Ecotoxicology and Environmental Safety 113 (2015) 469–476

Contents lists available at ScienceDirect

Ecotoxicology and Environmental Safety

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

Characterization of heavy-metal-contaminated sediment by using unsupervised multivariate techniques and health risk assessment

Yeuh-Bin Wang [a,b], Chen-Wuing Liu [a,*], Sheng-Wei Wang [c]

[a] Department of Bioenvironmental Systems Engineering, National Taiwan University, Taipei, Taiwan
[b] Department of Environmental Monitoring and Information Management, Environmental Protection Administration, Taipei, Taiwan
[c] Sinotech Environmental Technology, Ltd., Taipei, Taiwan

ARTICLE INFO

Article history:
Received 17 October 2014
Received in revised form
19 December 2014
Accepted 22 December 2014
Available online 5 January 2015

Keywords:
Sediment quality management
Self-organizing map (SOM)
Positive matrix factorization (PMF)
Health risk assessment

ABSTRACT

This study characterized the sediment quality of the severely contaminated Erjen River in Taiwan by using multivariate analysis methods—including factor analysis (FA), self-organizing maps (SOMs), and positive matrix factorization (PMF)—and health risk assessment. The SOMs classified the dataset with similar heavy-metal-contaminated sediment into five groups. FA extracted three major factors—traditional electroplating and metal-surface processing factor, nontraditional heavy-metal-industry factor, and natural geological factor—which accounted for 80.8% of the variance. The SOMs and FA revealed the heavy-metal-contaminated-sediment hotspots in the middle and upper reaches of the major tributary in the dry season. The hazardous index value for health risk *via* ingestion was 0.302. PMF further qualified the source apportionment, indicating that traditional electroplating and metal-surface-processing industries comprised 47% of the health risk posed by heavy-metal-contaminated sediment. Contaminants discharged from traditional electroplating and metal-surface-processing industries in the middle and upper reaches of the major tributary must be eliminated first to improve the sediment quality in Erjen River. The proposed assessment framework for heavy-metal-contaminated sediment can be applied to contaminated-sediment river sites in other regions.

1. Introduction

Heavy-metal-contaminated sediment is a severe environmental problem because the toxicity, persistence, and bioaccumulation of heavy metals can cause adverse health risks (MacDonald et al., 2000; Ingersoll et al., 2001). In particular, rapid industrialization and urbanization have resulted in the production of a high amount of industrial wastewater, which is discharged into rivers, causing heavy metals to accumulate in the river sediment (Guo and Chen, 2012; Fu et al., 2013; Jiang et al., 2013). The river sediment is a basic environment for aquatic sediment-dwelling organisms and a sink for discharged contaminants from anthropogenic activities and natural processes. An understanding of the spatial distribution, source apportionment, and the associated health risks of heavy metals in river sediment is critical for regulating the water environment (Comero et al., 2011,2014; Tian et al., 2013).

Erjen River in southwestern Taiwan was the most severely heavy-metal-contaminated area in the 1970s. Numerous electronic waste recovery industries used open burning and acid washing to recover metals from waste electrical circuit boards and cables along the riverside. Heavy metals thus accumulated in the sediment and aquatic organisms (Lee et al., 1996; Chen et al., 2004). The 1986 green oyster incident in the river estuary forced the Taiwanese Environmental Protection Administration (EPA) and local government to initiate several restoration measures to remediate the river. According to the database of the river water quality hosted by the Taiwanese EPA, the metal concentrations in the river water substantially decreased over the past three decades (Taiwan EPA, 2014). A recent study which investigated selected heavy metals in the water, sediment, and oysters in the estuary of the river, indicated that the environment of the estuary has been generally restored (Chen et al., 2014). However, the accumulated heavy metals in the river sediment have not been systematically assessed.

The unsupervised multivariate techniques are efficient tools in extracting meaningful, valuable, and precise information from environmental monitoring datasets (Tsakovski et al., 2011; Gredilla et al., 2013). Among such tools, factor analysis (FA) can be used to extract the latent factors that dominate the spatial distribution of heavy metals in the water environment from large monitoring datasets by plotting a factor-score map (Liu et al., 2003). FA can

* Corresponding author. Fax: +886 2 23639557.
E-mail address: cwliu@ntu.edu.tw (C.-W. Liu).

http://dx.doi.org/10.1016/j.ecoenv.2014.12.036
0147-6513/

also be applied in identifying the extent and magnitude of the contaminated hotspots with the elevated factor scores. FA results can assist researchers in suggesting remediation priorities or improvement measures for revamping contaminated sites (Riba et al., 2002; Rodriguez-Barroso et al., 2010; Casado-Martinez et al., 2009; Moller and Einax, 2013). A self-organizing map (SOM) is a type of artificial neural network method. SOMs differ from FA based on the linear relationships among the variables of the monitoring dataset. SOMs can be used to determine the similarity between the monitoring dataset and identify the contaminated hotspots as precedence for control measures (Alvarvez-Guerra et al., 2008; Tsakovski et al., 2011; Veses et al., 2013, 2014). Moreover, to apply effective control measures to the major pollution sources in contaminated hotspots, the sources apportionment of the major pollutants must be estimated. To protect the environment, the major pollution sources must be effectively controlled by quantifying the apportionment of health risk to the corresponding major pollution sources. The positive matrix factorization (PMF) technique is an advanced factor analysis method which has been recently used in the field of source apportionment of heavy metals in sediment (Chen et al., 2013; Pekey and Dogan, 2013; Comero et al., 2011, 2014). If the source apportionment of pollutants in sediment is identified, the apportionment of health risk caused by the corresponding major pollution sources can be predicted and effective remediation measures can be readily implemented (Tian et al., 2013). However, a few detailed studies have combined risk assessment and PMF for the apportionment of health risks caused by heavy-metal pollution.

This study first assessed the spatial characteristics of heavy-metal-contaminated sediment to determine the contaminated hotspots in Erjen River by using FA and SOM. PMF was then used to estimate the apportionment of major sources and associated health risks from the identified sources. Finally, control measures were suggested for further improving the sediment quality of the river.

2. Materials and methods

2.1. Study area

Erjen River is located in southwestern Taiwan, and it originates from Shanchuhu Lake at an altitude of 460 m and discharges into the Taiwan Strait. Fig. 1 shows the location of the river. The mainstream of the river is 61.2 km long, and the watershed area of the river is approximately 339.2 km^2. The downstream of the river is between the cities of Tainan and Kaohsiung, and Sanyegong Creek is the main tributary of the river. The regional climate is subtropical, with an annual average precipitation of 1700 mm and average temperature of 24.3 °C. The river water serves as the major water source for agricultural activities, fisheries, and ecological conservation.

In the 1970s, several polluting activities—such as open burning and acid washing of waste electrical circuit boards and cables to recover valuable metals—were conducted along the downstream riverbank. These activities also led to the discharge of various toxic substances into the river. The severely contaminated river had a notorious name: “the black dragon river” of Taiwan (Lee et al., 1996; Fu and Wu, 2006). To remediate Erjen River, the environmental authority implemented several measures, including (1) prohibiting the importation of waste electrical circuit boards and cables; (2) banning open waste burning; (3) enforcing stricter factory effluent standards and permit-management systems; (4) demolishing aluminum-smelting plants; and (5) removing abandoned waste electrical circuit boards from the riverbank. The current pollution status of the river water has been progressively improved (Chen et al., 2004; Taiwan EPA, 2014). However, the heavy-metal content in the river sediment remains unclear and the sediment quality has not been thoroughly assessed.

2.2. Data collection and processing

The Taiwanese EPA established 26 sampling sites along the mainstream of the Erjen River and its most vital tributary, Sanyegong Creek, to survey sediment quality from 2010 to 2012. Sediment samples were collected yearly in the dry and rainy seasons. At each sampling location, three samples were collected at 0–10 cm depth from the riverbed, mixed together and carefully preserved until laboratory processing. Eight heavy metals—namely, arsenic (As), mercury (Hg), cadmium (Cd), chromium (Cr), copper (Cu), nickel (Ni), lead (Pb), and zinc (Zn)—were analyzed in a Taiwan EPA-certified laboratory—Industrial Technology Research Institute. This study used the discussed sediment quality survey dataset provided by the Taiwan EPA.

Table S1 lists the descriptive statistics of the raw monitoring datasets. Most variables did not follow a normal distribution. To avoid analytical errors caused by differences in the variable unit or range, the raw datasets were standardized before SOM and FA were applied (Vesanto et al., 2000; Alvarvez-Guerra et al., 2008).

2.3. Analysis methods

The multivariate analysis methods—including factor analysis (FA), self-organizing maps (SOMs), and positive matrix factorization (PMF)—and health risk assessment were applied to characterize the sediment quality of the Erjen River. The flowchart of assessment framework for heavy-metal-contaminated sediment is shown in Fig. S1.

2.3.1. SOM

SOMs represent an unsupervised artificial neural network technique designed for the 2D visualization and analysis of multi-dimensional datasets without the rigid linearity assumptions of traditional statistical methods (Alvarvez-Guerra et al., 2008; Tsakovski et al., 2011). SOMs classify datasets according to their similarity through a training process. The learning algorithm applied in a standard SOM is divided into six steps (Peeters et al., 2006). In summary, the SOM network construction comprises an input layer and a layer of neurons, which are fully interconnected. Each of the input vectors is also connected to each neuron. Each neuron is represented by a weight vector. The weight vectors of the SOM were first initialized to random values in this study. Furthermore, the learning algorithm iteratively calculates and compares the Euclidean distances between each data vector and all the weight vectors of the SOM, and the so-called best-matching unit (BMU) or neuron whose weight vector exhibits the minimum distance is obtained. The SOM weight vectors are updated so that the BMU and its topological neighbors are moved toward the input vector. After the learning iterations, the BMUs of similar datasets will be close to each other on the final SOM obtained (Vesanto et al., 2000; Alvarvez-Guerra et al., 2008; Coz et al., 2008). In this study, SOM Toolbox 2.0 for Matlab (Vesanto et al., 2000) was used to perform the SOM analysis.

2.3.2. Factor analysis

FA provides the general relationship between measured variables by elucidating the multivariate patterns that might facilitate classifying the original data. FA can be used to determine the geographical distribution of resultant factors. The interpretation of factors may provide insight about the main sources that govern the distribution of chemical variables (Liu et al., 2003). The data are transformed into factors, and this study only retained factors with eigenvalues that exceed 1 (Reyment and Joreskog, 1993). The contribution of each factor (factor score) at each sampling site in