Organic Contaminants In Soil/Sediment As A Tracer For Pollution Sources

***Rozita Osman[1], Norashikin Saim[1], Md Pauzi Abdullah[2]**
[1]Faculty of Applied Sciences, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia
[2]Faculty of Science and Technology, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia

*Correspondence to: Rozita Osman, rozit471@salam.uitm.edu.my
Published online: May 24, 2010

Abstract

In this study, soil and sediment samples from known sources (sewage treatment plants, chicken farms, paddy fields and palm oil plantations) were analyzed for selected organic compounds (polycyclic aromatic hydrocarbons (PAHs), chlorpyrifos, cypermethrin and sterols). Samples were extracted using pressurized liquid extraction (PLE) techniques. The data sets obtained were subjected to chemometric techniques namely cluster analysis (CA), discriminant analysis (DA) and principal component analysis (PCA). Based on CA, two clusters were generated; cluster 1 consisting of PAHs, chlorpyrifos and cypermethrin while cluster 2 consisting of coprostanol, cholesterol, stigmasterol, stigmastanol, and β-sitosterol. DA plot rendered clear discrimination between the samples from different sources of contaminations. PCA analysis, applied to the data sets resulted in four latent factors explaining 63.8% of the total variance. The varifactors obtained from PCA indicated that the parameters responsible for source variations are sterols (coprostanol, cholesterol, stigmasterol, β-sitosterol and stigmastanol), which are strongly correlated to sewage and chicken farm samples, PAHs (naphthalene, acenaphtene, pyrene, benzo[a]anthracene and benzo[a]pyrene) and pesticides (chlorpyrifos and cypermethrin) which are comparatively correlated to samples from agricultural activities (paddy fields and palm oil plantations). The application of chemometric on selected organic compounds provides useful and promising techniques in tracing sources of contaminants as well as to reduce the complexity of the data interaction without lost of much important information.

Keywords: Chemometric; organic contaminants; cluster analysis; discriminant analysis; principal component analysis.

1. Introduction

Identifying the sources of organic contaminants in the environment is of key importance to our understanding of pollution patterns and for making decisions concerning site remediation. In this study, soil/sediment samples from point sources of pollution (sewage treatment plants, chicken farms, paddy fields and palm oil plantations) were analyzed for selected organic contaminants (polycyclic aromatic hydrocarbons (PAHs), chlorpyrifos, cypermethrin and sterols). These samples were used to determine the actual organic contaminants representative of that particular source. Based on previous studies, the use of faecal sterols can indicate the source of faecal matter because they vary in type and quantity in the faeces of humans and animals and the microbiota in their digestive tract [1,2].

The data sets obtained were subjected to various chemometric techniques including Cluster Analysis (CA), Discriminant Analysis (DA) and Principal Component Analysis (PCA). Environmental data normally contain huge amounts of concentration values of chemicals, spread at distant geographical sites and during different time periods. Moreover, the content of chemicals is also estimated at different environmental compartments (i.e., air, water, sediments and biota). All these data values are difficult to cope and evaluate in a simple and fast way using simple univariate statistical tools due to their large number and due to their multivariate correlation. In order to discover relevant patterns within large multivariate data sets, the application of chemometric methods based on statistical multivariate data analysis is proposed. Chemometric methods have often been used in exploratory data analysis tools for classification [3,4] of samples (observations) or sampling stations and identification of pollution sources [5,6]. In many other cases, the exploratory data analysis results will serve to achieve an insight into, e.g., the contamination situation of a certain location and to make a plan for remediation or to prepare more focused sampling plans. Principal components analysis (PCA) is an exploratory, multivariate, statistical technique that can be used to examine data variability. The principal components are ordered in such a way that the first PC explains most of the variance in the data, and each subsequent one accounts for the largest proportion of variability that has not been accounted for by its predecessors. Although the number of PCs equals the number of independent original variables, generally, most of the variation in the data set can be explained by the first few principal components that can be used to represent the original observations [7]. Multivariate techniques can consider a number of factors, which control data variability simultaneously [8] and therefore offer significant advantages over univariate techniques.

The pressurized liquid extraction technique (PLE) was used to achieve faster extraction of organic contaminants from soil/sediment samples. This study was carried out to fulfill the main objectives of using organic contaminants as a tracer for

pollution sources by chemometric methods within the known pollution area. The results can be a useful tool in developing appropriate strategies for the effective management of environmental problems.

2. Methods

2.1 Standards and reagents

Methanol and *n*-hexane were of pesticide residue grade and purchased from Merck (Darmstadt, Germany). Silica gel (70-230 mesh ASTM) was obtained from Merck (Darmstadt, Germany) and non-washed diatomaceous earth was purchased from Sigma-Aldrich (Steinheim, Germany). Silica gel was activated for 24 hours at 130 °C before used. This was cooled in dessicator prior to use. Chlorpyrifos PESTANAL® 99.5%, cypermethrin mix of isomers PESTANAL® 98% and *N, O*-bis(trimethylsilyl)trifluoroacetamide (BSTFA) + trimethylchlorosilane (TMCS), 99:1 were purchased from Sigma-Aldrich, GmbH (Seelze, Germany). Individual standards of PAHs: naphthalene, acenaphtene, antracene and pyrene were obtained from Dr. Ehrenstorfer, GmbH (Augsburg, Germany). Acenaphthylene, fluorine, benzo[a]anthracene and benzo[a]pyrene, were obtained from Supelco (Bellefonte, USA). Individual standards of sterols, specifically 5β-cholestan-3β-ol (coprostanol), 5-cholesten-3β-ol (cholesterol), 5β-cholestan-3α-ol (stigmasterol) and stigmastanol were purchased from Sigma Aldrich (Steinheim, Germany). Internal standards, phenantrene d_{10} (Supelco, Bellefonte, USA) and 5α-cholestane (Sigma Aldrich, Steinheim, Germany) were used for quantification. Certified reference material® 1944 (CRM1944) was obtained from National Institute of Standards and Technology (NIST), Gaithersburg, USA.

2.2 Sampling and analytical procedures

Soil and sediment samples were collected within known pollutants areas: chicken farms, sewage treatment plants, paddy fields and oil palm plantations. For each sampling location, several sampling points were selected in order to minimize bias. Contaminated soils were air dried at room temperature to a water content of less than 5% and sieved through a 600 μm pore sieve. Soil samples were stored in an air-tight container at 4 °C until analysis. Extractions were done using pressurized liquid extraction (PLE) techniques followed the method reported by Osman et al. [9]. Cell loading was done in the following sequence: a cellulose filter was placed at the bottom of the cell, followed by 5 g of activated silica, another cellulose filter and finally a soil sample (5 g) mixed with diatomaceous earth. The sample cells were then closed to finger tightness and placed into the carousel of the PLE system. Two solvents, *n*-hexane and methanol (MeOH), were utilised as extraction solvents. In the first cycle, *n*-hexane was pumped into the cell and was preheated for 2 min to reach the optimum setting temperature and pressure (125 °C, 1400 psi) followed by a static extraction of 10 min. At the end of the cycle, the pressure was released and the extract was collected in 60 mL glass vials. The cell was rinsed with fresh solvent (about 80% of the extraction cell volume) and purged using pure nitrogen for 60 s. For the second extraction cycle, the sample was extracted again using MeOH under the same conditions. The extract was collected into a second collection vial. Internal standards (phenantrene, d_{10} and 5α-cholestane, 1 mL each) were added to the extracts and the volume was reduced to 1 mL prior to gas chromatograph analysis. The PLE extraction method was validated using 5 g soil samples spiked with sterols (coprostanol, cholesterol, stigmasterol, β-sitosterol and stigmastanol), PAHs (naphthalene, acenaphthylene, acenaphtene, fluorine, pyrene, benzo[a]anthracene and benzo[a]pyrene), chlorpyrifos and cypermethrin at the levels 5, 10, 20, and 60 μg mL^{-1}.

2.3 Gas Chromatograph Systems

In this study, gas chromatograph with flame ionisation detector (GC-FID) and gas chromatography with electron capture detector (GC-ECD) were used in order to separate all compounds.

2.3.1 GC-FID Analysis

Gas chromatographic separation and identification of PAHs and sterols was performed using an HP6890 series II (Agilent Technologies Inc., Palo Alto, CA, USA) with splitless injection and flame ionisation detection. A 30 m x 0.25 mm id x 0.25 μm film thickness HP5-MS capillary column (Agilent technologies) was used to achieve separation of PAHs and sterols with the following temperature program: initial temperature, 50 °C; held for 2 min; increase by 18 °C min^{-1} to 250 °C, increase by 10 °C min^{-1} to 310 °C; held for 11 min. The detector temperature was set at 310 °C. PAH quantification was carried out using a five-point calibration plot containing 5, 10, 25, 50 and 100 mg L^{-1} PAH standard mixtures and 20 mg L^{-1} internal standard (phenantrene, d_{10}). Sterol quantification was carried out using five-point calibration plot containing 5, 10, 25, 50 and 100 mg L^{-1} sterol standard mixtures and 20 mg L^{-1} internal standards (5α-cholestane).

2.3.2 GC-ECD Analysis

Separation of chlorpyrifos and cypermethrin was achieved using HP7890A gas chromatographs equipped with ^{63}Ni electron capture detectors, GC-ECD (Agilent Technologies Inc., Palo Alto, CA, USA). A 30 m x 0.25 mm id x 0.25 μm film thickness HP5-MS capillary column (Agilent technologies) was used for the quantitative analysis of chlorpyrifos. The injection port and detector temperatures were set at 250 °C. The injection volume was 1 μL, and the splitless period following the injection was 2 min. The ECD detector utilised pure N_2 (>99.999%) as a carrier and make-up gas at a controlled constant velocity of 60 mL min^{-1}. The temperature program of the HP5-MS column was set to 150 °C for 1 min. then increased by 25 °C min^{-1} to 260 °C for 8 minutes. Compounds were identified based on the retention time of the standards and quantified by external standard calibration.

2.4 Chemometric approach

Environmental data sets are usually complex and contain a large amount of information with internal relationships among variables, often in a partially hidden structure. The goal of chemometric studies is to display the most significant patterns, looking for possible groupings and sources of data variation, as well as for their temporal and geographical distributions through resolution and modeling of the raw data [10]. After data conversion into a single matrix formed by concentration values for each combination of variables and cases, a stepwise statistical approach was used employing the following exploratory techniques: cluster analysis (CA), discriminant analysis (DA) and principal component analysis (PCA). Prior to analysis, non-detected values were replaced with half the detection limit [11].

3. Results and Discussion

3.1 Method Validation

The linearity of the extraction method was studied using soil samples spiked with standard PAHs (naphthalene, acenaphthylene, acenaphthene, fluorene, pyrene, benzo[a]anthracene and benzo[a]pyrene), sterols (coprostanol, cholesterol, stigmasterol, β-sitosterol and stigmastanol), chlorpyrifos and cypermethrin at levels 5, 10, 20, 40 and 60 μg mL^{-1}. Good linearity was obtained for all compounds with a correlation coefficient (r^2) in the range of 0.9910 to 0.9998 (Table 1). The method was found to be precise (RSD <9%) and accurate, with satisfactory recoveries for all compounds. Recoveries of PAHs found between 85 to 98%, sterols (97 to 104%), chlorpyrifos (84 to 103.4%) and cypermethrin (88 to 101%). The instrumental limit of detection (LOD) and limit of quantification (LOQ) were calculated based on 3:1 and 10:1 signal to noise ratios, respectively, using the standard solution containing the compounds at the lowest concentration levels. In order to estimate the accuracy and precision of the method developed, a reference soil (CRM1944) was extracted in triplicate using 5 g silica packed in PLE extraction cell at the optimum conditions. The extraction conditions and the results are presented in Table 2. All values do not have statistical significance at 95% confidence level ($p > 0.05$).

3.2 Point Source Pollution

Most of the research and studies conducted on the Malaysian context were on point source pollution as this pollution gave acute impacts on the ecosystem [12]. Controlling point source pollution is also much easier as "polluter-pay-principal" is applicable. As for non point source pollution, this area is still new in terms of number of study conducted due to much difficulty in acquiring information regarding the pollution as the impact is long term. This study was carried out to fulfill the main objectives of identifying and differentiating sources of organic contaminants using chemometric methods within point source of pollution. Soil/sediment samples from point sources (paddy field, oil palm plantations, sewage treatment plants and chicken farms) were used to determine the organic contaminants representative of that particular source. PAHs (naphthalene, acenaphthylene, acenaphtene, fluorene, pyrene, benzo[a]anthracene and benzo[a]pyrene), sterols (coprostanol, cholesterol, stigmasterol, β-sitosterol and stigmastanol), insecticides (chlorpyrifos and cypermethrin) were the selected organic contaminants. The profiles of these contaminants from point sources were subjected to chemometric techniques to provide information in identifying non-point source pollution. Table 3 shows the concentrations of the organic contaminants from different point sources.

It clearly showed that coprostanol is dominant in the sewage sample, followed by cholesterol, stigmasterol and stigmastanol. Similar findings were reported from previous studies whereby coprostanol has been widely used as a marker of human sewage contamination in faecal pollution [2,13]. Higher concentration of β-sitosterol was found in chicken farm soil samples compared to other samples. The concentration of PAHs such as naphthalene, acenaphthylene, pyrene and benzo[a]pyrene was relatively

high in agricultural land due to the burning of plant waste. The high concentration of chlorpyrifos in paddy field samples compared to that of oil palm may be due to the more frequent usage of this compound to control pest in paddy plantation.

Table 1. Analytical performance of one-step extraction using silica (5 g) packed in PLE extraction cell.

Compound	Linear response (r^2)	Instrumental LOD ($\mu g\ mL^{-1}$)	Instrumental LOQ ($\mu g\ mL^{-1}$)	Precision RSD (%) (n=7)
Naphtalene	0.9963	0.050	0.165	2.0
Acenaphthylene	0.9990	0.050	0.165	8.8
Acenaphthene	0.9970	0.050	0.165	2.7
Fluorene	0.9965	0.100	0.333	1.0
Pyrene	0.9910	0.100	0.333	0.7
Benzo[a]anthracene	0.9944	0.100	0.333	2.5
Benzo[a]pyrene	0.9993	0.100	0.333	2.4
Chlorpyrifos	0.9975	0.054	0.018	5.1
Cypermethrin	0.9985	0.250	0.833	7.6
Coprostanol	0.9997	0.100	0.333	1.8
Cholesterol	0.9992	0.100	0.333	1.9
Stigmasterol	0.9939	0.100	0.333	0.9
β-sitosterol	0.9972	0.200	0.667	3.8
Stigmastanol	0.9998	0.500	1.650	1.8

Table 2. PAH concentration found in the reference soil CRM1944 (NIST, New York/New Jersey Waterway Sediment), using selective PLE (0.5 g soil, 5 g silica) packed inside extraction cell, solvent: *n*-hexane, 125 °C, 1400 psi, 10 min static extraction. All values are expressed in mg kg^{-1} dry mass.

Compound	Concentration found (mg/kg dry mass), n=3	Certified concentration values (mg/kg dry mass)
Naphtalene	1.51 ± 0.17	1.65 ± 0.31
Pyrene	7.68 ± 0.41	9.70 ± 0.42
Benzo[a]anthracene	4.10 ± 0.29	4.72 ± 0.11
Benzo[a]pyrene	4.20 ± 0.01	4.30 ± 0.13

Table 3. Concentration of PAHs, sterols, chlorpyrifos and cypermethrin detected from point source of pollution.

Organic Contaminants	Concentration (μg/kg), n=140			
	Paddy Field	Palm Oil Plantation	Sewage Treatment Plant	Chicken Farm
Naphthalene	18.93 ± 6.84	3.85 ± 0.23	9.11 ± 4.75	4.89 ± 0.34
Acenaphthylene	17.86 ± 1.10	3.90 ± 0.31	6.08 ± 5.00	4.30 ± 0.31
Acenaphthene	83.50 ± 8.50	3.85 ± 0.52	5.78 ± 3.21	5.58 ± 0.68
Fluorene	10.70 ± 6.22	8.78 ± 0.52	10.04 ± 6.80	4.97 ± 0.33
Pyrene	68.10 ± 12.71	16.54 ± 3.78	11.56 ± 2.10	8.95 ± 0.79
Benzo[a]anthracene	27.20 ± 12.30	35.77 ± 5.72	10.63 ± 2.10	9.93 ± 0.34
Benzo[a]pyrene	115.21 ± 12.3	27.03 ± 7.90	18.21 ± 2.17	9.79 ± 0.89
Coprostanol	273.11 ± 14.20	44.02 ± 7.50	40684.86 ± 32.53	2153.18 ± 4.60
Cholesterol	492.82 ± 9.81	117.71 ± 8.45	22985 ± 38.20	7463.09 ± 10.82

Stigmasterol	1300 ± 4.82	397.10 ± 12.87	7457.42 ± 21.68	3665.82 ± 20.06
β-sitosterol	477.53 ± 13.72	44.02 ± 10.39	4224.29 ± 33.10	15696.60 ± 11.89
Stigmastanol	292.51 ± 11.60	101.95 ± 13.01	13002.25 ± 8.79	2898.24 ± 7.31
Chlorpyrifos	635.8 ± 7.70	114.86 ± 7.15	19.67 ± 6.10	10.57 ± 0.74
Cypermethrin	8.82 ± 1.91	35.91 ± 8.10	0.40 ± 0.34	0.22 ± 0.02

3.3 Cluster analysis (CA)

CA was applied to the data sets in order to identify similarities in organic contaminants contents between the samples. The CA analysis grouped the variables into two groups (Figure 1). Cluster 1 consists of acenaphthylene, naphthalene, fluorene, cypermethrin, pyrene, acenaphtene, chlorpyrifos, benzo[a]pyrene (BaP) and benzo[a]anthracene (BaA). The presence of chlorpyrifos, cypermethrin and PAHs in the same group (Cluster 1) most probably because these pesticides were used in agricultural land while PAHs resulted from agricultural activities such as burning of plant wastes. Sterols (coprostanol, cholesterol, stigmasterol, stigmastanol and β-sitosterol) were group together in Cluster 2. However, β-sitosterol forms its own subgroup showing that this compound may be a significant indicator for a specific source of pollution. Based on previous study on faecal contamination by Leeming et al. [2] and Saim et al. [1], β-sitosterol is the major sterol found in samples from bird species. CA on cases which involved the sampling sites (Figure 2) showed three clusters, representing oil palm and paddy (cluster 1), chicken (cluster 2) and sewage treatment plant (cluster 1). The sterols dominated in sewage and chicken farm were coprostanol, cholesterol, stigmasterol, stigmastanol and β-sitosterol. Variables in cluster 1 (chlorpyrifos, cypermethrin, benzo[a]anthracene, benzo[a]pyrene,pyrene, acenaphtene, acenaphthylene, fluorene, pyrene and naphthalene) significantly representing paddy and oil palm soil samples.

3.4 Discriminant Analysis (DA)

A total of 140 soil and sediment samples collected from point sources were subjected to discriminant analysis. As shown in Figure 3, DA plot rendered clear discrimination between the samples from different sources of contaminations. The sewage samples were clearly isolated from other samples suggesting that it can successfully be used as the source tracer for faecal contamination in an environmental compartment. The findings concurred with those reported by Isobe et al. [14] and Leeming et al. [2].

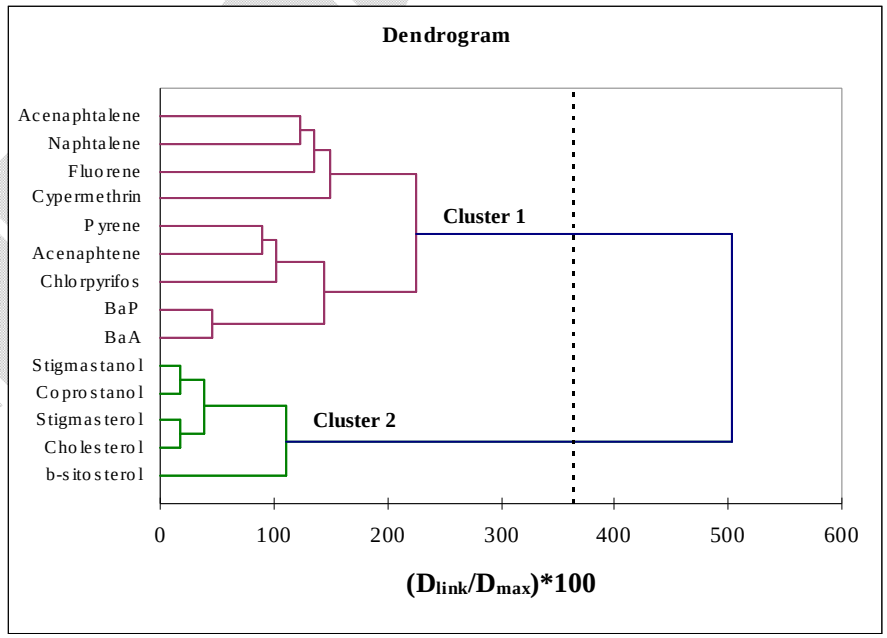

Figure 1. Dendrogram showing the clusters of organic contaminants from point source of pollutant.

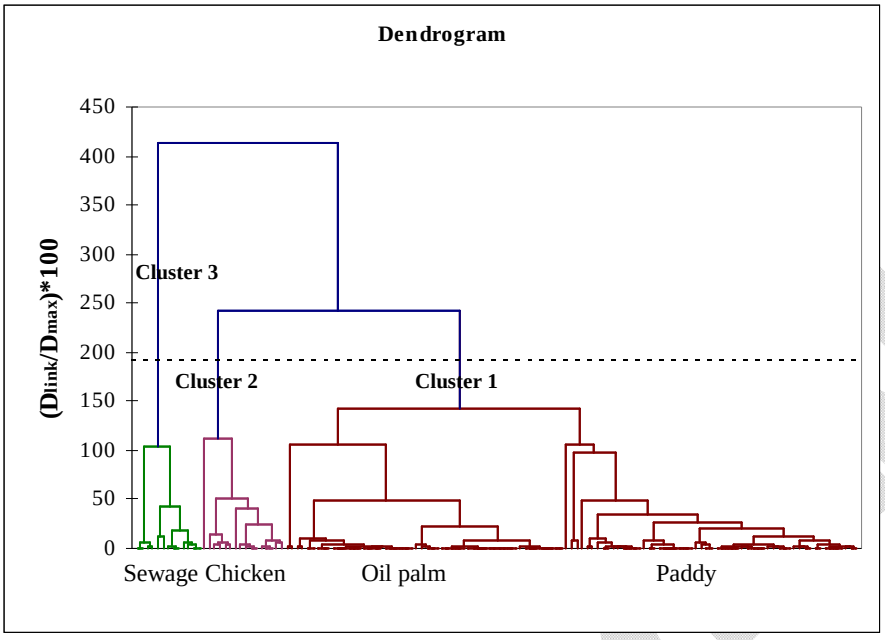

Figure 2. Dendrogram showing the cluster of sampling sites.

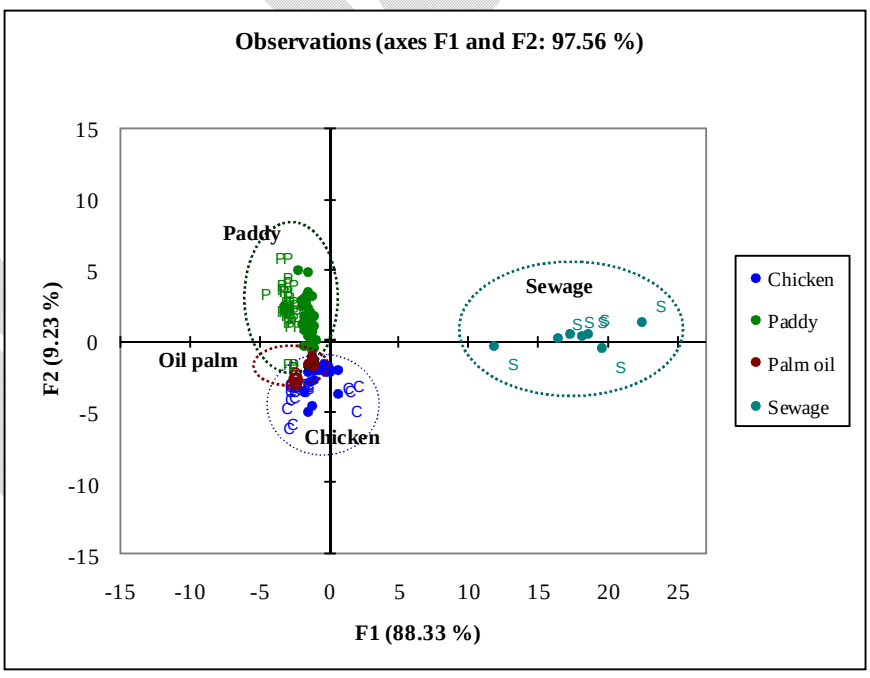

Figure 3. Plot of discriminant functions showing four sources of pollution (chicken, paddy, oil palm and sewage).

3.5 Principal Component Analysis (PCA)

The principal component analysis (PCA) was actually performed on the correlation matrix between the different parameters followed by varimax rotation. The parameter loading for the four components from PCA of the data set after varimax rotation are given in Table 4. The application of PCA in all compounds has provided more visually easier to interpret sources to samples. The PCA separated the compounds according to their sources and distinct clusters and vectors could be determined from the loading (Figure 1a) and score (Figure 1b) plots. The PCA readily sorted out the compounds that co-vary and from a priori knowledge, cluster source type could be identified. When superimposed on the scores, the sample sites could be assigned to particular sources. The initial component matrix shows that strong loading of coprostanol, cholesterol, stigmasterol and stigmastanol are associated in a first varifactor (VF1) which explains 28.0% of the variance. The results are strongly supported by CA analysis whereby these compounds formed one group (cluster 2) which represents faecal contamination. Second, third and fourth varifactors contain a mix-source of organic contaminants explaining 16.1%, 10.0% and 9.7% of the variance, respectively. The predominance of low molecular weight PAHs in the agricultural soils especially paddy fields may be due to the combustion products of low temperature pyrolytic processes such as biomass burning [15,16] and/or petrogenic sources [17]. The presence of these compounds indicates recent pollution, since the lighter PAHs are more biodegradable and less lipophilic and are not expected to persist or be sorbed as strongly as the heavier PAHs. The emission of vehicles used in plantation also contributed to PAHs such as pyrene, benzo[a]anthracene, benzo[a]pyrene [18,19] which explains 16.1% of the variance in VF2.

Table 4. Loadings of organic contaminants on the Varimax rotated PCs for soil samples collected from known source of pollution (paddy field, oil palm, chicken farm and sewage treatment plant).

Parameters	VF1	VF2	VF3	VF4
Naphtalene	-0.129	-0.236	0.419	***0.577***
Acenaphtalene	-0.019	-0.151	*0.478*	0.050
Acenaphtene	-0.184	0.189	***0.708***	0.162
Fluorene	0.260	-0.036	-0.083	***0.635***
Pyrene	-0.082	***0.648***	0.285	-0.187
Benzo[a]anthracene	0.017	**0.787**	-0.097	-0.042
Benzo[a]pyrene	-0.068	**0.875**	-0.010	0.025
Coprostanol	**0.863**	0.005	-0.057	0.218
Cholesterol	**0.956**	-0.096	-0.036	-0.111
Stigmasterol	**0.919**	-0.048	0.040	-0.142
β-sitosterol	0.520	-0.182	-0.008	***-0.540***
Stigmastanol	**0.939**	-0.012	-0.062	0.030
Chlorpyrifos	-0.277	***0.541***	0.279	0.342
Cypermethrin	-0.244	0.033	***-0.560***	0.248
Eigenvalue	4.163	2.136	1.501	1.137
Variability (%)	28.018	16.134	10.025	9.663
Cumulative %	28.018	44.151	54.176	63.839

Note: Strong loadings >0.75 are shown in bold; moderate loading 0.5-0.75 in italic bold.

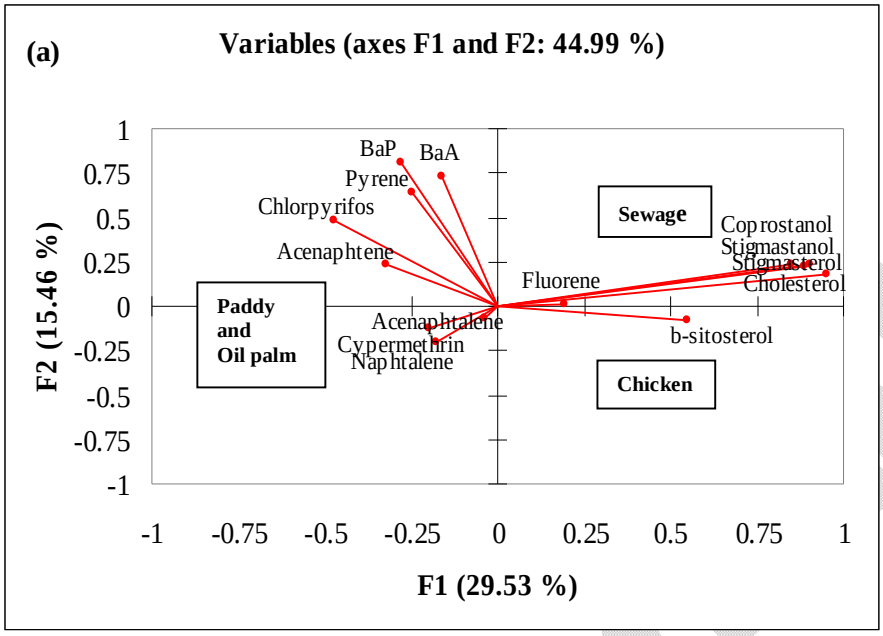

Figure 1. (a) The loadings and (b) the scores on the first two principal components for all compounds.

4. Conclusion

Chemometric techniques (cluster analysis, discriminant analysis and principle component analysis) were used to evaluate variations in organic contaminants (polycyclic aromatic hydrocarbons (PAHs), chlorpyrifos, cypermethrin and sterols) from several sources. Using cluster analysis, sampling sites were grouped into three clusters of similar organic contaminants characteristics. The varifactors obtained from PCA indicated that the parameters responsible in source variations are sterols (coprostanol, cholesterol, stigmasterol, β-sitosterol and stigmastanol), representing faecal contamination (sewage treatment plant and chicken farm) while PAHs (naphthalene, acenaphtene, pyrene, benzo[a]anthracene and benzo[a]pyrene) and pesticides (chlorpyrifos and cypermethrin) representing samples from agricultural activities (paddy fields and oil palm plantations). In order to obtain more precise pollution identification, measurements of other organic compounds from various pollution sources should be included in the study.

5. Competing Interests

The authors declare that they have no competing interests.

6. Author's Contributions

RO conducted the experimental work, sample collection, data analysis and preparation of the manuscript; NS assisted with preparation of the manuscript; MPA helped with chemometric analysis.

7. Acknowledgement

The authors would like to acknowledge financial support obtained from Ministry of Higher Education (MOHE), Malaysia for this project (Project number: 600-IRDC/ST/FRGS 5/3/1318).

References

1. Saim N, Osman R, Spian DRSA, Jaafar MZ, Juahir H, Abdullah MP, 2009. Chemometric approach to validating faecal sterols as source tracer for faecal contamination in water. Water Research, 43: 5023-5030.

2. Leeming R, Nichols P, 1996. Concentrations of coprostanol that correspond to existing bacterial indicator guideline limits. Water Research, 30(12): 2997-3006.

3. Brodnjak-Voncina D, Dobcnik D, Novic M, Zupan J, 2002. Chemometrics characterisation of the quality of river water. Analytica Chimica Acta, 462: 87-100.

4. Kowalkowski T, Zbytniewski R, Szpejna J, Buszewski B, 2006. Application of chemometrics in river water classification. Water Research, 40: 744-752.

5. Shrestha S, Kazama F, 2007. Assessment of surface water quality using multivariate statistical techniques: A case study of the Fuji river basin, Japan. Environmental Modeling and Software, 22: 464-475.

6. Vega M, Pardom R, Barrado E, Ddebaan L, 1998. Assessment of seasonal and polluting effects on the quality of river water by exploratory data analysis. Water Research, 32(12): 3581-3592.

7. Abdul-Wahab SA, Bakheit CS, Al-Alawi SM, 2005. Principal component and multiple regression analysis in modeling of ground-level ozone and factors affecting its concentrations. Environmental Modeling and Software, 20: 1263–1271.

8. Boruvka L, Vecek O, Jenlika, 2005. Principal component analysis as a tool to indicate the origin of potentially toxic elements in soil. Geoderma, 128: 289–300.

9. Osman R, Saim N, Abdullah MP, 2008. Selective accelerated solvent extraction for the analysis of polycyclic aromatic hydrocarbons and sterols from soil. The Malaysian Journal of Analytical Sciences, 12(2): 352-356.

10. Tauler R, Peré-Trepat E, Lacorte S, Barceló D, 2004. Chemometrics Modeling of Environmental Data. Department of Environmental Chemistry, Institute of Chemical and Environmental Research IIQAB-CSIC, Spain.

11. Reimann C, Filzmoser P, 2000. Normal and lognormal data distribution in geochemistry: death of a myth. Consequences for the statistical treatment of geochemical and environmental data. Environmental Geology, 39: 1001–1014.

12. Ong YS, Mahmud AR, Ahmad N, Sulaiman NA, 2009. GIS-based non point source pollution modeling of Selangor Basin. Malaysia GIS: 1-4.

13. Carreira RS, Wagener ALR, Readman JW, 2004. Sterols as markers of sewage contamination in a tropical urban estuary (Guanabara Bay, Brazil): space–time variations. Estuarine, Coastal and Shelf Science, 60: 587–598.

14. Isobe KO, Tarao M, Chiem NH, Minh LY, 2004. Effect of environmental factors on the relationship between concentrations of coprostanol and fecal indicator bacteria in tropical (Mekong Delta) and temperate (Tokyo) freshwaters. Applied and Environmental Microbiology, 70(2): 814–821.

15. Garcia-Falcon MS, Soto-Gonzalez B, Simal-Gandara J, 2006. Evolution of the concentrations of polycyclic aromatic hydrocarbons in burnt woodland soils. Environmental Science and Technology, 40: 759–763.

16. Jenkins BM, Jones AD, Turn SQ, Williams RB, 1996. Emission factors for polycyclic aromatic hydrocarbons from biomass burning. Environmental Science and Technology, 30: 2462–2469.

17. Ping LF, Luo YM, Zhang HB, Li QB, Wu LH, 2007. Distribution of polycyclic aromatic hydrocarbons in thirty typical soil profiles in the Yangtze River Delta region, East China. Environmental Pollution, 147: 358–365.

18. Li J, Zhang G, Li XD, Qi SH, Liu GQ, Peng XZ, 2006. Source seasonality of polycyclic aromatic hydrocarbons (PAHs) in a subtropical city, Guangzhou, South China. Science of the Total Environment, 355: 145–155.

19. Omar NYMJ, Abas MRB, Ketuly KA, Tahir NM, 2002. Concentrations of PAHs in atmospheric particles (PM-10) and roadside soil particles collected in Kuala Lumpur, Malaysia. Atmospheric Environment, 36: 247–254.