國立交通大學

環境工程研究所

碩士論文

應用多方向量分析分析污染檢測數據

The Analyses of Sampling Data Using Polytopic Vector Analysis

研究生：莊敏筠

指導教授：葉弘德 教授

中華民國九十六年八月

應用多方向量分析分析污染檢測數據

The Analyses of Sampling Data Using Polytopic Vector Analysis

研究生：莊敏筠　　Student：Min-Yun Chuang

指導教授：葉弘德　　Advisor：Hund-Der Yeh

國立交通大學

環境工程研究所

碩士論文

A Thesis
Submitted to Institute of Environmental Engineering
College of Engineering
National Chiao Tung University
in Partial Fulfillment of the Requirements
for the Degree of
Master of Science
in
Environmental Engineering
August, 2007
Hsinchu, Taiwan

中華民國九十六年八月

應用多方向量分析分析污染檢測數據

研究生：莊敏筠　　　　指導教授：葉弘德

國立交通大學環境工程研究所

摘要

鑒定化學污染源與瞭解污染團在時空上的分佈，在環境污染領域是重要且新受矚目的問題。在多重污染源地區，環境工程師面臨著如何鑒定複雜地形的污染源型態。如果污染源具有獨特的化學指紋，即可藉由多變數統計分析方法來進一步解決環境指紋問題。

多方向量分析(Polytopic Vector Analysis)是一種鑑定污染源化學指紋的多變量分析統計方法。假定某一混合系統，多方向量分析可用來計算三個重要參量：(1)來源(後端成分)個數、(2)各後端成分的化學指紋組成、以及(3)不同樣本中來自各來源的濃度相關比例。

本研究應用多方向量分析，分別針對兩個研究案例做分析(I)客雅溪流域重金屬檢測數據，以及(II)新竹科學園區地下水有機污染檢測資料。

近年來，在新竹香山海域養殖的牡蠣及沈積物的銅含量皆出現顯著增高的趨勢。綠牡蠣體內伴隨著高濃度的銅含量，不但威脅消費者健康且影響漁民每年三百多萬美金產值。香山牡蠣銅污染的來源是長久被關切且極需解決的問題。雖然在過去的研究曾指出香山海域銅污染可能源自於新竹科學園區或香山工業區，然而卻缺少強而有力的證據指出確切的污染者。於 2006 年檢測客雅溪及其支流的十二組數

據，這十二組數據分別來自於不同兩組研究團隊的檢測成果。本研究目的在於利用統計方法-多方向量分析，鑑定綠牡蠣組織內的重金屬銅之來源。第一組及第二組數據，分別在客雅溪沿岸檢測共6個水樣及7個懸浮固體樣本，採樣時間為2006年四月，每個取樣分析18種重金屬項目(鋅、銅、鉛、鐵、鋁、錳、鎘、鍶、鉻、鋇、鎳、銀、錫、砷、釩、鎢、鎵及鉬)。第三組到第八組數據，沿客雅溪流域7個取樣點，採樣時間為2006年二月至七月；第九組到第十二組數據，沿客雅溪流域8個取樣點，採樣時間於2006年八月至十一月，每個水樣分析六個項目(懸浮固體、溶氧、生化需氧量、氨氮、銅及砷)。由多方向量分析結果顯示，新竹科學園區為綠牡蠣中銅的主要來源。

在第二個研究案例，係應用多方向量分析針對新竹科學園區地下水污染檢測資料作分析，利用計算係數以及因子負荷係數判定該工業區地下水有機污染物資料，鑒定出有六個有機污染物指紋，進而計算出各後端成分在地理上的分佈。因此，利用多方向量分析可決定後端成分的個數，並且發展出一個地下水污染物的混合模式，以模擬地下水污染源化學指紋與污染團在空間上的分佈。

The analyses of sampling data using Polytopic vector analysis

Student: Min-Yun Chuang **Advisor: Hund-Der Yeh**

Institute of Environmental Engineering
National Chiao Tung University

ABSTRACT

Identification of chemical contaminant sources is an important environmental problem. Given multiple sources in a field area, the environmental engineer is faced with the challenge of identification and mapping multiple plumes with overlapping geographic distribution. If a contaminant source is characterized by a distinctive spectrum of chemicals, then the environmental chemical fingerprinting problem may be addressed through a multivariate statistical approach.

Polytopic vector analysis (PVA) is a statistical pattern recognition technique for multivariate data used to identify fingerprints of contaminant sources. Given a mixed system, PVA is used to determine three parameters of interest in a mixed system: (1) the number of sources (end-members), (2) the composition of each source, and (3) the relative proportions of each source in each sample.

In this study, polytopic vector analysis is applied to analyze two problems. In the first problem, the copper source of green oyster in Hsianshan wetland is identified while in the second problem the groundwater contamination by organic chemicals in Hsinchu science park is analyzed.

In recent two decades, the copper concentrations in both the oyster organs and sediment were very high in the Hsianshan coastal area, Taiwan. The oyster with high concentration of copper poses not only a threat to human health but also results in an annual loss of about 3.1 million US dollars. What is the source of copper in the Hsianshan oyster is a question of long standing and tough problem to be solved. Previous studies indicated that the copper source originated from either the Hsianshan industrial park (HIP) or the Hsinchu Science Park (HSP) is responsible for the contamination in Hsianshan coastal area. Although several investigations had been conducted for identifying the copper source of green oyster; however, there was no clear evident or result to show who is responsible for the copper pollution. In order to search for the source of copper, the water and suspended solid samples were collected and analyzed. Water samples mainly collected from the Keya stream and some suspected source locations in 2006. Samples of suspended solid were also

collected from the Keya stream and coastal area of Hsianshan. Totally 12 sets of water and suspended solid sample data were available. The Data Set 1 contains water samples taken from 6 different sites along the Keya stream, Hsinchu while the Data Set 2 has suspended solid samples taken from the same stream at 7 different sites. Both two data sets are analyzed for eighteen heavy metals including Zn, Cu, Pb, Fe, Al, Mn, Cd, Sr, Cr, Ba, Ni, Ag, Sn, As, V, W, Ga, and Mo. The Data Sets 3 to 8 were sampled at 7 locations along the Keya stream during February to July, 2006. The Data Sets 9 to 12 were collected at 8 locations, where 7 locations were the same as those taken above and one extra location was at Yuchegou creek, during August to November, 2006. Those ten data sets, Data Sets 3 to 12, were taken and analyzed for six items (SS, DO, BOD, NH_3-N, Cu, As). The purpose of this study is to identify the copper source based on those sample data using a multivariable statistical analysis called polytopic vector analysis (PVA). The results of PVA indicate that a significant part of copper comes from HSP and a minor part of copper is from Yuchegou creek.

In the second problem, the study applies PVA to analyze the concentration data taken from HSP, Taiwan. Based on inspecting the coefficient of determine (CD scatter) and the factor loadings index, six separate organic compound fingerprints are

identified. Moreover, the geographic distribution of each end-member can be presented. Thus, this method (PVA) can develop a groundwater mixing model by searching for end-members. In addition, this mixing model can identify chemical contaminant sources and map multiple plumes with overlapping geographic distributions.

致謝

這本碩士論文，我要獻給我最愛的父母親。感謝您們不辭辛勞的工作，就為了給我們最好的教育環境。也因為您們的教誨，讓我有向上的動力。

能夠完成這本論文，首先要感謝我的指導教授葉弘德老師，老師總能在適時提供我論文的方向，並細心的指正論文的每個細節，才能讓論文有今天這麼完整的面貌。在老師那邊也學習到很多做研究的謹慎的態度，更獲得許多寶貴的經驗。另外，也要感謝海洋大學的陳明德教授，教導統計觀念及提供軟體讓我能突破研究最開始的瓶頸。再來要非常感謝我的可愛的GW Group-智澤學長、彥禎學長、雅琪學姊、彥如學姊、士賓、博傑、其珊、仲豪、珖儀及跟我一起打拼兩年的隊長毓婷！謝謝你們平時的鼓勵及照顧，我們一起登玉山、遊司馬庫斯的美好經驗，相信永遠都會存在我們的心中！

再來我要特別感謝，這六年來一路陪伴我的瑞元，在最重要的時刻你總是給我最棒的精神鼓勵及支持，讓我有勇氣突破！

希望我的致謝能夠真實的傳達我最真摯的感謝。

在交大的六年，如今要畫下一個句點。但，這個句點代表著下個人生階段的起點

敏筠　謹致於

交通大學環境工程研究所

2007年8月

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

The main subject of thesis is to identify the source of contamination using the method of polytopic vector analysis (PVA). PVA is a statistical pattern recognition technique for multivariate data used to identify fingerprints of contaminant sources. This thesis is to study two problems. The first problem is to identify the source of green oyster in Hsianshan wetland while the second problem is to analyze groundwater contamination by organic chemicals in Hsinchu science park (HSP).

PROBLEM I

Green Oyster in Hsianshan wetland

CHAPTER 1 INTRODUCTION

1.1 Hsinchu city

Hsinchu city, known as wind city, is situated in the northwest part of Taiwan, comprising an area of 104 km^2 and a population of about 390,000 people in 2007. The spring and summer are rainy seasons and the average annual precipitation and temperature are about 1,780 mm and 22.7^{o}C, respectively. During summer seasons, it is full of rain with monsoon climate. The northeasterly monsoon prevails from September to May in the Hsinchu area. The rainfall occurred often with typhoon generally reaches the highest during June and September and then significantly reduces afterward.

The Hsinchu plain is mainly formed by the alluvium deposit of the Touchien river, which is the largest stream in the Hsinchu area and has a total length of 63.4 km and a drainage area of 565.97 km^2. Six rivers in Hsinchu area are mostly originated from the east side of mountain ridge to the west of Taiwan strait. Figure 1 shows that from the north to the south they are Touchien river, Keya stream, Sanxingong stream, and Yengang stream. In addition, two creeks, Nanman creek and Yuchegou creek, are the major tributaries of Keya stream. The Keya stream is the second largest stream within the Hsinchu city. The total length of Keya stream is about 26.6 km, covering a basin area of 45.6 km^2. The stream receives the agricultural wastewater in the upstream reach and the wastewater from HSP and municipal wastewater from commercial and residential areas in the middle reach. In the downstream reach, the stream receives the wastewater from the agricultural activity and HIP and the municipal wastewater from residential area as well. The estimated flow rate of Keya stream ranges from 0.91 to 1.58 m^3/sec (BEPH, 2006).

The city of Hsinchu was classified into three districts namely East, North, and Hsianshan districts in 1982. The Hsianshan district is located in the west of Hsinchu city as shown in Fig. 1. In fact, it is a costal village with an area about 55 km^2 and a population of 69,300. On the west of Hsianshan district, there is a gently sand coast comprising shoal, swamp, and wetland which is a famous coast fishery aquaculture area in Taiwan. On the other hand, there is the Hsianshan Industry Park (HIP) situated on the east of Hsianshan district.

1.2 Hsinchu science park and Hsianshan industry park

Hsinchu science park (HSP) located at the eastern Hsinchu city occupies 658 hectares. The HSP according to its land development can be divided into three parts called the 1st, 2nd, and 3rd periods of the park as illustrated in Fig.1. The development of industry in the HSP is rapidly increased in high technology within recent years. At present, the industry in HSP, like Silicon Valley in California, includes semiconductor manufacturing companies, computer and peripherals, telecommunications, optoelectronics, precision machinery, and biotechnology, etc. Among those categories, the major industry is semiconductor manufactures. By the end of December 2004, there were 164 integrated circuit companies in the Park, with total sales revenue of US$22,309 million, which represented growth of 32% from 2003. Regarding IC design, 21 new IC design companies were founded in the park in 2004. These companies focused on designing ICs for computers and peripherals, telecommunications, and consumer electronic appliances. The raw materials in wafer process of semiconductor manufactures are inorganic metal, inorganic acid/alkali, developer, photoresist, etc. At present, the highest water consumed in the HSP is about 175,000 ton/day. The HSP wastewater discharge point is located at Nanman creek, a tributary of Keya stream also shown in Fig.1, and about 35 m upstream form the confluence of Nanman creek and Keya stream. The water in Keya stream in non-rainy days is mainly from the untreated municipal wastewater and the effluent of about 86,000 ton/day produced from the HSP sewage treatment plant.

Hsianshan industry park (HIP) located within the Hsianshan district is in the west of Hsinchu city. The major industry in HIP includes plastic and paper manufacturers, glass factory, electroplating plant, iron and steel plants, and machining industry, etc.

The major contaminants produced from those factories may include inorganic acid and some heavy metals such as Cu, Zn, Mn, Cr, Fe, Pb, Cr, Ba, Sn, and Hg. The wastewater of HIP is mainly discharged into Yuchegou creek which is a tributary of Keya stream. In addition to the wastewater produced from HSP and HIP, it is known that some factories such as electroplating plant distributed in the drainage area of Keya stream also discharges industrial wastewater into Keya stream. The industrial wastewater discharged to the Keya stream flows southward after entering the sea and migrates to the Hsianshan coastal area by the alongshore current driven by northeasterly monsoon in the winter and spring seasons.

1.3 Hsianshan wetland and green oysters

The area of Hsianshan wetland, located right at the south of the Keya estuary, is the largest intertidal zones at the north part of Taiwan. The environment in the shallow waters along the coast is very suitable for the aquaculture of shellfishes. The aquaculture area in Hsianshan intertidal zones is about 342.5 hectare. The major shellfish in aquaculture are oysters and clams. In fact, the oyster field here has a history of more than 100 years, and it provides a spectacular scene in this area. The cultured oyster species at the Hsianshan wetland is Crassistrea *gigas*. The annual production rate of the oyster was about 400 tons in the past and drops to about 50 tons in 2004 and 16 tons in 2006. Accordingly the estimated loss in the profit of Hsianshan oyster was up to 3.1 million US dollars annually.

Environmental investigations in the Hsianshan wetland have been made for more than 20 years. The impact of copper on the wetland ecology has been the focus of the investigations (BEPH, 2006; Lin et al., 1999; Huang et al., 2006). The problem of green oyster occurred primary as result of high copper concentration in oyster organs. In fact, high copper concentrations were observed in both the oyster organs and sediments along the Hsianshan coastal area. Table 1 shows the statistics of the copper concentrations of oyster and sediment in Hsianshan area sampled from 1987 to 2006. The table shows that the Hsianshan oyster had copper concentration with average values ranging from 131 ppm to 790 ppm in the year from 1987 to 2006 except in 1998 which had an extremely high copper concentration of 2070 ppm. Note that the copper concentrations reached 5000 ppm in 1997 and 4750 ppm in 1998 in some oyster samples. It is interesting to point out that the Cu concentration in Siensan oyster has high Cu concentration in the winter and spring and low Cu concentration in the summer, which is the rainy season in Taiwan. The reason for this phenomenon may be attributed to the ingesting way of thin oysters in winter (BEPH, 2006). In the wintertime of Taiwan, the sea sediment is disturbed by powerful northeasterly monsoon. Consequently, the suspended sediments with high heavy metal concentration mixed with the planktons are consumed by the oyster through filtering ingestion. Long term ingestion of those substances with high Cu concentration may result in a significant effect of bioaccumulation.

The regulation standards for Cu concentration in aquatic product are 100ppm and 30 ppm in Canada and Australia, respectively. Obviously, the problem of copper pollution in Hsianshan oyster poses a serious threat to the local fishing community and human health (Lin et al., 1999). Han et al. (1998) pointed out that the copper, zinc, and arsenic concentrations in oysters were significantly higher than those in other seafood in Taiwan. They also mentioned that those metals in the seafood possibly comes from the wastewater discharged from various factories in the Keya stream basin and the impact of metal pollution in the seafood and the health risk from consuming the contaminated seafood desires public attention and further assessment (Han et al., 1998).

1.4 Oysters character

There are a number of different groups of mollusks including oyster and shell which grow in marine or brackish water. Some of these groups are highly prized as food, both raw and cooked. Oysters are filter-feeders that draw water in over their gills through the beating of cilia. Suspended food plankton and particles are trapped in the mucus of the gills and transported to the mouth, where they are eaten, digested and expelled as feces or pseudofeces. Feeding activity is greatest in oysters when water temperatures are above 50°F (10°C). Healthy oysters consume algae and other water-borne nutrients, each one filtering up to five litres of water per hour. Today that feeding activity would happen almost a year. Nowadays, the sediment, nutrients, and algae have contamination problems in local waters. Oysters filter these pollutants, and either eat them or shape them into small packets that are deposited on the bottom where they are harmless.

Oysters breathe much like fish, using both gills and mantle. The mantle is lined with many small, thin-walled blood vessels which extract oxygen from the water and expel carbon dioxide. A small, three-chambered heart, lying under the abductor muscle, pumps colorless blood, with its supply of oxygen, to all parts of the body. At the same time two kidneys located on the underside of the muscle purify the blood of any waste products they have collected.

Heavy metal would accumulate in oyster organs. Oyster would get killed while a abnomal water pollution occurred. The routes of copper into oyster organs divided into two ways. The first way is that Cu in the seawater absorbed by algae, and then ingested and accumulated in oysters. Otherwise, the second way is that oyster directly ingests copper from seawater. In general, the appearance of greenish oyster would happen while Cu concentration in oyster organs exceeds 500ppm.

1.5 Two green oyster investigation projects

Recently, there were two investigations in identifying the copper source for Hsianshan green oyster. In April 2006, Huang et al. took water and sediment samples from HSP and Keya stream for analyzing heavy metal concentrations (Huang et al., 2006). In addition, they also measured the heavy metal concentrations in oysters near Keya estuary for identifying the potential copper source in the Hsianshan oyster. Table 2 shows the relative distance from the confluence of Nanman creek to Keya stream or the locations of these 7 sampling points. The sampling point P is inside the 2nd period of HSP for which the samples are pretreatment wastewater mostly discharged from semiconductor manufacturing companies before entering the wastewater treatment plant of HSP. The location of N_0 is right at the discharge point of the HSP at Nanman creek. The sampling point S is located at a municipal sewer in the eastern part of Hsinchu city where the municipal wastewater is going to enter the Keya stream. The sampling points K_{d2} and K_{d4} are located at Keya bridge and Shiangya bridge, respectively, which are both at the downstream of the Keya stream. Note that these two bridges are situated at about 710 m and 8,900m, respectively, from the downstream of the confluence of Nanman creek and Keya stream. The sampling point I is at the intertidal zone of Keya stream. The sampling point C is at the coastal area of Hsianshan. The water samples taken from those six sampling points were analyzed for 18 heavy metals, including Zn, Cu, Pb, Fe, Al, Mn, Cd, Sr, Cr, Ba, Ni, Ag, Sn, As, V, W, Ga, and Mo. The concentration data of those 18 heavy metals, called as Data Set 1 given in (Huang et al., 2006), are graphed in Fig. 3, showing that all the heavy metal concentrations are in ppb level and the Cu concentrations sampled at P, N_0, S, K_{d2}, K_{d4}, and I are 20, 340, 5, 15, 10, and 12 ppm, respectively. The concentration data of Data Set 2 are graphed in Fig. 4, showing that the Cu concentrations sampled at P, N_0, S, K_{d2}, K_{d4}, I and C are 4600, 1400, 200, 500, 300, 250 and 100, respectively.

Huang et al. pointed out that the concentrations of Zn, Cu, Fe, Al, Sr, Ni, Cd, Sn, W, Ga, and Mo at the N_0, the discharge point of HSP, are higher than those at other sampling points (Huang et al., 2006). Particularly, the Cu concentration at N_0 is several tens fold if compared with that at other sampling points, indicating that the HSP is the Cu source for water in the downstream of Keya stream (Huang et al., 2006). They also found very high concentrations of Zn, Cu, Pb, Fe, Al, Sr, Ba, Ni, W, and Ga in the organs of oyster growing in the Hsianshan coast. In addition, the mean concentrations of Zn (2584 ppm), Cu (2366 ppm), Fe (1348 ppm) and Al (1883 ppm) are especially high among those 18 heavy metals. Note that the concentrations of Cu in the sediment ranged from 70 to 275 ppm with an average of 170 ppm are also listed in Table 1. They mentioned that the Cu source of oyster should relate to the high Cu concentration discharged from HSP wastewater and the coastal sediment. In addition to the Cu concentration, they also mentioned that the concentrations of Zn, Fe, Al, Sr, Ni, W, and Ga are also high both in HSP wastewater and oyster, revealing some relationship between the HSP wastewater and contaminated oyster. However, they remarked that they did not have direct evidence to prove that that HSP is the source of heavy metals of the Hsianshan wetland contamination (Huang et al., 2006).

An investigation was also conducted for identifying the copper source of Hsianshan green oyster by the Bureau of Environmental Protection Hsinchu (BEPH) (2006). The water samples were collected at 7 sampling points, N_u, N_0, N_d, K_u, K_{d1}, K_{d3}, and K_{d5} during the period of February to August, 2006 and analyzed for 6 chemical items including SS, DO, BOD, NH_3-N, Cu, As. One additional sampling point Y located at Yuchegou creek was added and its water samples were taken from August to November, 2006. The purpose of adding this sampling point is to assess whether the HIP is a Cu source or not. The concentration data of those 6 chemical items for Data Sets 3 to 12 respectively taken from February to November in 2006 are given in Table 4. Figure 2 shows the sampling locations, where N_u, N_0, N_d are at Nanman creek, Y

is at Yuchegou creek, and $K_{u,}$ K_{d1}, K_{d3}, K_{d5} are at Keya stream. The sampling points N_u and N_d are, respectively, located at the upstream and downstream of the discharge point of HSP wastewater treatment plant at Nanman creek. The K_u is at the upstream and K_{d1}, K_{d3}, K_{d5} are at the downstream of the confluence of Nanman creek and Keya stream. Table 2 shows the relative distances of those 8 sampling points measured from the confluence of Nanman creek and Keya stream.

Most of the Cu concentrations in the sampling points listed in Table 3 are under the environmental standard of 0.03 mg/L in the regulation for protecting human health promulgated by Taiwan Environmental Protection Agency. The study of BEPH pointed out that the Cu concentrations at N_0 and N_d are relatively and equally high if compared with those at N_u except in March and October, 2006. Table 3 shows that, the concentration of Cu at K_{d1}, which is indeed located right at the downstream of N_0 and N_d, is generally higher than those of other the downstream sampling points. It is noticed that the Cu concentrations at N_u in March and October, at K_u in June and October, and at Y in November are relatively high. However, there are no industrial factories in the drainage basins located above the sampling points N_u and $K_{u.}$ Those data may be caused by illegal dumping of chemical wastes which happened sometimes before in different rivers. The Cu concentration detected at Y reflects that the wastewater released from HIP sometime contains heavy metal. The study of

BEPH also mentioned that the Cu concentrations in oyster are higher in the winter and spring and lower in the summer as indicated in Table 1. Note that this study also estimates that the discharged Cu contained in the wastewater of the HSP to Nanman creek is about 1263 kg/yr (BEPH, 2006).

1.6 Polytopic vector analysis

Polytopic vector analysis (PVA) is a hybrid of several algorithms developed over a period of 30 years, primarily within the mathematical geology literature (Imbrie, 1963; Miesch, 1976a; Miesch, 1976b; Klovan et al., 1976; Full et al., 1981; Full et al., 1982; Ehrlich et al., 1987; Evans et al., 1992; Johnson, 1997). It is relatively new in the environmental and chemometric area yet has been extensively used in the geological sciences such as sedimentology, petrophysics, and igneous petrology. The formalism of the PVA algorithm is outlined in detail by Johnson (1997). The PVA is self-training method used to determine the three parameters of interest in a mixed system: (1) the number of end-members; (2) the chemical composition of each end-member (in percent); and (3) the relative contribution of each end-member in each sample (in percent) (Johnson et al., 2002).

Johnson et al. (2000) used PVA to resolve a source degradation model developed for polychlorinated biphenyls (PCBs) in surface water of San Francisco Bay and

Estuary, California. Chemical fingerprints were identified related to known industrial formulations of PCB mixtures, as well as to natural degradation processes of the original PCB source profiles. Barabas (2004a, b) et al. developed a modified algorithm (M-PVA) to resolve a dioxin dechlorination fingerprint, indicative of biotic/abiotic transformations in field samples of sediments. Bright et al. (1999) used PVA to apportion sources of dioxins and furans to Howe Sound and the lower Strait of Georgia marine ecosystem, British Columbia, Canada, based on the deposition in recent sediments. DeCaprio et al. (2005) used PVA to determine the number, composition, and relative proportions of independent congener patterns that contributed to the overall serum PCB profile. Huntley et al. (1998) utilized PVA to determine the number of dominant fingerprint patterns of sediment cores collected throughout the Newark Bay Estuary and analyzed for polychlorinated dibenzo-p-dioxins and dibenzofurans (PCDD/Fs). The PVA results indicated that comparison of end-member patterns with source-specific fingerprint patterns found three PCDD/F congener patterns common to all models: combustion source, sewage sludge sources, and sources associated with polychlorinated biphenyls (PCBs).

PVA was performed in two steps. The first is a principal components analysis performed on the transformed matrix. Evaluation of the number of significant principal components retaining for the mixing model and evaluation of outliers are done using the goodness-of-fit criteria of Miesch (1976) and Johnson (1997). The second step in PVA determines the chemical composition of the end-member sand and the relative proportions of each end-member in each sample. This step uses the DENEG algorithm of Full et al. (1981, 1982).

1.7 Objectives

The problem of green oyster in Hsianshan wetland results in several ten million (in US dollars) loss in the fishery, and draws a lot of attention from public and government in Taiwan. Several studies had been made (BEPH, 2006; Lin et al., 1999; Huang et al., 2006) in identifying the copper source of green oyster incidents, yet, no persuasive result or clear evidence points out the source of the copper so far. Therefore, the objective of this thesis is to identify the source of copper in green oyster which has been a serious and unsolved problem in the past two decades. Note that in the summer and autumn, i.e., the rainy seasons in Hsinchu; the river stage is high and the contaminant concentration in water samples were diluted by the large amount of stream water. Therefore, we utilize PVA to analyze 12 data sets taken in April, 2006 from (Huang et al., 2006) and in February to November, 2006 (BEPH, 2006).

CHAPTER 2 MATERIAL AND METHOD

2.1 PVA Formalism

This section presents the full PVA algorithm running under default conditions; that is (1) the Q-mode factor analysis algorithm originally presented as the Fortran IV program EXTENDED CABFAC (Klovan and Miesch, 1976) and (2) the iterative oblique vector rotation algorithm originally presented as the Fortran IV program EXTENDED QMODEL (Full et al. 1981). While any number of alternative data transformation and calculation options are available and are often implemented, these algorithms represent the core of PVA as it is presently implemented under default options by the commercial version of the Quick Start software system.

2.2 Matrix adjustments

Prior to applying the PVA algorithm, several matrix adjustments to the data are required. The sampled concentrations in the Data Sets 1 and 2 are all greater than the detection limits; therefore, no adjustment is needed. However, the matrix of the Data Sets 3 to 12 contains a number of data where their concentrations are below the detection limits. In order to have a robust data set for the statistical analysis, the concentrations of those data below the detection limits are assumed equal to one half of the reported detection limits.

Three data transformations are performed. First, the data are normalized as constant row-sum sample vectors (i.e., concentrations are expressed as percent of total concentration). Then the range transformation is applied to all the variables such that their values fall in the range from 0 to 1. Finally, each of the sample vectors is transformed to having equal Euclidean length (Miesch, 1976).

2.3 Polytopic Vector Analysis

In this study, the PVA of "Quick start" package is used for the analyses. The "Quick start" is a shorter version of the manual using the most likely default values for PVA. The software was written based on the original algorithms of Klovan and Miesch (1976) and Full et al. (1981). The full PVA algorithm running under default conditions; that is (1) the Q-mode factor analysis algorithm originally presented as the Fortran IV program EXTENDED CABFAC (Klovan and Miesch, 1976) and (2) the iterative oblique vector rotation algorithm originally presented as the Fortran IV program EXTENDED QMODEL (Full et al., 1981).

The objective of PVA is to resolve three parameters of concern in a mix system: (1) the number of components (termed end-members) in the mixture, (2) the identity (i.e., composition) of each end-member, and (3) the relative proportions of each end-member in each sample. To state the problem mathematically, given a matrix X of *m* samples and *n* variables, and the true (but as yet unknown) number of end-members *k*, a feasible solution may be found and expressed as:

$$\underset{(m\times n)}{X} = \underset{(m\times k)}{A}\,\underset{(k\times n)}{F} \tag{1}$$

where *A* is a matrix of mixing proportions (the relative proportions of each of *k* end-members in each sample), and *F* is a matrix of *k* end-member compositions.

The PVA is performed in two steps. The Klovan and Miesch (1976) algorithm is used to determine the number of end-members (*k* end-members). The number of significant eigenvectors is equal to the number of end-members in the model, and is evaluated using (1) normalized varimax loadings of Klovan and Miesch (1976) and (2) the coefficient of determination (CD) of Miesch (1976) and the CD scatter plots of Johnson (1997). The second step in PVA modeling involves an iterative algorithm (Full et al., 1981) that determines the chemical composition of the end-members and the relative proportions of each end-member in each sample.

2.4 Determining the number of end-members

Of the many varieties of eigenvector methods, no topic has created more argument and controversy than the criteria used to determine the correct number of eigenvectors (i.e., k: the number of factors, or end-members). In this section, a number of methods commonly used to determine the numbers of significant eigenvectors are presented, and a graphical method is proposed.

2.4.1 Average Eigenvalue Criterion

This criterion is based on the rationale that only those principal components whose eigenvalues are above the average eigenvalue should be retained for the model. This criterion was first propose be Kaiser (1960). In those cases where principal components analysis (PCA) is performed index on the correlation matrix, the average eigenvalue will be 1.0. As such, this index is also referred to as the eigenvalue-one criterion (Malinowski, 1991).

2.4.2 Cumulative Present Variance

The rationale for this criterion is simple: a large percentage of the variance in the original matrix should be accounted for by a reduced dimensional model. However, workers in multivariable statistics, chemometrics and mathematical geology generally acknowledge that any proposed cutoff criterion is arbitrary (Deane, 1992; Reyment and Joreskog, 1993; Malinowski, 1991). The lack of a clear criterion makes the cumulative variance method dubious.

2.4.3 Scree Test

The scree test of Cattel (1966) is based on the supposition that the residual variance, not accounted for by a k principal component model should level off at the point where the principal components begin accounting for number, the point where the curve begins to level off should show a noticeable inflection point, or “knee”.

2.4.4 Normalized Varimax Loadings

The factor analysis program of Klovan and Miesch (1976) included a subroutine that calculate the number of samples with normalized varimax factor loadings greater than 0.1. The analyst looks for a sharp drop in the index as an indication of the appropriate number of eigenvectors.

2.4.5 The coefficient of determination (CD) and the CD scatter plots

Given an *m* sample by *n* variable data matrix *X* the index used by Miesch (1976) was the “coefficient of determination” (CD) between each variable in the original data matrix (*X*) and its back-calculated reduced dimensional equivalent ($\hat{X}$). For each number of potential eigenvectors, 1, 2 …n, Miesch calculated a $n \times 1$ vector as:

$$r_j^2 \cong \frac{s(x)_j^2 - (d_j)^2}{s(x)_j^2} \tag{2}$$

where $s(x)_j^2$ is the variance in the *j*th column of *X* and $(d_j)^2$ is the variance of residuals between column *j* of *X* and column *j* of $\hat{X}$. The Miesch CD is the r^2 with respect to a line of one-to-one back-calculation between X_j and $\hat{X}_j$.

A graphical extension of Miesch’s method has recently been implemented by Johnson, the CD scatter-plots (Johnson, 2002). The appropriate graphic to illustrate the fit of the Miesch CD is a series of n scatter plots that show the measured value for each variable X_j plotted against the back-calculated values from the k proposed principal component ($\hat{X}_j$).

CHAPTER 3 RESULTS AND DISCUSSION

3.1 PVA results based on Huang et al.'s data

3.1.1 The chemical pattern of Data Set 1

The chemical composition and geographic distribution of each end-member (EM) obtained from Data Set 1 for water samples are shown in Tables 5 and 6, respectively. Table 5 presents the end-member fingerprint compositions (in percent) analyzed through PVA of 6×18 data matrix with a 6 end-member model. Table 6 shows the end-member fingerprint compositions (in decimal percentages) analyzed through PVA with the same data matrix and end-member model as the previous one.

EM-1 represents the chemical fingerprint in the sampling point *I* which is located at the intertidal zone of Keya stream. It is characterized by high proportions of Pb (10.1%), Mn (13.8%), As (13%), and Mo (10.2%). EM-2 represents the fingerprint in *S* for the municipal sewer in the eastern part of Hsinchu city with high proportions of Pb(14.5%), Cr(18.2%), Ba(18.2%), and Sn(12.7%). Those heavy metals in the municipal wastewater may be from some factories such as electroplating plant or steel plant in the city. EM-3 stands for the HSP wastewater (N_0) fingerprint, characterized by high proportions of Zn(8.5%), Cu(8.5%), Fe(8.5%), Al(8.5%), Sr(8.5%), Ni(8.5%), W(7.4%), and Ga(8.5%). EM-4 represents the fingerprint at *P* (2nd period of HSP), characterized by Pb(9.5%), Cd(10.9%), Ag(10.4%), Sn(10.9%), W(10.9%), Ga(10.4%), and Mo(10.9%). EM-5 represents the fingerprint at K_{d2} (Keya Bridge), characterized by Pb(9.1%), Ag(12.4%), As(10.0%), V(12.4%), and W(10.7%). EM-6 represents the chemical fingerprint at K_{d4} (Shiangya Bridge), characterized by Pb(12.9%), Mn(10.2%), As(12.9%), W(9.3%), and Mo(9.9%).

Huang et al. (2006) mentioned that the high concentrations of heavy metals, including Zn, Cu, Pb, Fe, Al, Sr, Ba, Ni, W, and Ga, in the organs of oysters were found along the Hsianshan coast. From PVA results, EM-3 represents the HSP wastewater fingerprint characterized by high proportions of Zn, Cu, Fe, Al, Sr, Ni, W, and Ga. These results strongly support the conclusion that the heavy metals in oyster organs were mainly from the wastewater of HSP. In addition, EM-2 is characterized by high proportions of Pb, Cr, Ba, and Sn. Accordingly, the source of Pb and Ba in oyster organs may come from the municipal wastewater sewer. On the other hand, the heavy metals of the other EMs are recovered from a small number of the factories distributed in the Hsinchu city.

3.1.2 The chemical pattern of Data Set 2

Tables 7 and 8 respectively show the chemical composition and geographic distribution of each end-member (EM) obtained from Data Set 2 for suspended solid samples. Table 7 presents the end-member fingerprint compositions (in percent) analyzed through PVA of 7×18 data matrix with a 6 end-member model. Table 8 shows the end-member fingerprint compositions (in decimal percentages) analyzed through PVA with the same data matrix and end-member model as the previous one.

EM-1 represents the chemical fingerprint in the sampling point *C* which is located at the coastal area of Hsianshan. It is characterized by high proportions of Fe(14.1%), Al(14.1%), Mn(14.1%), Ba(10.2%), As(14.1%) and V(14.1%). EM-2 represents the fingerprint for the HSP wastewater discharge point (N_0) with high proportions of Ag(20.1%), Sn(10.9%), Ga(17.1%) and a relative high proportion Cu(5.8%) compared with other end-members. EM-3 stands for *S* representing the municipal sewer in the eastern part of Hsinchu city with high proportions of Zn(28.2%), Pb(11.8%), Sr(17.5%) and Mo(15.5%). EM-4 represents the fingerprint at *P* (2nd period of HSP), characterized by Zn(9.5%), Cu(9.5%), Cd(9.5%), Sr(9.5%), Ni(9.5%), Sn(9.5%), W(9.5%) and Ga(9.5%). EM-5 represents the fingerprint at *I*, which is located at the intertidal zone of Keya stream, with high proportions of Zn(8.8%), Pb(9.3%), Fe(8.1%), Mn(8.7%), Sr(9.2%), Cr(9.5%), As(8.0%) and Mo(9.5%). EM-6 represents the chemical fingerprint at K_{d2} (Keya Bridge), characterized by Zn(11.0%), As(10.2%), and Ga(13.7%). EM-7 represents the chemical fingerprint at K_{d4} (Shiangya Bridge), characterized by Zn(8.4%), Pb(9.9%), Sr(9.9%), and Ba(9.9%).

Relative high Cu proportions with 5.8% and 9.4% in EM-2 and EM-4, respectively, can be observed. EM-2 represents the HSP wastewater fingerprint and EM-4 represents the fingerprint at *P* (2nd period of HSP).

3.2 On the PVA results from Data Sets 1 and 2

Results from both the water samples and suspended solid samples indicate that high Cu proportion appearing from the HSP discharge point into the Keya stream. In addition, the PVA results of Data Set 1 show that eight heavy metals concentrations are high in both HSP wastewater point and oyster organs. Based on these results, it is reasonable to conclude that the HSP is a major Cu source to the green oyster.

3.3 PVA results for the data from BEPH project

3.3.1 The chemical pattern of Data Set 3

Data Sets 3 to 12 are water samples collected from Keya stream, Nanman creek, and Yuchegou creek. The chemical composition and geographic distribution of each end-member (EM) of Data Set 3 obtained from PVA are shown in Table 9 and Table

10, respectively. Table 9 represents the end-member fingerprint compositions (in percent) analyzed through PVA of 7×6 data matrix with the 6 end-member model. Table 10 is the end-member fingerprint compositions (in decimal percentages) analyzed through PVA from the same data matrix and end-member model. EM-3 represents the chemical pattern for samples taken from N_u. It is characterized by high proportions of SS (21.9%), BOD (29.2%) and NH_3-N (24.1%). The pattern in N_0 is characterized by EM-1 with high proportions of Cu (23.8%), SS (20.7%) and BOD (19.6%). EM-6 represents the chemical pattern for N_d, characterized by BOD (24.2%), Cu (18.8%) and As (20.0%). This conforms that the metal, Cu, comes from the discharge point of HSP (N_0) and contributes to it downstream point N_d.

EM-5 represents the fingerprint for K_u, characterized by high proportions of SS (36.9%), BOD (17.0%) and DO (17.3%). Note that the proportion of Cu in EM-2 is only 8.0%. It indicates that the major copper source may not be from the upstream reach of Keya stream (K_u). The pattern represents for K_{d1} is EM-1 and EM-4. The proportion of copper is obviously high in EM-1 and EM-6, characterized by 23.8% and 8.4%, respectively. The result indicates that K_{d1} has high Cu proportion originated from the Nanman creek. In other words, the Cu in K_{d1} also comes from the HSP discharge point. However, there are low copper proportions in K_{d3} and K_{d5} with 8.4% and 0.4% where the water may be diluted by large amount of wastewater discharged from municipal sewers and Yuchegou creek.

3.3.2 The chemical pattern of Data Set 4

Table 11 and Table 12 respectively the chemical composition and geographic distribution of each end-member (EM) of Data Set 4 obtained from PVA are shown in. Table 11 represents the end-member fingerprint compositions (in percent) analyzed through PVA of 7×6 data matrix with the 6 end-member model. Table 12 is the end-member fingerprint compositions (in decimal percentages) analyzed through PVA from the same data matrix and end-member model. EM-3 represents the chemical pattern for samples taken from N_u. It is characterized by high proportions of SS (85.7%). The pattern represents for N_0 is EM-1 and EM-6. The proportion of copper is obviously high in EM-1 and EM-6, characterized by 13.6% and 10.5%, respectively. EM-1 represents the chemical pattern for N_d, characterized by NH_3-N (36.2%), Cu (13.6%) and As (34.3%). This conforms that the metal, Cu, comes from the discharge point of HSP (N_0) and contributes to it downstream point N_d.

EM-2 represents the fingerprint for K_u, characterized by high proportions of SS (44.4%), BOD (13.0%) and As (11.2%). Note that the proportion of Cu in EM-2 is only 2.9%. It indicates that the major copper source is not from the upstream reach of Keya stream (K_u). The pattern in K_{d1} is characterized by EM-6 with high proportions of NH_3-N (29.6%), Cu (10.5%) and As (31.6%). The result indicates that K_{d1} has high Cu proportion originated from the Nanman creek. In other words, the Cu in K_{d1} also comes from the HSP discharge point. Similarly, the copper proportion in K_{d5} is high (25.2%). Yet, the Cu proportions in K_{d3} is low which may be diluted by large amount of wastewater discharged from municipal sewers and Yuchegou creek.

3.3.3 The chemical pattern of Data Set 5

Table 13 and Table 14 respectively represent the chemical composition and geographic distribution of each end-member (EM) of Data Set 5 obtained from PVA. Table 13 represents the end-member fingerprint compositions (in percent) analyzed through PVA of 7×6 data matrix with the 6 end-member model. Table 14 is the end-member fingerprint compositions (in decimal percentages) analyzed through PVA from the same data matrix and end-member model. EM-3 represents the chemical pattern for samples taken from N_u. It is characterized by high proportions of SS (52.2%) and BOD (26.5%). The pattern represents for N_0 is EM-1. The proportion of copper is obviously high, characterized by 19.9% in EM-1 revealing that the Cu may come from the HSP. EM-6 represents the chemical pattern for N_d, characterized by NH_3-N (37.6%) and Cu (21.8%). This conforms that the metal, Cu, comes from the discharge point of HSP (N_0) and contributes to it downstream point N_d.

EM-2 represents the fingerprint for K_u, characterized by high proportions of SS (84.1%). Note that the proportion of Cu in EM-2 is only 2.3%. It indicates that the major copper source is not from the upstream reach of Keya stream (K_u). The pattern in K_{d1} is comprised by EM-1 with 42.9% and EM-4 with 40.4%, the proportions of Cu in EM-1 and EM-4 are 19.9% and 24.9%, respectively. The result indicates that K_{d1} has high Cu proportion originated from the Nanman creek. In other words, the Cu in K_{d1} also comes from the HSP discharge point. Similarly, the high copper proportions in K_{d3} and K_{d5} are 19.6% and 24.9%, respectively. We are, therefore, reasonably to conclude that the copper found at K_{d3} and K_{d5} located at the downstream reaches of Keya River comes from the HSP discharge point.

3.3.4 The chemical pattern of Data Set 6

The chemical composition and geographic distribution of each end-member (EM) of Data Set 6 obtained from PVA are respectively shown in Table 15 and Table 16. Table 15 represents the end-member fingerprint compositions (in percent) analyzed through PVA of 7×6 data matrix with the 6 end-member model. Table 16 lists the end-member fingerprint compositions (in decimal percentages) analyzed through PVA from the same data matrix and end-member model. EM-2 represents the chemical pattern for samples taken from N_u characterized by high proportions of SS (46.5%) and BOD (21.9%). The pattern represents for N_0 is EM-6. The proportion of copper is relative high, characterized by 11.2% in EM-6 revealing that the Cu may also come from the HSP. EM-4 represents the chemical pattern for N_d, characterized by NH_3-N (51.6%) and Cu (16.7%). Table 15 also shows that the proportions of NH_3-N are higher than 50% in EM-1, EM-3, EM-4, and EM-6. Those analyzed results reflect that both Cu and NH_3-N come from the discharge point of HSP (N_0) and contributes to its downstream sampling points.

The pattern in K_u is comprised by EM-3 with 45.5% and EM-6 with 69.4%, the proportions of Cu in EM-1 and EM-4 are 8.5% and 11.2%, respectively. It indicates that the major copper source is not from the upstream reach of Keya stream (K_u). The pattern in K_{d1} is characterized by EM-1 with high proportion of NH_3-N (53.2%) and Cu (13.9%). The result indicates that K_{d1} has high Cu proportion originated from the Nanman creek. In other words, the Cu in K_{d1} also comes from the HSP discharge point. Similarly, the copper proportion in K_{d5} is 13.7%. However, the Cu proportion in K_{d3} is 8.5% which may reflect the problem of dilution. We are, therefore, reasonably to conclude that the copper found at K_{d3} and K_{d5} located at the downstream reaches of Keya River comes from the HSP discharge point (N_0).

3.3.5 The chemical pattern of Data Set 7

The chemical composition and geographic distribution of each end-member (EM) of Data Set 7 obtained from PVA are shown in Table 17 and Table 18, respectively. Table 17 represents the end-member fingerprint compositions (in percent) analyzed through PVA of 7×6 data matrix with the 6 end-member model. Table 18 is the end-member fingerprint compositions (in decimal percentages) analyzed through PVA from the same data matrix and end-member model. EM-4 represents the chemical pattern for samples taken from N_u. It is characterized by high proportions of SS (34.7%) and DO (24.9%). The pattern represents for N_0 is EM-5 which has a high proportion of copper, characterized by 15.9% in EM-5. EM-2 represents the chemical pattern for N_d, characterized by SS (87.4%).

The pattern in K_u is characterized by EM-6 with high proportion of NH_3-N (38.1%), As (34.0%) and Cu (13.8%). The pattern in K_{d1} is characterized by EM-1 with high proportion of NH_3-N (42.2%), As (22.5%) and Cu (16.7%). The pattern in K_{d3} is comprised by EM-2 with 75.2% and EM-6 with 90.1%, the proportions of Cu in EM-1 and EM-4 are 4.1% and 13.8%, respectively. The pattern in K_{d5} is characterized by EM-3 with high proportion of SS (70.2%) and Cu (18.2%).

The water samples in Data Set 7 were collected in 5 June, 2006. However, the rainfall record shown in Fig. 5 indicates that it were the rainy days from 28 to 31, May and 2 to 4, June. This is the reason why the proportions of SS at sampling points are very high in the sampling points of EM-2 and EM-3. It is difficult to observe the trend of Cu proportion based on the result of PVA for Data. Table 3 shows the original copper data taken in July, 2006 from various sampling points. The Cu concentrations have the same order of magnitude for N_0 and N_d at Nanman creek and for K_u, K_{d1}, and K_{d5} are at Keya stream. This fact indicates that N_0 is one of the Cu source to its downstream sampling points.

3.3.6 The chemical pattern of Data Set 8

Table 19 and Table 20 respectively show the chemical composition and geographic distribution of each end-member (EM) of Data Set 8 obtained from PVA. Table 19 represents the end-member fingerprint compositions (in percent) analyzed through PVA of 7×6 data matrix with the 6 end-member model. Table 20 is the end-member fingerprint compositions (in decimal percentages) analyzed through PVA from the same data matrix and end-member model. EM-3 represents the chemical pattern for samples taken from N_u. It is characterized by high proportions of SS (53.3%). The pattern in N_0 is major characterized by EM-6 with 83.5%, the proportions of Cu in EM-6 is 40.6%. Obviously, the wastewater from HSP wastewater point has high Cu concentration. EM-6 represents the chemical pattern for N_d, characterized by Cu (40.6%). This conforms that the metal, Cu, comes from the discharge point of HSP (N_0) and contributes to it downstream point N_d.

EM-5 represents the chemical pattern for samples taken from K_u. It is characterized by high proportions of SS (45.5%). It indicates that the major copper source is not from the upstream reach of Keya stream (K_u). The pattern in K_{d1} is characterized by EM-1 with high proportion of Cu (36.9%) and SS (19.7%). This result indicates that K_{d1} has high Cu proportion originated from the Nanman creek. In other words, the Cu in K_{d1} also comes from the HSP discharge point. Similarly, the high copper proportion in K_{d3} and K_{d5} are 21.6% and 17.3%, respectively. Therefore, it is reasonably to conclude that the copper found at K_{d3} and K_{d5} located at the downstream reaches of Keya River comes from the HSP discharge point.

3.3.7 The chemical pattern of Data Set 9

Table 21 and Table 22 respectively show the chemical composition and geographic distribution of each end-member (EM) of Data Set 9 obtained from PVA. The end-member fingerprint compositions (in percent) analyzed through PVA of 8×6 data matrix with the 6 end-member model are listed in Table 21. Table 22 is the end-member fingerprint compositions (in decimal percentages) analyzed through PVA from the same data matrix and end-member model. EM-2 represents the chemical pattern for samples taken from N_u. It is characterized by high proportions of SS (33.5%), DO (18.3%) and BOD (18.6%). The pattern in N_0 is major characterized by EM-1 (75.6%) and EM-3 (41.8%) with the proportions of Cu in EM-1 is 5.4 % and in EM-3 is 6.3%. EM-6 represents the chemical pattern for N_d, characterized by Cu (46.8%).

EM-1 represents the chemical pattern for samples taken from K_u. It is characterized by high proportions of NH_3-N (36.2%) and As (34.4%). It indicates that the major copper source is not from the upstream reach of Keya stream (K_u). The pattern in K_{d1} mainly consists of EM-1 and EM-6 with Cu proportion of 5.4% and 46.8%, respectively. Similarly, the copper proportion in K_{d3} and K_{d5} are 6.3% and 10.1%, respectively. EM-4 represents the chemical pattern for the sample collected from *Y*. It is characterized by high proportion of SS (20.2%), DO (24.5%) and BOD (22.7%). Accordingly, it is to conclude that the copper found at K_{d3} and K_{d5} located at the downstream reach of Keya River come from the sampling points of N_d and K_{d1} and, very likely, also come from HSP discharge point.

3.3.8 The chemical pattern of Data Set 10

The chemical composition and geographic distribution of each end-member (EM) of Data Set 10 obtained from PVA are respectively shown in Table 23 and Table 24. The end-member fingerprint compositions (in percent) analyzed through PVA of 8×6 data matrix with the 6 end-member model is shown in Table 23. Table 24 is the end-member fingerprint compositions (in decimal percentages) analyzed through PVA from the same data matrix and end-member model. Note that this data set was sampled in Sept. 12 and Sept. 8 to 12 were rainy days according to the rainfall record shown in Fig. 5. The proportions of Cu shown in Table 23 are very low for all the sampling points except K_{d5} and Y indicating that Y may be a minor copper source.

3.3.9 The chemical pattern of Data Set 11

The chemical composition and geographic distribution of each end-member (EM) of Data Set 11 obtained from PVA are respectively shown in Table 25 and Table 26. Table 25 represents the end-member fingerprint compositions (in percent) analyzed through PVA of 8×6 data matrix with the 6 end-member model. Table 26 is the end-member fingerprint compositions (in decimal percentages) analyzed through PVA from the same data matrix and end-member model.

N_u is mainly composed by EM-5 (55.2%) and EM-6 (65.4%). EM-5 is major characterized by NH_3-N (56.6%) and Cu (16.2%). EM-6 is major characterized by NH_3-N (26.7%) and Cu (21.9%). The pattern in N_0 is characterized by EM-4 with high proportions of As (42.9%) and SS (15.4%). EM-5 represents the chemical pattern for N_d, characterized by NH_3-N (56.6%).

EM-2 represents the chemical pattern for samples taken from K_u. It is characterized by high proportions of SS (33.5%) and Cu (33.9%). The pattern in K_{d1} is mainly composed by EM-5 and EM-6 with Cu proportion of 16.2% and 21.9%, respectively. Similarly, the copper proportion in K_{d3}, Y, and K_{d5} are 21.9%, 26.5%, and 13.5%, respectively. Therefore, the results indicate that the copper found at K_{d1}, K_{d3} and K_{d5} mainly comes from upstream of Nanman creek and upstream of Keya stream.

3.3.10 The chemical pattern of Data Set 12

Table 27 and Table 28 respectively show the chemical composition and geographic distribution of each end-member (EM) of Data Set 12 obtained from PVA. Table 27 represents the end-member fingerprint compositions (in percent) analyzed through PVA of 8×6 data matrix with the 6 end-member model. Table 28 is the end-member fingerprint compositions (in decimal percentages) analyzed through PVA from the same data matrix and end-member model.

EM-2 represents the chemical pattern for N_u, characterized by SS (67.6%). N_0 is mainly represented by EM-1 (285.6%). EM-1 is major characterized by NH_3-N (42.4%) and As (38.2%). EM-6 represents the chemical pattern for N_d, characterized by NH_3-N (39.3%) and As (42.7%).

EM-4 represents the chemical pattern for samples taken from K_u. It is characterized by high proportions of SS (45.8%) and Cu (20.3%). EM-1 represents the chemical pattern for samples taken from K_{d1}. It is characterized by high proportions of NH_3-N (42.4%) and As (38.2%). The pattern in K_{d3} is composed by EM-1 and EM-4 with Cu proportion of 5.9% and 20.3%, respectively. There is a high copper proportion in Y with 71.0%. The pattern in K_{d5} is characterized by EM-5 with high proportions of SS (41.6%) and Cu (19.2%). Therefore, it is to conclude that the copper found at K_{d5} located at the downstream may mainly comes from Youchegou creek (Y).

3.4 On the PVA results from Data Sets 3 to 12

From PVA results of Data Sets 3 to 9, it appears that high or relatively high proportion of copper can be found at N_0 and most of its downstream sampling points N_d, K_{d1}, K_{d3}, and K_{d5}. This result also supports the finding that the source of copper comes from the discharge point of wastewater from the HSP.

The water samples of Data Set 10 were taken after a long period of rainy days.

Therefore, the proportions of Cu obtained from PVA are generally very low for all the sampling points except K_{d5} and Y, indicating that Y may be a minor copper source.. The PVA results from Data Set 11 indicate that Cu source may mainly come from the upstream of the sampling points N_u and K_u. The results of Data Set 12 shows that the Cu source mainly comes from Yuchegou creek.

CHAPTER 4 CONCLUSIONS

The problem of green oyster in Hsianshan wetland has been reported for several years, yet the potential source of the green oyster still remains unsolved. The concentrations of Zn, Cu, Pb, Fe, Al, Sr, Ba, Ni, W, and Ga in the organs of oysters growing in the Hsianshan wetland were found very high. The PVA results of Data Set 1 for the water samples clearly indicate that the heavy metals in oyster organs are originated from the wastewater discharge point of HSP where has high proportions of Zn, Cu, Fe, Al, Sr, Ni, W, and Ga. The PVA results of Data Set 2 for suspended solid samples, one can observe that relative high Cu proportions at the HSP wastewater discharge point and the sampling point inside the HSP, indicating that HSP is the major Cu source.

From PVA results of Data Sets 3 to 9, it appears that high or relatively high proportion of copper can be found at N_0 and most of its downstream sampling points N_d, K_{d1}, K_{d3}, and K_{d5} indicating that the source of copper comes from the HSP wastewater discharge point. The PVA results obtained from those 7 data sets support previous finding that HSP is the major Cu source for the green oyster in Hsianshan wetland. However, the PVA results obtained from Data Sets 10 to 12 indicate that there is a minor Cu source in Yuchegou creek.

PROBLEM II

Groundwater Contamination by Organic Chemicals in Hsinchu Science Park

CHAPTER 1 INTRODUCTION

1.1 Hsinchu Science Park

Hsinchu science park (HSP) was founded in 1970. Recently, HSP has become high electronic technology product (semiconductor manufacturing companies, computer and peripherals, telecommunications, optoelectronics, precision machinery, and biotechnology, etc.) research and development center. According to the development of HSP technology industry, environmental protection becomes an important topic. In 1988, HSP initiated the groundwater monitoring and investigation works, including finding groundwater resource, planning and setting up groundwater monitor wells. In 1999, Bureau of Environmental Protection Hsinchu (BEPH) performed the investigation of the contamination for the soil and groundwater in HSP. They found that the sample collected from monitor well 19 contained the concentration of tricloroethene (TCE) exceeding 105.2 μg/L. Though the Science Park Administration had installed more than fifty monitor wells and put lots of efforts in investigating the contaminant source and contamination plume extent. However, the source of TCE contaminations and who is responsibe for the contamination still remain unsolved.

1.2 Polytopic vector analysis

Polytopic vector analysis (PVA) is a hybrid of several algorithms developed over a period of 30 years, primarily within the mathematical geology literature (Imbrie, 1963; Miesch, 1976a; Miesch, 1976b; Klovan et al., 1976; Full et al., 1981; Full et al., 1982; Ehrlich et al., 1987; Evans et al., 1992; Johnson, 1997). It is relatively new in the environmental and chemometric area yet has been extensively used in the geological sciences such as sedimentology, petrophysics, igneous petrology. The formalism of the PVA algorithm is outlined in detail by Johnson (1997). The PVA is self-training method used to determine three parameters of interest in a mixed system: (1) the number of end-members; (2) the chemical composition of each end-member (in percent); and (3) the relative contribution of each end-member in each sample (in percent) (Johnson et al., 2002).

Huntley et al. (1998) utilized PVA to determine the number of dominant fingerprint patterns of sediment cores collected throughout the Newark Bay Estuary and analyzed for polychlorinated dibenzo-p-dioxins and dibenzofurans (PCDD/Fs). The PVA results indicated that comparison of end-member patterns with source-specific fingerprint patterns found three PCDD/F congener patterns common to all models: combustion source, sewage sludge sources, and sources associated with polychlorinated biphenyls (PCBs). Bright et al. (1999) used PVA to apportion sources of dioxins and furans to Howe Sound and the lower Strait of Georgia marine ecosystem, British Columbia, Canada, based on the deposition in recent sediments. Johnson et al. (2000) used PVA to resolve a source degradation model developed for polychlorinated biphenyls (PCBs) in surface water of San Francisco Bay and Estuary, California. Chemical fingerprints were identified related to known industrial formulations of PCB mixtures, as well as to natural degradation processes of the original PCB source profiles. Barabas et al. (2004a; 2004b) developed a modified algorithm (M-PVA) to resolve a dioxin dechlorination fingerprint, indicative of biotic/abiotic transformations in field samples of sediments. DeCaprio et al. (2005) used PVA to determine the number, composition, and relative proportions of independent congener patterns that contributed to the overall serum PCB profile.

PVA was performed in two steps. The first is a principal components analysis performed on the transformed matrix. Evaluation of the number of significant principal components to retain for the mixing model and evaluation of outliers are done using the goodness-of-fit criteria of Miesch (1976a) and Johnson (1997). The second step in PVA determines the chemical composition of the end-member sand the relative proportions of each end-member in each sample. This step uses the DENEG algorithm of Full et al. (1981, 1982).

1.3 Objective

This study uses the multivariable statistic method (PVA) to analyze the groundwater sampling data for 41 samples by 59 organic compounds collected in HSP in July, 2002. This thesis aims at discussing the relationship between the end-members and the samples and between the end-members and the organic compound variables for the data sampled from HSP.

CHAPTER 2 MATERIAL AND METHOD

2.1 Matrix adjustment

The original data matrix for HSP was 41 samples by 59 organic compounds. This original matrix contains a number of elements with missing data or data reported below method detection limits (“non-detects”). In order to have a robust data set for statistical analysis, it is necessary to reduce this matrix. The variables with all missing data and non-detect data are removed from the data matrix. For PVA modeling purposes, the concentration of censored data is assumed present at a concentration equal to one half the reported detection limit. This resulted in a final matrix of 41 samples and 13 organic compounds (i.e., 1,1,1- trichloroethane, 1,1-dichloroethene, 1,1- dichloroethane, 1,2,4- trimethylbenzene, 1,3,5- trimethylbenzene, Benzene, Chlorobenzene, Chloroform, cis-1,2-dichloroethene, m,p-xylene, Toluene, Trichloroethylene (TCE), Trichlorofluoromethane).

Prior to PVA, three data transformations are performed. The data are first normalized as constant row-sum sample vectors (i.e., concentrations are expressed as percent of total concentration). Accordingly, a second transformation typically used is a range transformation (Miesch, 1976) that fixed the value of each variable between values of 0 to 1. A final transformation is performed to force each of the sample vectors to have equal Euclidean length (Miesch, 1976).

2.2 Polytopic Vector Analysis

In this study, the PVA of "Quick start" package is used for the analyses. The "Quick start" is a shorter version of the manual using the most likely default values for PVA. The software was written based on the original algorithms of Klovan and Miesch (1976) and Full et al. (1981). The full PVA algorithm running under default conditions; that is (1) the Q-mode factor analysis algorithm originally presented as the Fortran IV program EXTENDED CABFAC (Klovan and Miesch, 1976) and (2) the iterative oblique vector rotation algorithm originally presented as the Fortran IV program EXTENDED QMODEL (Full et al., 1981).

The objective of PVA is to resolve three parameters of concern in a mix system: (1) the number of components (termed end-members) in the mixture, (2) the identity (i.e., composition) of each end-member, and (3) the relative proportions of each end-member in each sample. To state the problem mathematically, given a matrix X of m samples and n variables, and the true (but as yet unknown) number of end-members k, a feasible solution may be found and expressed as:

$$\underset{(m\times n)}{\mathrm{X}} = \underset{(m\times k)}{\mathrm{A}}\,\underset{(k\times n)}{\mathrm{F}} \tag{1}$$

where A is a matrix of mixing proportions (the relative proportions of each of k

end-members in each sample), and F is a matrix of k end-member compositions.

The PVA is performed in two steps. The Klovan and Miesch (1976) algorithm is used to determine the number of end-members (k end-members). The number of significant eigenvectors is equal to the number of end-members in the model, and is evaluated using (1) normalized varimax loadings of Klovan and Miesch (1976) and (2) the coefficient of determination (CD) of Miesch (1976) and the CD scatter plots of Johnson (1997). The second step in PVA modeling involves an iterative algorithm (Full et al., 1981) that determines the chemical composition of the end-members and the relative proportions of each end-member in each sample.

2.3 Determining the number of end-members

Of the many varieties of eigenvector methods, no topic has created more argument and controversy than the criteria used to determine the correct number of eigenvectors (i.e., k: the number of factors, or end-members). In this section, a number of methods commonly used to determine the numbers of significant eigenvectors are presented, and a graphical method is proposed.

2.3.1 Average Eigenvalue Criterion

This criterion is based on the rationale that only those principal components whose eigenvalues are above the average eigenvalue should be retained for the model. This criterion was first propose be Kaiser (1960). In those cases where principal components analysis (PCA) is performed index on the correlation matrix, the average eigenvalue will be 1.0. As such, this index is also referred to as the eigenvalue-one criterion (Malinowski, 1991).

2.3.2 Cumulative Present Variance

The rationale for this criterion is simple: a large percentage of the variance in the original matrix should be accounted for by a reduced dimensional model. However, workers in multivariable statistics, chemometrics and mathematical geology generally acknowledge that any proposed cutoff criterion is arbitrary (Deane, 1992; Reyment and Joreskog, 1993; Malinowski, 1991). The lack of a clear criterion makes the cumulative variance method dubious.

2.3.3 Scree Test

The scree test of Cattel (1966) is based on the supposition that the residual variance, not accounted for by a k principal component model should level off at the point where the principal components begin accounting for number, the point where the curve begins to level off should show a noticeable inflection point, or “knee”.

2.3.4 Normalized Varimax Loadings

The factor analysis program of Klovan and Miesch (1976) included a subroutine that calculate the number of samples with normalized varimax factor loadings greater than 0.1. The analyst looks for a sharp drop in the index as an indication of the appropriate number of eigenvectors.

2.3.5 The coefficient of determination (CD) and the CD scatter plots

Given an m sample by n variable data matrix X the index used by Miesch (1976) was the “coefficient of determination” (CD) between each variable in the original data matrix (X) and its back-calculated reduced dimensional equivalent ($\hat{X}$). For each number of potential eigenvectors, 1, 2 …n, Miesch calculated a $n \times 1$ vector as:

$$r_j^2 \cong \frac{s(x)_j^2 - (d_j)^2}{s(x)_j^2} \tag{2}$$

where $s(x)_j^2$ is the variance in the jth column of X and $(d_j)^2$ is the variance of residuals between column j of X and column j of $\hat{X}$. The Miesch CD is the r^2 with respect to a line of one-to-one back-calculation between X_j and $\hat{X}_j$.

A graphical extension of Miesch’s method has recently been implemented by Johnson, the CD scatter-plots (Johnson, 2002). The appropriate graphic to illustrate the fit of the Miesch CD is a series of n scatter plots that show the measured value for each variable X_j plotted against the back-calculated values from the k proposed principal component ($\hat{X}_j$).

CHAPTER 3 RESULTS AND DISCUSSION

3.1 Determination of number of chemical fingerprints

The 13 eigenvalues and the cumulative percent variance for each are shown in Table 29. The groundwater mixing model consisting of more than four end-members would account for greater than 91% of the total observed variance in the data set as shown in Table 29. The number of end-members was determined by evaluation of normalized varimax loadings and CD scatter plots in “Quick Start”.

Normalized varimax loadings are listed in Table 29. There is a sharp drop from Factor 5 to Factor 6 and no factor beyond six has steep drop. This indicates a minimum of 5 end-members.

The “coefficient of determination” (CDs) of all variable is presented in Table 30 which indicates that a minimum of six end-members are required to accurately back-calculate all 13 analytes. The final analyte to obtain a reasonably high CD is cis-1,2-dichloroethene which progresses from a CD of 0.36 at 5 end-members to a CD of 0.66 at 6 end-members. CD scatter plots for five and six end-members are presented in Figs. 6 and 7. Figure 1 indicates a good fit for most analytes, but poor fit for a number of analytes, most notably cis-1,2-dichloroethene (CD = 0.36). For six end-members, cis-1,2-dichloroethene improves to 0.66 as presented in Fig. 7.

In summary, the criteria used to determine the number of fingerprints in the model indicate a minimum of five end-members, and as many as six. The data were submitted to "Quick Start" for resolution of a six end-member model.

3.2 End-member fingerprint composition and geographic distribution

End-member compositions and sample mixing proportions were resolved for a six end-member model. End-member fingerprint compositions shown in a bar chart of Fig. 8 to Fig. 13.

Maps showing the geographic distribution of the six end-members (EM-1 to EM-6) are presented in Fig. 8 to Fig. 13. These figures are "bubble-maps". The shaded circles or bubbles shown at each sampling station have a diameter proportional to the amount of that fingerprint in each sample. The chemical composition, geographic distribution, and interpretation of each fingerprint are discussed below.

3.2.1 End-Member 1

Through the analyses of PVA, the corresponding results show that the mixing proportions of EM-1 in each sample are less than 35% as indicated in Fig.8. In other words, the concentrations of EM-1 in each sample are small. The most likely explanation is that the EM-1 may be considered as background concentration. The background concentration is actually derived from the data transformation that the "non-detect" values are replaced by one-half of the reported limit of detection. The background concentration generally exists in samples, and their values are usually very small. Accordingly, we conclude that EM-1 is the background concentration.

3.2.2 End-Member 2

This fingerprint is related to a predominantly TCE source. The highest concentration of EM-2 is observed in the monitoring wells (MWs) 19, 29, 30, 31, 32, and 39 located in the center area shown in Fig. 9. This suggests a contribution from TCE.

3.2.3 End-Member 3

EM-3 is dominated by Trichlorofluoromethane and Chloroform. EM-3 is noted in highest proportions in the MWs 9 and 11 and spread in the center area shown in Fig. 10.

3.2.4 End-Member 4

This fingerprint is dominated by 1,1-dichloroethene. In geographic distribution, EM-4 is noted in highest proportions in the MWs 8, 9, 17, 20, 27 and 28 spreading in the center area shown in Fig. 11.

3.2.5 End-Member 5

This fingerprint is related to a predominantly Chloroform source. EM-5 is noted in highest proportions in the MW 25 located in the west region shown in Fig. 12. This suggests a contribution from Chloroform.

3.2.6 End-Member 6

This fingerprint is related to a predominantly 1,1-dichloroethene source. EM-6 is noted in highest proportions in the MW 35 located in the center area shown in Fig. 13. This suggests a contribution from 1,1-dichloroethene.

CHAPTER 4 CONCLUSION

In groundwater pollution investigation, the measured items for sampling water may not be easy to identify the chemical contaminant sources and map multiple plumes with overlapping geographic distributions. Through the use of PVA, a six end-member (eigenvector) mixing model was developed to analyze the chemical composition or geographical distribution of contributing sources. The results of this study demonstrate the application of PVA in chemical data obtained from a complex environmental system and the end-members pattern can be used to identify the contaminant source. The analyzed results indicate that the method could produce a groundwater mixing model by searching for end-members. In addition, identification of chemical contaminant sources and mapping multiple plumes with overlapping geographic distributions can be simulated by the groundwater mixing model.

REFERENCES

Barabas, N.; Adriaens, P.; Goovaerts, P. (2004a), Modified Polytopic vector analysis to identify and quantify a Dioxin dechlorination signature in sediments.1. Theory, *Environmental Science & Technology*, 38, 1813-1820.

Barabas, N.; Goovaerts, P.; Adriaens, P. (2004b), Modified Polytopic vector analysis to identify and quantify a Dioxin dechlorination signature in sediments. 2. Application to the Passaic River, *Environmental Science & Technology* 2004, 38, 1821-1827.

Bright, D. A.; Cretney, W. J.; Macdonald, R. W.; Ikonomou, M. G.; Grundy, S. L. (1999), Differentiation of Polychlorinated Dibenzo-p-dioxin and Dibenzofuran Sources in Coastal British Columbia, Canada. , *Environmental Toxicology and Chemistry*, 18, 1097-1108.

Bureau of Environmental Protection Hsinchu (2006), An Investigation of the Source of Copper and Arsenic in Discharged Wastewater and the Effect to Oyster Living in Hsiensan Area, Final report. (in Chinese)

Cattell, R. B. (1966), The scree test for the number of factors, Multivariate Behav. Res., 1, 245-276.

Deane, J. M. (1992), Data reduction using principal components analysis, In Brereton, R. G. (ed.), Multivariate pattern recognition in chemometrics, Elsevier, New York, 125-177.

DeCaprio, A. P.; Johnson, G. W.; Alice, M. T.; Carpenter, D. O.; Chiarenzelli, J. R.; Morse, G. S.; Santiago-Ricera, A. L.; Schymura, M. J. (2005), Polychlorinated biphenyl (PCB) Exposure Assessment by Multivariate Statistical Analysis of Serum Congener Profiles in an Adult Native American Population, *Environmental Research*, 98, 284-302.

Ehrlich, R.; Full, W. (1987), In: *Use and Abuse of Statistical Methods in Earth Sciences*, Size, W., Ed., Oxford, London.

Evans, J. C.; Ehrlich, R.; Krantz, D.; Full, W. (1992), A Comparison between Polytopic Vector Analysis and Empirical Orthogonal Function Analysis for Analyzing Quasi-geostrophic Potential Vorticity, *J. Geophys. Res.*, 97, 2365-2378.

Full, W. E.; Ehrlich, R.; Klovan, J. E. (1981), Extended QMODEL – Objective Definition of External End Members in the analysis of Mixtures, *J. Math. Geol.*, 13, 331-344

Full, W. E.; Ehrlich, R.; Bezdek, J. C. (1982), Fuzzy QModel – A New Approach for Linear Unmixing, *J. Math. Geol.*, 14, 259-270.

Han, B. C.; Leng, W. L.; Chen, R. Y.; Fang, G. T.; Hung, T. C. and Tseng, R. J. (1998), Estimation of Target Hazard Quotients and Potential Health Risks for Metals by Consumption of Seafood in Taiwan, *Archives of Environmental Contamination and*

Toxicology, 35, 711-720.

Huang, S, J; Hsu, S. P.; Hsieh, H. L.; Chen, C. P. and Shiu, S. J. (2006), The Heavy Metal Distribution and Relationship among Hisnchu Science Park, Keya Stream and green Oysters, *Conference on Biodiversity of Hsinchu Seacoast*, 9-18. (in Chinese)

Huntley, S. L.; Carlson-Lynch, H.; Johnson, G. W.; Paustenbach, D. J.; Finley, B. L. (1998), Identification of Historical PCD/F Sources in Newark Bay Estuary Subsurface Sediments using Polytopic Vector Analysis and Radioisotope Dating Techniques, *Chemosphere*, 36, 1167-1185.

Imbrie, J. (1963), *Factor and Vector Analysis Programs for Analyzing Geologic Data*, Technical Report No. 6, Office of Naval Research

Johnson, G. W. (1997), Ph.D. Dissertation, University of South Carolina.

Johnson, G. W. and Ehrlich, R. (2002), State of Art Report on Multivariable Chemometric Methods in Environmental Forensics, *Environmental Forensics*, 3, 59-79.

Johnson, G. W.; Jarman, W. M.; Bacon, C. E.; Davis, J. A.; Ehrlich, R.; Risebrough, R. W. (2000), Resolving polychlorinated biphenyl source fingerprints in suspended particulate matter of San Francisco bay, *Environmental Science & Technology*, 34, 552-559

Kaiser, H. F. (1958), The varimax criterion for analytic rotation in factor analysis, *Psychometrika*, 23, 187-200.

Klovan, J. E.; Miesch, A. T. (1976), EXTENDED CABFAC and QMODEL, Computer Programs for Q-mode Factor Analysis of Compositional Data, *Computers and Geosciences*, 1, 161-178.

Lin, S. and Hsieh, I. J. (1999), Occurrences of Green Oyster and Heavy Metals Contaminant Levels in the Sien-San Area, Taiwan, *Marine Pollution Bulletin*, 38, 960-965.

Malinowski, E. R. (1991),*Factor analysis in chemistry*, John Wiley and Sons, Inc., New York, 350 p.

Miesch, A. T. (1976a), Q-mode Factor Analysis of Geochemical and Petrologic Data Matrixes with Constant Row Sums, *U.S. Geological Survey Professional Paper*, 574-g:47

Miesch, A. T. (1976b), Interactive Computer Programs for Petrologic Modeling with Extended Q-mode Factor Analysis, *Computers and Geosciences*, 2, 439-492.

Reyment, R.A. and Joreskog, K. G. (1993), *Applied factor analysis in the natural sciences*. Cambridge university Press, Cambridge, 371 p.

TABLE 1. The copper concentrations of oyster and sediment in Hsianshan area form 1987 to 2006

Sampling time	**Copper concentration (ppm)**				**source**
(year/month)	oyster		sediment		
	Range	Mean	Range	Mean	
1987/10	221-993	478	-	-	Lin et al., 1999
1987/11	93.2-1080	503	-	-	Lin et al., 1999
1987/12	252-973	679	-	-	Lin et al., 1999
1993	40-200	-	0.183-2.4	-	Lin et al., 1999
1997	200-5000	-	-	200	Lin et al., 1999
1998/1-2	506-4750	2070	-	-	Lin et al., 1999
2005/1-2	-	112.6	-	-	BEPH, 2006
2005/3-4	-	585	-	-	BEPH, 2006
2005/5-6	-	209.7	-	-	BEPH, 2006
2005/7-8	-	146.7	-	-	BEPH, 2006
2006/3	726-852	790	36-73	55	BEPH, 2006
2006/4	-	2633	70-275	170	Huang et al., 2006
2006/4	505-569	527	-	-	BEPH, 2006
2006/5	250-353	339	14-24	17	BEPH, 2006
2006/7	106-174	131	-	-	BEPH, 2006
2006/8	-	-	33-70	45	BEPH, 2006
2006/10	164-202	186	16-123	51.85	BEPH, 2006

Note: The symbol – represents that the data is not available.

TABLE 2. The location of sampling points [a].

Name	Distance[a] (m) / Location	Reference source
N_u	-40	BEPH, 2006
N_0	-35	BEPH, 2006, Huang et al., 2006
N_d	-5	BEPH, 2006
K_u	-40	BEPH, 2006
K_{d1}	80	BEPH, 2006
K_{d2}	710	Huang et al., 2006
Y	7,780 (at Yuchegou creek)	BEPH, 2006
K_{d3}	7,900	BEPH, 2006
K_{d4}	8,900	Huang et al., 2006
K_{d5}	9,300	BEPH, 2006
P	2nd period science park	Huang et al., 2006
S	a municipal sewer in the eastern part of Hsinchu city	Huang et al., 2006
I	intertidal zone of Keya stream	Huang et al., 2006
C	coast area of Hsianshan	Huang et al., 2006

Note: The distance is measured from the confluence of Nanman creek and Keya stream. The minus symbol represents the upstream direction.

TABLE 3. The copper concentrations from February to December, 2006 collected from 8 sampling points along Keya stream (BEPH, 2006).

Sampling date	**Sampling points (ppm)**							
2006	N_u	N_0	N_d	K_u	K_{d1}	Y	K_{d3}	K_{d5}
February	0.002	0.023	0.016	0.003	0.018	-	0.005	ND
March	0.017	0.018	0.017	0.001	0.015	-	ND	0.014
April	0.006	0.021	0.02	0.004	0.023	-	0.015	0.022
May	0.004	0.009	0.016	0.01	0.012	-	0.009	0.008
June	0.003	0.012	0.009	0.013	0.018	-	0.005	0.021
July	0.007	0.024	0.024	0.007	0.023	-	0.018	0.021
August	0.005	0.006	0.082	0.006	0.085	0.004	0.005	0.007
September	0.003	0.004	0.005	0.010	0.008	0.005	0.007	0.013
October	0.022	0.006	0.019	0.017	0.018	0.011	0.015	0.028
November	0.004	0.015	0.007	0.011	0.007	0.075	0.005	0.014

Note: ND (non-detects) represents the value below method detection limits. The symbol – represents that the data is not available.

TABLE 4 The concentration data of 6 chemical items of Data Sets 3 to 12 (BEPH, 2006)

Sampling date 2006	**Sampling point**	**Item (ppm)**					
		SS	**DO**	**BOD**	**NH_3-N**	**Cu**	**As**
	Detected limitation	2.5	0.5	2.0	0.013	0.0010	0.00043
27, February	N_u	8.4	5.9	11.2	9.22	0.002	0.0016
	N_0	20	5.5	18.9	12.2	0.023	0.0169
	N_d	13.6	5.6	20.6	12.4	0.016	0.0171
	K_u	13.9	6.5	6.4	4.16	0.003	0.0037
	K_{d1}	14.4	5.5	14.9	12.3	0.018	0.0162
	K_{d3}	20.8	5.4	6.3	12.3	0.005	0.01
	K_{d5}	127	5.6	**1**	2.13	**0.0005**	0.0025
14, March	N_u	253	5.2	9.4	6	0.017	0.0045
	N_0	39.5	5.3	3.9	44.1	0.018	0.0476
	N_d	11.1	5.7	3	45.3	0.017	0.0429
	K_u	15.4	6.6	4.5	3.3	0.001	0.0039
	K_{d1}	27.7	5.6	7.2	42.3	0.015	0.0452
	K_{d3}	17.2	5.3	4.3	43.5	**0.0005**	0.0245
	K_{d5}	26	5.5	6.6	0.72	0.014	0.0027
11, April	N_u	42	5.2	21.3	3.4	0.006	0.0025
	N_0	22.5	5.5	6.9	41.7	0.021	0.0081
	N_d	17.7	5.6	5.2	34.6	0.02	0.0088
	K_u	145	6.5	9.3	3.47	0.004	0.0042
	K_{d1}	14.3	5.6	6.5	29.1	0.023	0.0071
	K_{d3}	12.5	5.2	8.2	29.7	0.015	0.0059
	K_{d5}	51	5.3	4.8	0.92	0.022	0.0045
10, May	N_u	28	6.8	13.2	6.45	0.004	0.0018
	N_0	7.3	6.8	4.8	40.4	0.009	0.0119
	N_d	2.7	6.6	6.2	49.5	0.016	0.0149
	K_u	5.2	6.7	5	54.6	0.01	0.0121
	K_{d1}	4.4	6.6	5.5	45.9	0.012	0.0119
	K_{d3}	7.6	6	6.1	68.5	0.009	0.0092
	K_{d5}	16.8	6.4	12.4	10.9	0.008	0.0038

TABLE 4(conti.) The concentration data of 6 chemical items of Data Sets 3 to 12 (BEPH, 2006)

Sampling date 2006	**Sampling point**	**item (ppm)**					
		SS	**DO**	**BOD**	**NH_3-N**	**Cu**	**As**
	Detected limitation	2.5	0.5	2.0	0.013	0.0010	0.00043
5, June	N_u	9.2	6.6	3.2	3.72	0.003	0.0008
	N_0	5	6.5	5.1	2.54	0.012	0.0445
	N_d	193	6.5	6.1	0.65	0.009	0.0056
	K_u	2.2	6.4	4.7	35.9	0.013	0.032
	K_{d1}	7.4	6.4	6.3	45.6	0.018	0.0243
	K_{d3}	55.5	6.2	5.5	27.2	0.005	0.0247
	K_{d5}	81	6.4	**1**	1.35	0.021	0.0047
10, July	N_u	25.6	5.9	5.1	1.55	0.007	0.0029
	N_0	10.9	6.3	8.2	2.5	0.024	0.0098
	N_d	6.8	6.1	9.7	2.47	0.024	0.0101
	K_u	25.2	6.2	10.1	1.63	0.007	0.0052
	K_{d1}	12.3	5.9	8.1	2.48	0.023	0.0105
	K_{d3}	29.3	6.1	19.3	2.47	0.018	0.0081
	K_{d5}	78	6.2	8.4	2.44	0.021	0.0056
10, August	N_u	12.8	7	7.1	4.12	0.005	0.0022
	N_0	6.4	6.4	3.8	42.5	0.006	0.0402
	N_d	6.4	6.5	5.1	41.2	0.082	0.0339
	K_u	14.2	6.8	5.5	39.9	0.006	0.0379
	K_{d1}	10.4	6.6	6.6	44	0.085	0.0497
	Y	5.6	6.8	6.3	1.23	0.004	0.0038
	K_{d3}	10.4	6.9	5.7	29.7	0.005	0.0217
	K_{d5}	26.8	6.7	18.3	6.73	0.007	0.0036

TABLE 4(conti.) The concentration data of 6 chemical items of Data Sets 3 to 12 (BEPH, 2006)

Sampling date 2006	**Sampling point**	**item (ppm)**					
		SS	DO	BOD	NH_3-N	Cu	As
	Detected limitation	2.5	0.5	2	0.013	0.001	0.00043
12, September	N_u	23.2	6.8	6.1	4.5	0.003	0.0026
	N_0	4	6.2	19.9	33.1	0.004	0.0028
	N_d	7.6	6.4	5.2	43.1	0.005	0.0268
	K_u	122	6.7	5.2	0.82	0.01	0.0027
	K_{d1}	17.2	6.5	5.2	38.3	0.008	0.0208
	Y	5.5	6.7	16.4	1.32	0.005	0.0032
	K_{d3}	21.6	6.8	4.6	36.3	0.007	0.0279
	K_{d5}	22.4	6.6	5.4	4.89	0.013	0.004
4, October	N_u	27.6	6.3	10.4	64.3	0.022	0.0021
	N_0	6.4	5.9	2.4	3.08	0.006	0.0179
	N_d	9.3	5.9	3.1	66.4	0.019	0.0136
	K_u	16.8	6.2	2.5	2.92	0.017	0.0047
	K_{d1}	7.2	5.9	2.9	57.9	0.018	0.0171
	Y	12	6.2	5.3	4.08	0.011	0.003
	K_{d3}	14	5.7	4.5	18.3	0.015	0.0111
	K_{d5}	150	5.8	9.4	8.35	0.028	0.0056
6, November	N_u	48.2	6.8	4.7	5.55	0.004	0.0021
	N_0	4.8	7.3	1	47.9	0.015	0.0322
	N_d	4	7	2.4	44.4	0.007	0.0482
	K_u	24.8	7.2	3.6	2.66	0.011	0.0049
	K_{d1}	5.5	7.2	3.2	49.9	0.007	0.045
	Y	8.7	7.2	8.5	4.09	0.075	0.0022
	K_{d3}	7.6	7.4	4.8	39.2	0.005	0.0256
	K_{d5}	30.3	6.9	10.1	8.11	0.014	0.0034

TABLE 5. End-member fingerprint compositions (in percent) analyzed through PVA of Data Set 1, 6×18 data matrix, 6 end-member model. Each column sums to 100%.

Item	**EM-1**	**EM-2**	**EM-3**	**EM-4**	**EM-5**	**EM-6**
Zn	3	0	**8.5**	3.8	3.2	3.9
Cu	0.3	0	**8.5**	0.5	0.4	0.2
Pb	**10.1**	**14.5**	0	**9.5**	**9.1**	**12.9**
Fe	2.8	0	**8.5**	6.3	3.8	2.6
Al	0	9.3	**8.5**	0.9	6.1	4.2
Mn	**13.8**	8.1	1.8	0	3.4	**10.2**
Cd	6.7	8.4	0	**10.9**	7	6.6
Sr	5.8	1.9	**8.5**	0	3.9	2.7
Cr	6.7	**18.2**	6.3	2.2	0	4.6
Ba	4.3	**18.2**	4.2	0	2.5	3.4
Ni	4.6	0	**8.5**	4.9	5.2	4.7
Ag	0	8.7	3.6	**10.4**	**12.4**	0.6
Sn	0	**12.7**	5.6	**10.9**	1.6	1.1
As	**13**	0	2.4	6.6	**10**	**12.9**
V	7.5	0	3.1	1	**12.4**	7.3
W	8.9	0	**7.4**	**10.9**	**10.7**	**9.3**
Ga	2.2	0	**8.5**	**10.4**	3	2.9
Mo	**10.2**	0	6	**10.9**	5.2	**9.9**

TABLE 6. End-member fingerprint compositions (in decimal percentages) analyzed through PVA of Data Set 1, 6×18 data matrix, 6 end-member model.

Station	**EM-1**	**EM-2**	**EM-3**	**EM-4**	**EM-5**	**EM-6**
P	0	0	0	1	0	0
N_0	0	0	1	0	0	0
S	0	1	0	0	0	0
K_{d2}	0	0	0	0	1	0
K_{d4}	0	0	0	0	0	1
I	1	0	0	0	0	0

TABLE 7. End-member fingerprint compositions (in percent) analyzed through PVA of Data Set 2, 7×18 data matrix, 7 end-member model. Each column sums to 100%.

Item	**EM-1**	**EM-2**	**EM-3**	**EM-4**	**EM-5**	**EM-6**	**EM-7**
Zn	0.0	0.0	28.2	9.5	8.8	11.0	8.4
Cu	0.0	5.8	1.2	9.5	0.3	1.2	0.4
Pb	7.8	0.0	11.8	0.9	9.3	6.8	9.9
Fe	14.1	0.5	0.0	3.2	8.1	6.0	7.9
Al	14.1	1.6	0.0	0.9	5.2	4.8	7.1
Mn	14.1	0.0	3.7	0.7	8.7	5.2	6.6
Cd	0.0	8.9	7.8	9.5	1.6	2.8	1.9
Sr	0.0	9.7	17.5	9.5	9.2	7.6	9.9
Cr	3.5	3.8	0.0	5.3	9.5	3.8	5.0
Ba	10.2	0.0	4.9	2.9	7.8	5.7	9.9
Ni	0.4	1.2	0.0	9.5	2.3	3.7	2.7
Ag	0.0	20.1	5.9	1.4	1.3	7.8	2.1
Sn	0.5	10.9	0.0	9.5	0.8	2.3	1.2
As	14.1	8.8	0.0	5.7	8.0	10.2	7.2
V	14.1	0.4	2.8	0.0	3.9	3.9	5.3
W	0.0	5.1	0.9	9.5	0.4	1.3	0.5
Ga	7.1	17.1	0.0	9.5	5.4	13.7	7.5
Mo	0.0	6.2	15.5	3.3	9.5	2.2	6.4

TABLE 8. End-member fingerprint compositions (in decimal percentages) analyzed through PVA of Data Set 2, 7×18 data matrix, 7 end-member model.

Station	**EM-1**	**EM-2**	**EM-3**	**EM-4**	**EM-5**	**EM-6**	**EM-7**
P	0	0	0	1	0	0	0
N_0	0	1	0	0	0	0	0
S	0	0	1	0	0	0	0
K_{d2}	0	0	0	0	0	1	0
K_{d4}	0	0	0	0	0	0	1
I	0	0	0	0	1	0	0
C	1	0	0	0	0	0	0

TABLE 9. End-member fingerprint compositions (in percent) analyzed through PVA of Data Set 3, 7×6 data matrix, 6 end-member model. Each column sums to 100%.

Item	**EM-1**	**EM-2**	**EM-3**	**EM-4**	**EM-5**	**EM-6**
SS	20.7	91.5	21.9	34.8	36.9	15.9
DO	5.7	4.0	15.4	9.0	17.3	6.6
BOD	19.6	0.7	29.2	10.5	17.0	24.2
NH_3-N	12.6	1.5	24.1	20.6	11.0	14.5
Cu	23.8	0.4	5.2	8.4	8.0	18.8
As	17.5	1.8	4.2	16.7	9.8	20.0

TABLE 10. End-member fingerprint compositions (in decimal percentages) analyzed through PVA of Data Set 3, 7×6 data matrix, 6 end-member model.

Station	**EM-1**	**EM-2**	**EM-3**	**EM-4**	**EM-5**	**EM-6**
N_u	0	0	1	0	0	0
N_0	1	0	0	0	0	0
N_d	0	0	0	0	0	1
K_u	0	0	0	0	1	0
K_{d1}	0.76	-0.1073	-0.0925	0.2891	0.0646	0.0861
K_{d3}	0	0	0	1	0	0
K_{d5}	0	1	0	0	0	0

TABLE 11. End-member fingerprint compositions (in percent) analyzed through PVA of Data Set 4, 7×6 data matrix, 6 end-member model. Each column sums to 100%.

Item	**EM-1**	**EM-2**	**EM-3**	**EM-4**	**EM-5**	**EM-6**
SS	8.9	44.4	85.7	46.8	18.0	19.4
DO	4.6	19.0	1.8	9.9	5.6	3.9
BOD	2.4	13.0	3.2	11.9	4.5	5.0
NH_3-N	36.2	9.5	2.0	1.3	45.6	29.6
Cu	13.6	2.9	5.8	25.2	0.5	10.5
As	34.3	11.2	1.5	4.9	25.7	31.6

TABLE 12. End-member fingerprint compositions (in decimal percentages) analyzed through PVA of Data Set 4, 7×6 data matrix, 6 end-member model.

Station	**EM-1**	**EM-2**	**EM-3**	**EM-4**	**EM-5**	**EM-6**
N_u	0	0	1	0	0	0
N_0	0.6561	0.0241	0.2185	-0.1042	-0.1591	0.3646
N_d	1	0	0	0	0	0
K_u	0	1	0	0	0	0
K_{d1}	0	0	0	0	0	1
K_{d3}	0	0	0	0	1	0
K_{d5}	0	0	0	1	0	0

TABLE 13. End-member fingerprint compositions (in percent) analyzed through PVA of Data Set 5, 7×6 data matrix, 6 end-member model. Each column sums to 100%.

Item	EM-1	EM-2	EM-3	EM-4	EM-5	EM-6
SS	21.3	84.1	52.2	57.6	16.3	19.3
DO	5.2	3.8	6.5	6	6.8	6.1
BOD	6.5	5.4	26.5	5.4	10.7	5.7
NH_3-N	39.5	2	4.2	1	38.8	37.7
Cu	**19.9**	2.3	7.5	**24.9**	**19.6**	**21.8**
As	7.7	2.4	3.1	5.1	7.7	9.6

TABLE 14. End-member fingerprint compositions (in decimal percentages) analyzed through PVA of Data Set 5, 7×6 data matrix, 6 end-member model.

Station	EM-1	EM-2	EM-3	EM-4	EM-5	EM-6
N_u	0	0	1	0	0	0
N_0	1	0	0	0	0	0
N_d	0	0	0	0	0	1
K_u	0	1	0	0	0	0
K_{d1}	0.4291	-0.2728	-0.0074	0.4043	0.3438	0.1031
K_{d3}	0	0	0	0	1	0
K_{d5}	0	0	0	1	0	0

TABLE 15. End-member fingerprint compositions (in percent) analyzed through PVA of Data Set 6, 7×6 data matrix, 6 end-member model. Each column sums to 100%.

Item	**EM-1**	**EM-2**	**EM-3**	**EM-4**	**EM-5**	**EM-6**
SS	5.1	46.5	7.1	2.8	28.8	9.1
DO	7.6	11.3	5.6	6.9	11.0	8.5
BOD	6.4	21.9	5.7	6.5	21.3	6.0
NH_3-N	53.2	10.7	64.4	51.6	18.7	50.4
Cu	13.9	6.6	8.5	16.7	13.7	11.2
As	13.8	3.0	8.6	15.5	6.5	14.8

TABLE 16. End-member fingerprint compositions (in decimal percentages) analyzed through PVA of Data Set 6, 7×6 data matrix, 6 end-member model.

Station	**EM-1**	**EM-2**	**EM-3**	**EM-4**	**EM-5**	**EM-6**
N_u	0	1	0	0	0	0
N_0	0	0	0	0	0	1
N_d	0	0	0	1	0	0
K_u	-0.1522	-0.1462	0.4552	0.0286	0.1207	0.6939
K_{d1}	1	0	0	0	0	0
K_{d3}	0	0	1	0	0	0
K_{d5}	0	0	0	0	1	0

TABLE 17. End-member fingerprint compositions (in percent) analyzed through PVA of Data Set 7, 7×6 data matrix, 6 end-member model. Each column sums to 100%.

Item	**EM-1**	**EM-2**	**EM-3**	**EM-4**	**EM-5**	**EM-6**
SS	6.9	87.4	70.2	34.7	6.6	2.3
DO	5.9	2.9	5.5	24.9	8.6	6.8
BOD	5.8	2.8	0.9	12.1	6.7	5.0
NH_3-N	42.2	0.3	1.2	14.0	3.4	38.1
Cu	16.7	4.1	18.2	11.3	15.9	13.8
As	22.5	2.5	4.1	3.0	58.8	34.0

TABLE 18. End-member fingerprint compositions (in decimal percentages) analyzed through PVA of Data Set 7, 7×6 data matrix, 6 end-member model.

Station	**EM-1**	**EM-2**	**EM-3**	**EM-4**	**EM-5**	**EM-6**
N_u	0	0	0	1	0	0
N_0	0	0	0	0	1	0
N_d	0	1	0	0	0	0
K_u	0	0	0	0	0	1
K_{d1}	1	0	0	0	0	0
K_{d3}	-0.2957	0.7521	-0.3093	0.0333	-0.0818	0.9014
K_{d5}	0	0	1	0	0	0

TABLE 19. End-member fingerprint compositions (in percent) analyzed through PVA of Data Set 8, 7×6 data matrix, 6 end-member model. Each column sums to 100%.

Item	**EM-1**	**EM-2**	**EM-3**	**EM-4**	**EM-5**	**EM-6**
SS	19.7	64.1	53.3	35.2	45.5	11.5
DO	9.5	5.1	12.3	7.3	11.2	10.3
BOD	13.0	6.9	10.6	23.2	18.3	16.4
NH_3-N	4.0	2.0	3.2	3.0	2.9	4.2
Cu	36.9	17.3	14.6	21.6	12.7	40.6
As	16.9	4.6	6.0	9.7	9.4	17.1

TABLE 20. End-member fingerprint compositions (in decimal percentages) analyzed through PVA of Data Set 8, 7×6 data matrix, 6 end-member model.

Station	**EM-1**	**EM-2**	**EM-3**	**EM-4**	**EM-5**	**EM-6**
N_u	0	0	1	0	0	0
N_0	0.1373	0.1003	0.1253	-0.1209	-0.0766	0.8346
N_d	0	0	0	0	0	1
K_u	0	0	0	0	1	0
K_{d1}	1	0	0	0	0	0
K_{d3}	0	0	0	1	0	0
K_{d5}	0	1	0	0	0	0

TABLE 21. End-member fingerprint compositions (in percent) analyzed through PVA of Data Set 9, 8×6 data matrix, 6 end-member model. Each column sums to 100%.

Item	**EM-1**	**EM-2**	**EM-3**	**EM-4**	**EM-5**	**EM-6**
SS	12.9	33.5	13.1	20.2	38.8	3.7
DO	6.2	18.3	8.7	24.5	9.7	3.7
BOD	5.0	18.6	7.2	22.7	26.5	2.9
NH_3-N	36.2	10.8	37.4	4.4	9.7	23.5
Cu	5.4	13.1	6.3	14.4	10.1	46.8
As	34.4	5.8	27.3	13.7	5.2	19.4

TABLE 22. End-member fingerprint compositions (in decimal percentages) analyzed through PVA of Data Set 9, 8×6 data matrix, 6 end-member model.

Station	**EM-1**	**EM-2**	**EM-3**	**EM-4**	**EM-5**	**EM-6**
N_u	0	1	0	0	0	0
N_0	0.7556	-0.3462	0.4184	0.1763	-0.0308	0.0267
N_d	0	0	0	0	0	1
K_u	1	0	0	0	0	0
K_{d1}	0.7385	-0.0528	-0.6948	0.0828	0.0412	0.8852
K_{d3}	0	0	1	0	0	0
K_{d5}	0	0	0	0	1	0
Y	0	0	0	1	0	0

TABLE 23. End-member fingerprint compositions (in percent) analyzed through PVA of Data Set 10, 8×6 data matrix, 6 end-member model. Each column sums to 100%.

Item	**EM-1**	**EM-2**	**EM-3**	**EM-4**	**EM-5**	**EM-6**
SS	8.1	14.4	82.8	39.8	5.7	50.2
DO	6.8	17.6	4.5	11.7	8.9	14.7
BOD	5.5	43.0	3.5	9.6	28.4	13.2
NH3-N	45.8	3.5	0.6	8.7	47.3	9.7
Cu	5.3	13.1	6.8	23.1	5.7	6.5
As	28.5	8.4	1.8	7.1	4.0	5.6

TABLE 24. End-member fingerprint compositions (in decimal percentages) analyzed through PVA of Data Set 10, 8×6 data matrix, 6 end-member model.

Station	**EM-1**	**EM-2**	**EM-3**	**EM-4**	**EM-5**	**EM-6**
N_u	0	0	0	0	0	1
N_0	0	0	0	0	1	0
N_d	1	0	0	0	0	0
K_u	0	0	1	0	0	0
K_{d1}	0.7194	-0.0904	0.0677	0.2024	0.1193	-0.0185
K_{d3}	0.9321	0.0945	0.1914	0.031	-0.1633	-0.0858
K_{d5}	0	0	0	1	0	0
Y	0	1	0	0	0	0

TABLE 25. End-member fingerprint compositions (in percent) analyzed through PVA of Data Set 11, 8×6 data matrix, 6 end-member model. Each column sums to 100%.

Item	**EM1**	**EM2**	**EM3**	**EM4**	**EM5**	**EM6**
SS	28.9	33.5	72.4	15.4	7.9	20.4
DO	14.9	12.4	2.8	14.2	5.0	8.3
BOD	12.7	5.0	4.5	5.8	2.6	6.6
NH_3-N	9.8	5.8	4.0	7.4	56.6	26.7
Cu	26.5	33.9	13.5	14.4	16.2	21.9
As	7.2	9.4	2.7	42.9	11.6	16.2

TABLE 26. End-member fingerprint compositions (in decimal percentages) analyzed through PVA of Data Set 11, 8×6 data matrix, 6 end-member model.

Station	**EM1**	**EM2**	**EM3**	**EM4**	**EM5**	**EM6**
N_u	0.4443	-0.4571	0.1497	-0.3428	0.552	0.6539
N_0	0	0	0	1	0	0
N_d	0	0	0	0	1	0
K_u	0	1	0	0	0	0
K_{d1}	-0.1168	-0.0044	-0.0421	0.061	0.822	0.2805
K_{d3}	0	0	0	0	0	1
K_{d5}	0	0	1	0	0	0
Y	1	0	0	0	0	0

TABLE 27. End-member fingerprint compositions (in percent) analyzed through PVA of Data Set 12, 8×6 data matrix, 6 end-member model. Each column sums to 100%.

Item	**EM1**	**EM2**	**EM3**	**EM4**	**EM5**	**EM6**
SS	4.7	67.6	8.2	45.8	41.6	3.5
DO	6.1	9.5	6.8	13.3	9.5	6.2
BOD	2.7	6.6	8.0	6.6	13.9	2.1
NH3-N	42.4	7.8	3.9	4.9	11.1	39.3
Cu	5.9	5.6	71.0	20.3	19.2	6.2
As	38.2	2.9	2.1	9.0	4.7	42.7

TABLE 28. End-member fingerprint compositions (in decimal percentages) analyzed through PVA of Data Set 12, 8×6 data matrix, 6 end-member model.

Station	**EM1**	**EM2**	**EM3**	**EM4**	**EM5**	**EM6**
N_u	0	1	0	0	0	0
N_0	2.8558	0.0251	0.156	0.2922	-0.4468	-1.8823
N_d	0	0	0	0	0	1
K_u	0	0	0	1	0	0
K_{d1}	1	0	0	0	0	0
K_{d3}	2.7377	-0.3347	-0.1258	0.4129	0.1687	-1.8587
K_{d5}	0	0	0	0	1	0
Y	0	0	1	0	0	0

TABLE 29. The 13 eigenvalues, the cumulative percent variance and the normalized varimax loadings for the matrix of 41 samples and 13 organic compounds used indices for determination of the number of significant eigenvectors.

Eigenvector No.	**Eigenvalue**	**Cumulative Percent variance**	**Normalized varimax loadings**
1	26.9341	65.69	36
2	6.0193	80.37	9
3	2.6238	86.77	16
4	1.8429	91.27	9
5	1.2905	94.42	**15**
6	0.781	96.32	3
7	0.6578	97.93	1
8	0.3933	98.88	3
9	0.2135	99.4	1
10	0.1736	99.83	2
11	0.0699	100	1
12	0.0003	100	0
13	0	100	0

TABLE 30. Miesch coefficients of determination (CDs) calculated from a matrix of 41 samples and 13 organic compounds. The final analyte to obtain a reasonably high CD is cis-1,2-dichloroethene which progresses from a CD of 0.36 at 5 end-members to a CD of 0.66 at 6 end-members.

Variable	**Number of end-members**										
	2	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
1,1,1-trichloroethane	0.33	0.44	0.53	0.55	0.76	0.81	0.81	0.89	1	1	1
1,1-dichloroethene	0.07	0.43	0.76	0.97	0.97	0.98	0.99	0.99	1	1	1
1,1-dichloroethane	0.5	0.53	0.75	0.87	0.94	0.96	0.97	1	1	1	1
1,2,4-trimethylbenzene	0.56	0.55	0.93	0.98	0.98	0.98	0.99	0.99	0.99	1	1
1,3,5-trimethylbenzene	0.55	0.55	0.93	0.98	0.98	0.98	1	1	1	1	1
Benzene	0.51	0.54	0.9	0.98	0.98	0.99	0.99	0.99	0.99	1	1
Chlorobenzene	0.51	0.53	0.89	0.99	0.99	0.99	1	1	1	1	1
Chloroform	0.13	0.68	0.79	0.98	0.98	0.99	1	1	1	1	1
cis-1,2-dichloroethene	0.1	0.09	0.17	**0.36**	**0.66**	0.83	0.92	0.95	1	1	1
m,p-xylene	0.36	0.44	0.56	0.6	0.83	0.87	0.9	0.95	1	1	1
Toluene	0.45	0.45	0.72	0.75	0.75	0.9	0.99	1	1	1	1
Trichloroethylene	0.88	0.95	1	1	1	1	1	1	1	1	1
Trichlorofluoromethane	-0.02	0.18	0.31	0.95	0.95	0.98	1	1	1	1	1

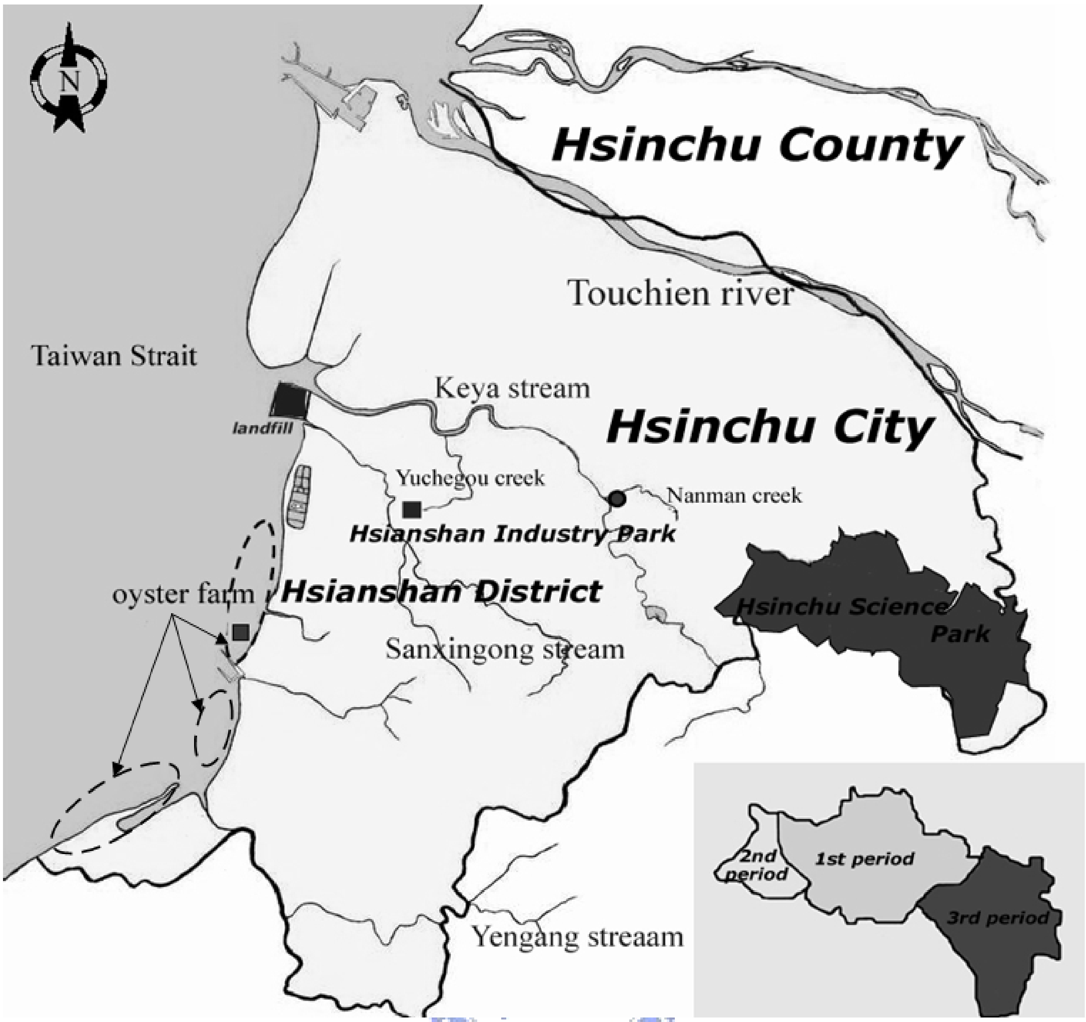

FIGURE 1. Study area locations.

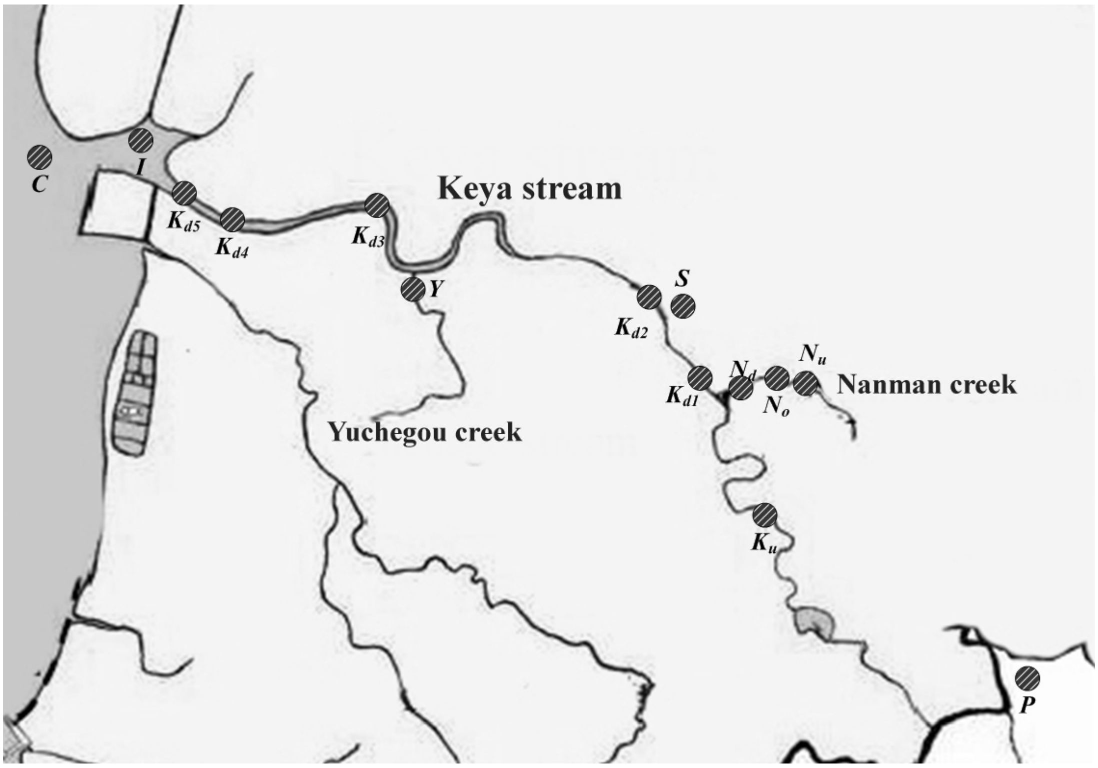

FIGURE 2. Sampling points location along Keya stream, Nanman creek, and Yuchegou creek.

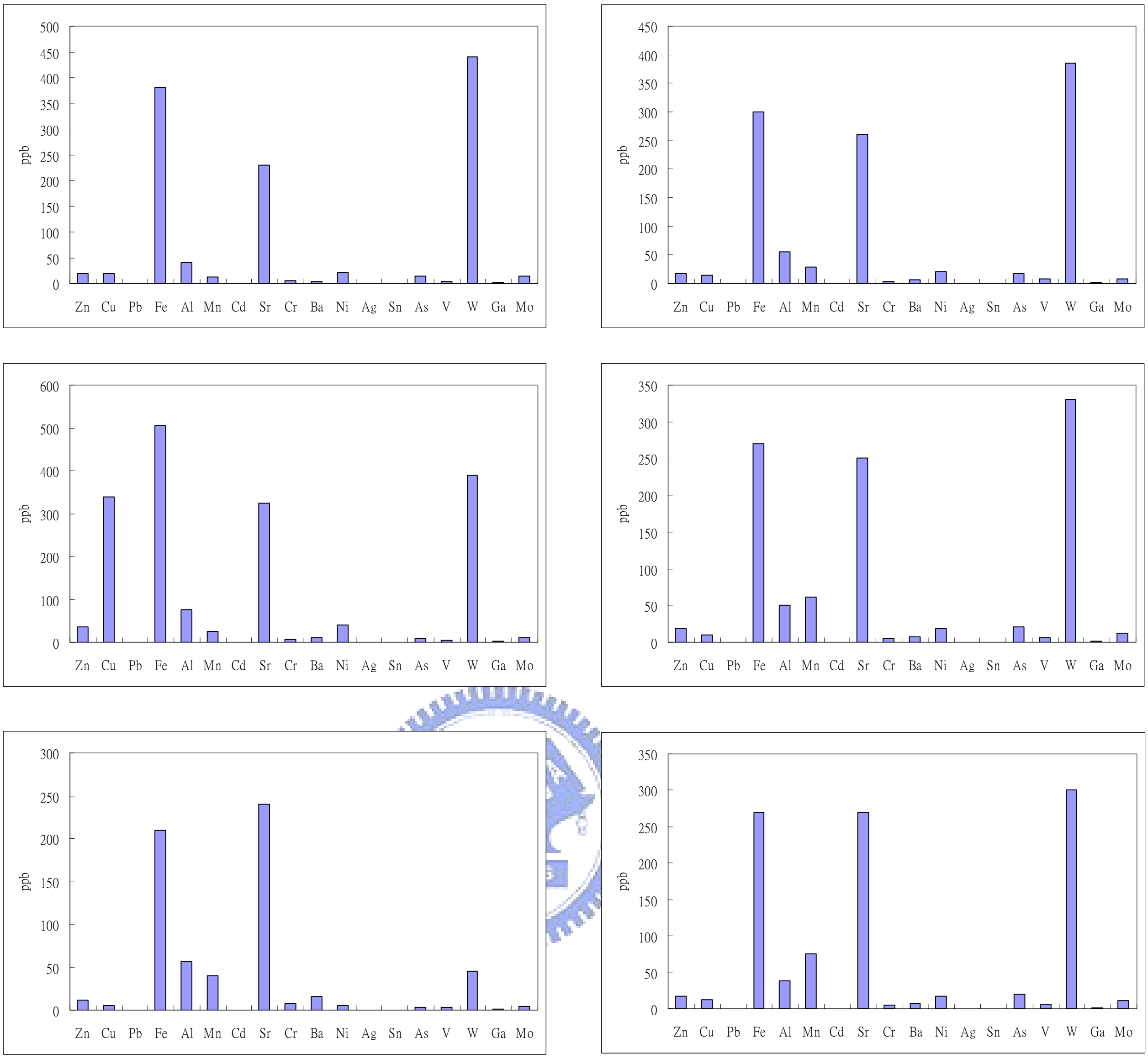

FIGURE 3. The concentration data of 18 heavy metals of water samples collected at *P*, N_0, *S*, K_{d2}, K_{d4}, and *I* in April, 2006 (Huang et al., 2006).

FIGURE 4. The concentration data of 18 heavy metals of suspended solid samples collected at P, N_0, S, K_{d2}, K_{d4}, I, and C in April, 2006 (Huang et al., 2006).

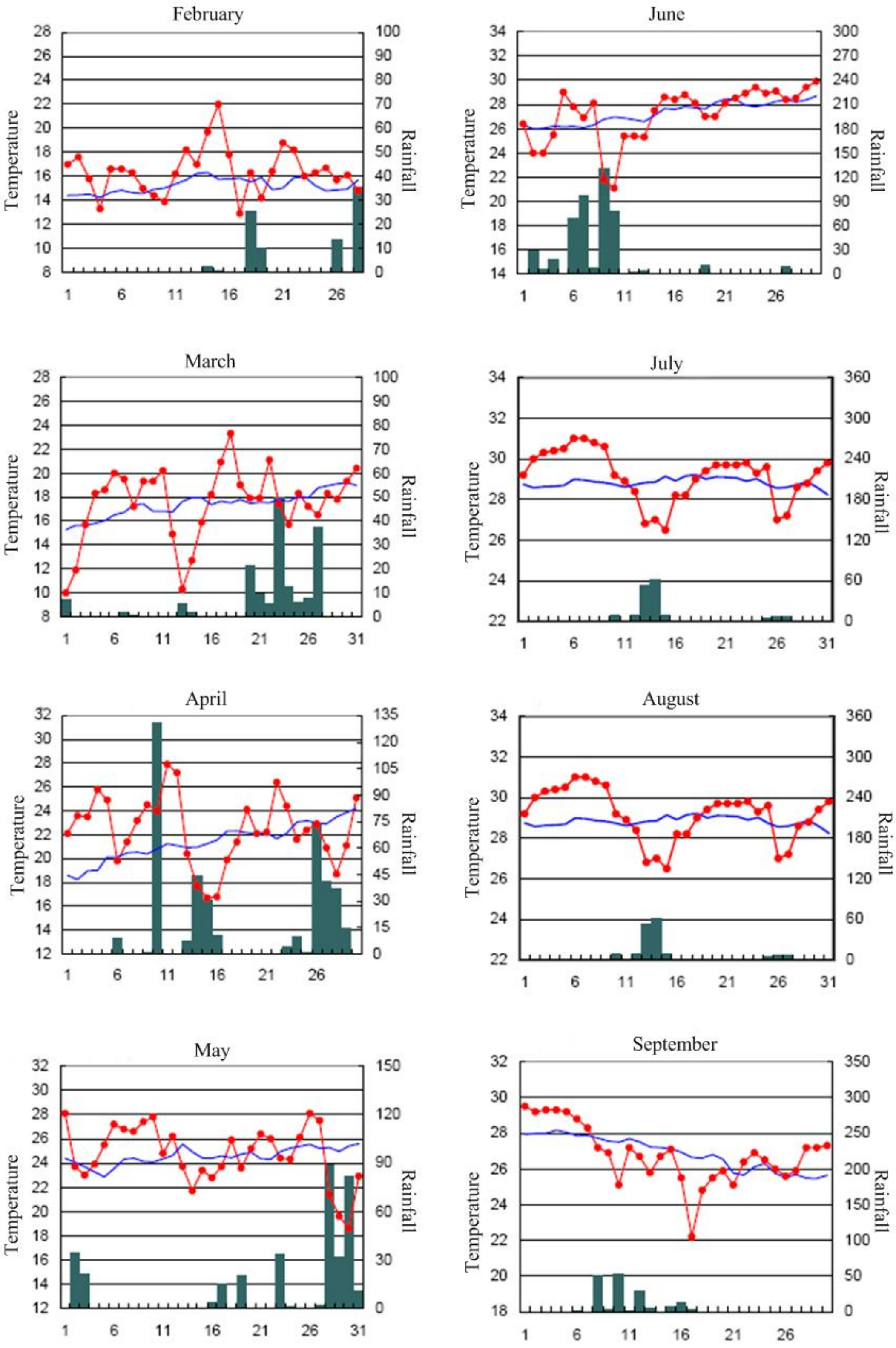

FIGURE 5. Statistical information of temperature and rainfall in Hsinchu during February to November, 2006

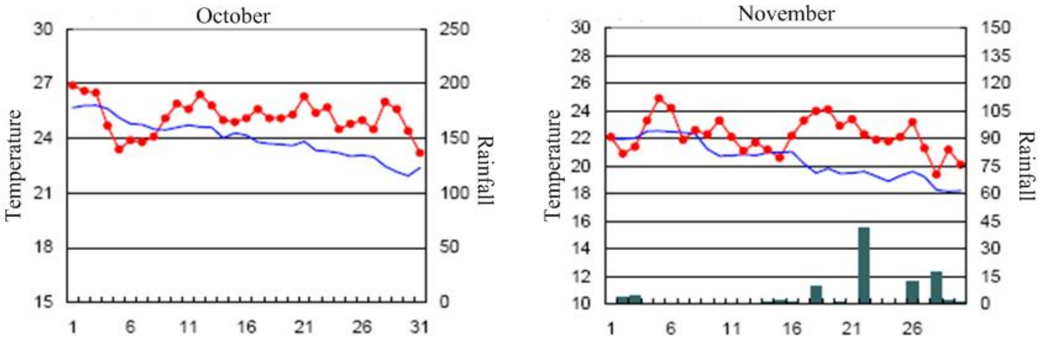

FIGURE 5. (conti.) Statistical information of temperature and rainfall in Hsinchu during February to November, 2006

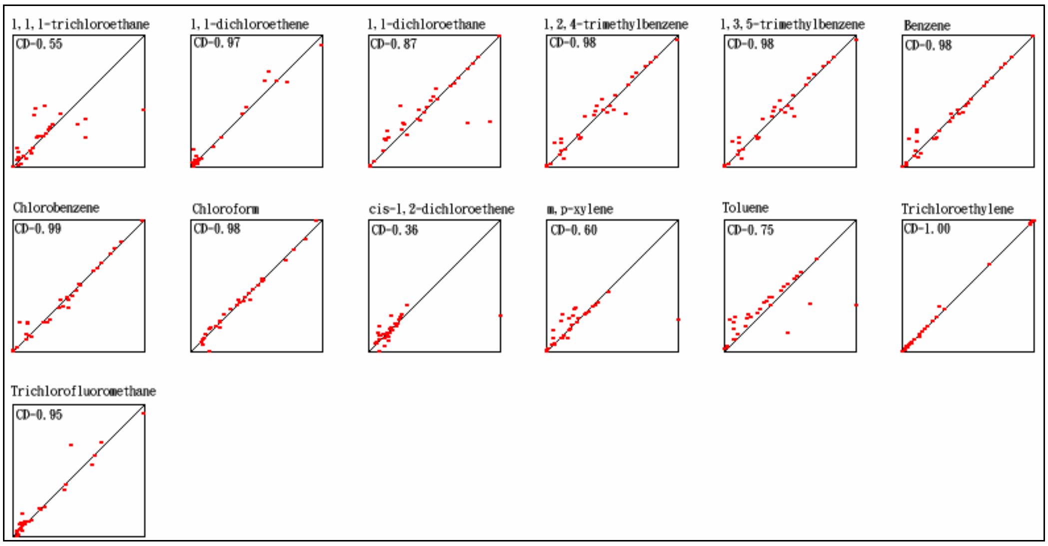

FIGURE 6. The CD scatter plots for 5 end-member model: an industry park in the northern part of Taiwan, 13x41 matrix. Figure 1 indicates a good fit for most analytes, but poor fit for a number of analytes, most notably cis-1,2-dichloroethene (CD=0.36).

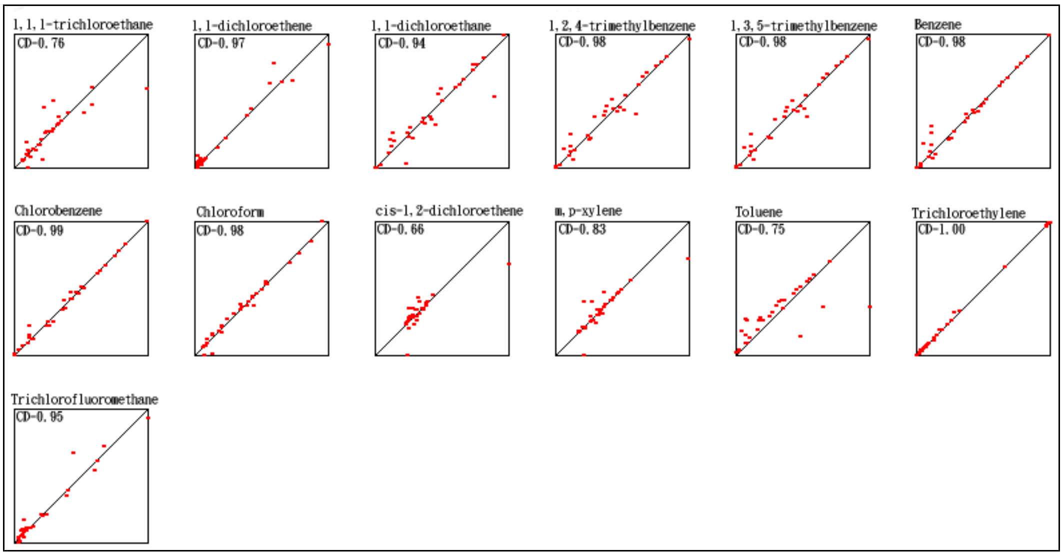

FIGURE 7. The CD scatter-plots for 6 end-member model: an industry park in the northern part of Taiwan, 13x41 matrix. For six end-members, cis-1,2-dichloroethene improves to 0.66.

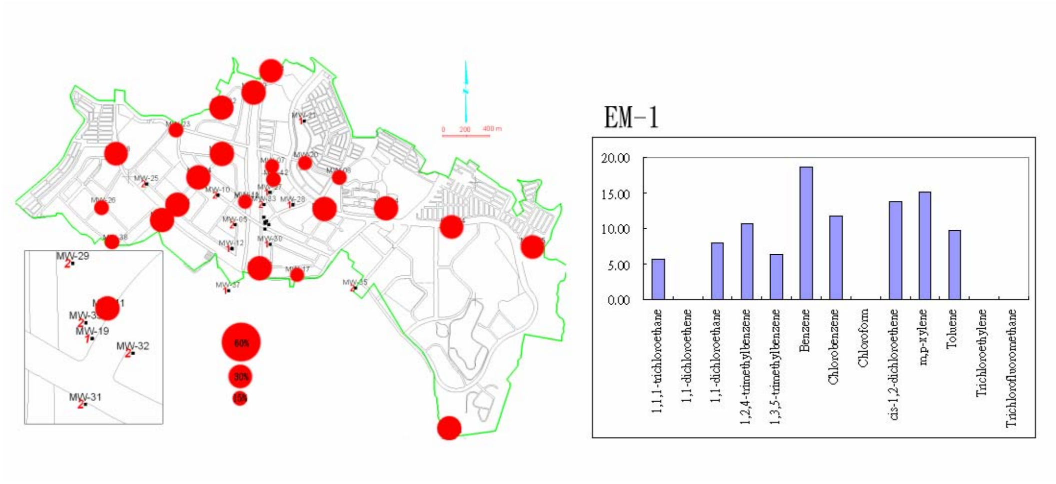

FIGURE 8. Fingerprint composition and geographic distribution for end-member 1.

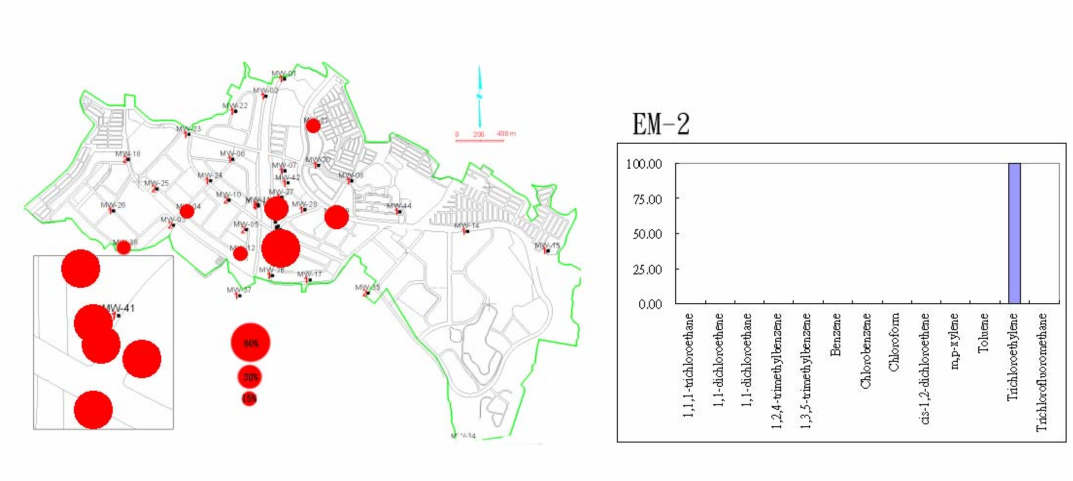

FIGURE 9. Fingerprint composition and geographic distribution for end-member 2.

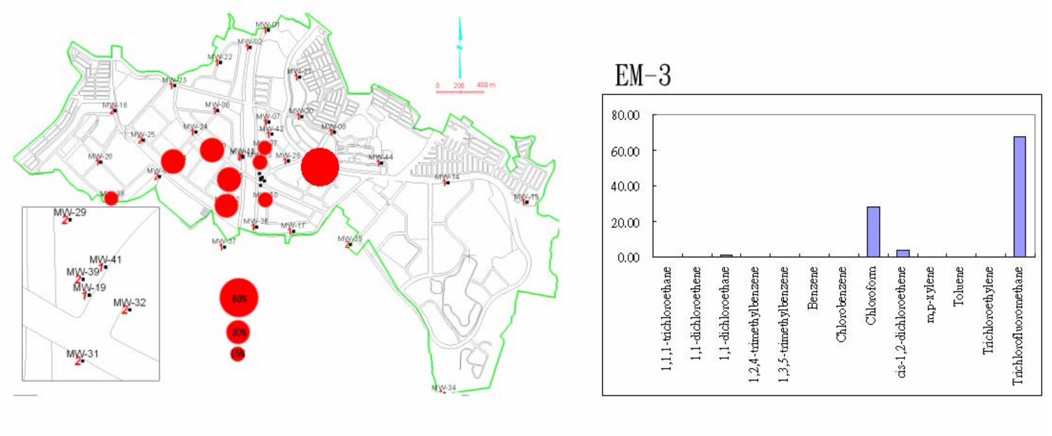

FIGURE 10. Fingerprint composition and geographic distribution for end-member 3.

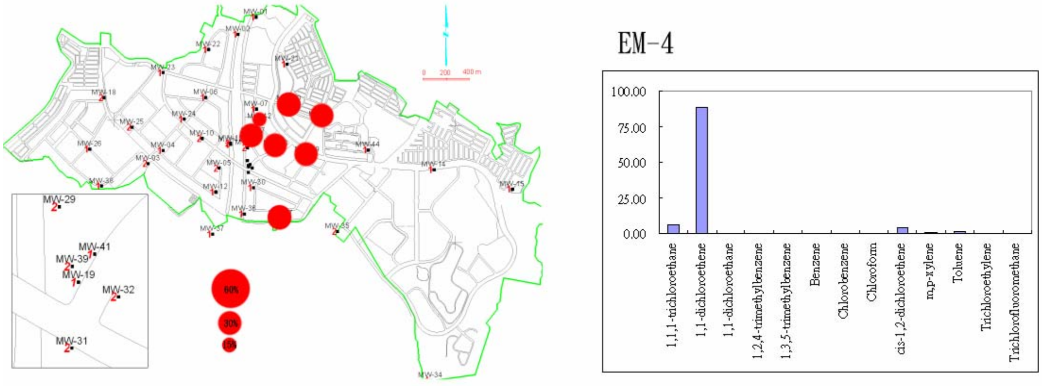

FIGURE 11. Fingerprint composition and geographic distribution for end-member 4.

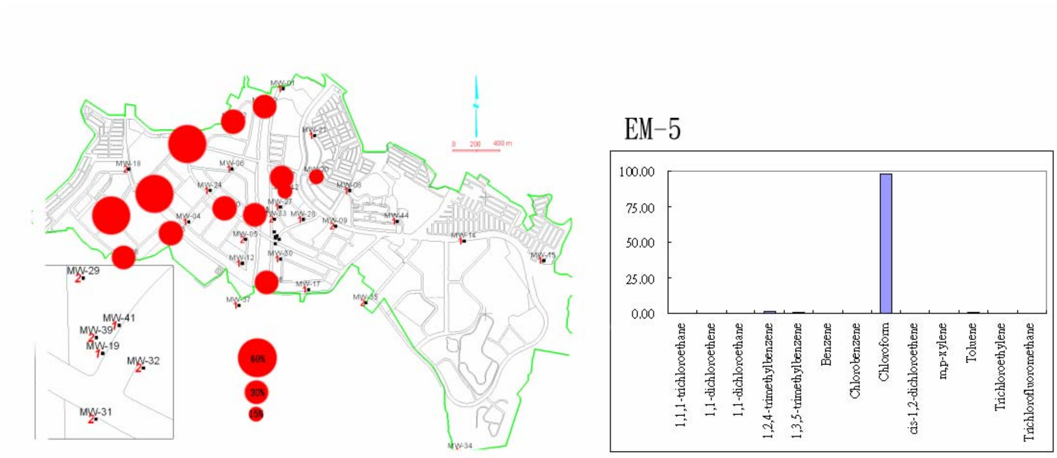

FIGURE 12. Fingerprint composition and geographic distribution for end-member 5.

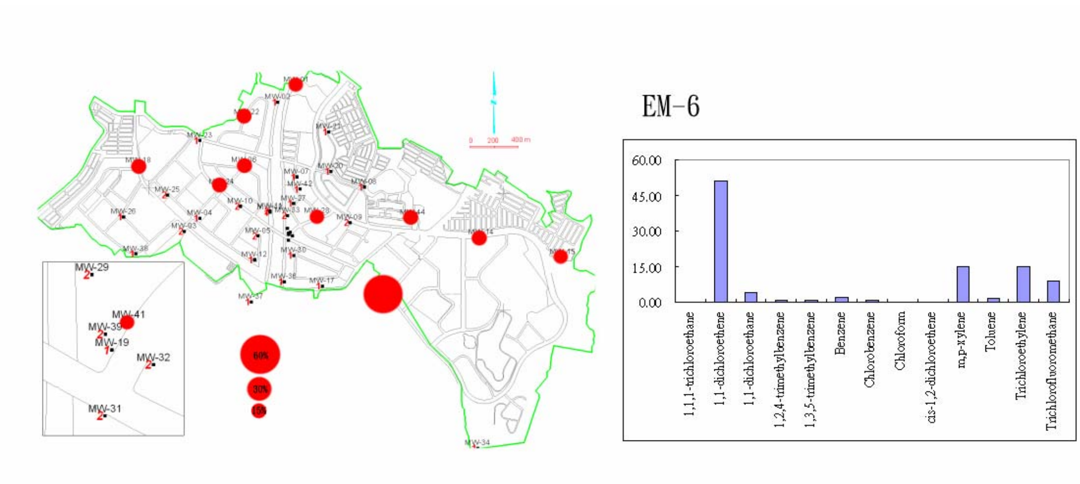

FIGURE 13. Fingerprint composition and geographic distribution for end-member 6.

個人資料

姓名：莊敏錡

生日：民國 72 年 2 月 24 日

出生地：新竹縣

聯絡電話：0952620135

email：abandon.ev94g@nctu.edu.tw

住址：新竹縣峨眉鄉峨眉村 2 鄰峨眉街 25 號

學歷：

民國 94 年畢業於國立交通大學土木工程學系

民國 96 年畢業於國立交通大學環境工程研究所