.
- Jiang JP, Khan AU, Shi B, Tang SJ, Khan J. Application of positive matrix factorization to identify potential sources of water quality deterioration of Huaihe River, China. *Appl Water. Sci* 2019;9:63.
- Zanotti C, Rotiroli M, Fumagalli L, Stefania GA, Canonaco F, Stefanelli G, et al. Groundwater and surface water quality characterization through positive matrix factorization combined with GIS approach. *Water Res*. 2019;159:122–34.
- Vu CT, Lin C, Shern CC, Yeh G, Le VG, Tran HT. Contamination, ecological risk and source apportionment of heavy metals in sediments and water of a contaminated river in Taiwan. *Ecol Indic*. 2017;82:32–42.
- Bzdusek PA, Christensen ER. Comparison of a new variant of PMF with other receptor modeling methods using artificial and real sediment PCB data sets. *Environmetrics*. 2006;17:387–403.
- Bzdusek PA, Lu JH, Christensen ER. PCB congeners and dechlorination in sediments of Sheboygan River, Wisconsin, determined by matrix factorization. *Environ Sci Technol*. 2006;40:120–9.
- Hemann JG, Brinkman GL, Dutton SJ, Hannigan MP, Milford JB, Miller SL. Assessing positive matrix factorization model fit: a new method to estimate uncertainty and bias in factor contributions at the measurement time scale. *Atmos Chem Phys*. 2009;9:497–513.
- TEPA. Sampling Methods for Industrial Effluent (NIEA W109). Taipei: Environmental Protection Administration; 2019 [in Chinese]. <https://www.epa.gov.tw/niea/C3090125D2C84BD8>.
- TEPA. Method for the Determination of Metals and Trace Elements in Water by Inductively Coupled Plasma Mass Spectrometry (NIEA W313). Taipei: Environmental Protection Administration; 2019 [in Chinese]. <https://www.epa.gov.tw/niea/2BD0287A7CA402E7> (in Chinese).
- Norris G, Duvall R, Brown S, Bai S. EPA Positive Matrix Factorization (PMF) 5.0 Fundamentals and User Guide. Washington, DC: US Environmental Protection Agency; 2014.
- Lee E, Chan CK, Paatero P. Application of positive matrix factorization in source apportionment of particulate pollutants in Hong Kong. *Atmos Environ*. 1999;33:3201–12.
- WHO. Guidelines for drinking-water quality: fourth edition incorporating the first addendum. Geneva: World Health Organization; 2017.
- Shuhaimi-Othman M, Nadzifah Y, Nur-Amalina R, Umirah NS. Deriving freshwater quality criteria for copper, cadmium, aluminum and manganese for protection of aquatic life in Malaysia. *Chemosphere*. 2013;90:2631–6.
- Zhang YH, Zang WC, Qin LM, Zheng L, Cao Y, Yan ZG, et al. Water quality criteria for copper based on the BLM approach in the freshwater in China. *PLoS One*. 2017;12:e0170105.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.