Environmental Science and Pollution Research
https://doi.org/10.1007/s11356-021-12772-6

RESEARCH ARTICLE

Estimation of rainwater harvesting by the reflectance of the purity index of rainfall

Siti Nor Fazillah Abdullah[1] · Azimah Ismail[1,2] · Hafizan Juahir[1] · Fathurrahman Lananan[1] · Nor Muzlinda Hashim[1,3] · Nadiana Ariffin[1,4] · Tengku Azman Tengku Mohd[4]

Received: 26 April 2020 / Accepted: 29 January 2021

Abstract
Rainwater harvesting is an effective alternative practice, particularly within urban regions, during periods of water scarcity and dry weather. The collected water is mostly utilized for non-potable household purposes and irrigation. However, due to the increase in atmospheric pollutants, the quality of rainwater has gradually decreased. This atmospheric pollution can damage the climate, natural resources, biodiversity, and human health. In this study, the characteristics and physicochemical properties of rainfall were assessed using a qualitative approach. The three-year (2017–2019) data on rainfall in Peninsular Malaysia were analysed via multivariate techniques. The physicochemical properties of the rainfall yielded six significant factors, which encompassed 61.39% of the total variance as a result of industrialization, agriculture, transportation, and marine factors. The purity of rainfall index (PRI) was developed based on subjective factor scores of the six factors within three categories: good, moderate, and bad. Of the 23 variables measured, 17 were found to be the most significant, based on the classification matrix of 98.04%. Overall, three different groups of similarities that reflected the physicochemical characteristics were discovered among the rain gauge stations: cluster 1 (good PRI), cluster 2 (moderate PRI), and cluster 3 (bad PRI). These findings indicate that rainwater in Peninsular Malaysia was suitable for non-potable purposes.

Keywords Atmospheric pollutant · Physicochemical · Rain gauge stations · Rainwater harvesting · Purity of rainfall index

Introduction

Water is an essential component of our ecosystem, particularly for anthropogenic activities such as domestic supply, irrigation, and industry (Kamarudin et al. 2019; Suhaila and Jemain 2007). In recent years, the increasing population and anthropogenic activities have resulted in a drastic reduction in freshwater supply (Jasrotia et al. 2009). This water scarcity is not only restricted within affected regions but has also become a global issue due to the rapid growth of global population and climatic change (GhaffarianHoseini et al. 2016). Overexploitation of groundwater has caused significant environmental pressures that has led to stress on water supply in many countries. Thus, managing resources sustainably has become a major global concern (Ghimire et al. 2019). An increase in the efficiency and quality of water supply is crucial to address and reduce the growing challenges in managing sustainable water consumption (Kamarudin et al. 2020; Sunardi et al. 2020; Santos and Taveira-Pinto 2013).

The diverse and reliable use of rainwater has been accepted as an alternative solution to reduce water scarcity in several countries (Ghisi et al. 2007). Thus, rainwater harvesting (RWH) can help to provide water which is safer and more sustainable (Man et al. 2014). The RWH system referred to the rainwater that falls on the roof surface is channelled into a gutter and collected in simple storage tanks or barrels via the rainwater downpipe. The RWH for non-potable use has no filtering and purification systems. The RWH system is a part

Responsible Editor: Philippe Garrigues

Azimah Ismail
azimahismail@unisza.edu.my

[1] East Coast Environmental Research Institute, Universiti Sultan Zainal Abidin, Gong Badak Campus, 21300 Terengganu, Malaysia

[2] Faculty of Innovative Design & Technology, Universiti Sultan Zainal Abidin, Gong Badak Campus, 21300 Terengganu, Malaysia

[3] Muadzam Shah Polytechnic, Lebuhraya Tun Razak, 26700 Muadzam Shah, Pahang, Malaysia

[4] Kuala Terengganu Polytechnic, Jalan Sultan Ismail, 20200 Kuala Terengganu, Terengganu, Malaysia

of green technologies as this rainwater, which is a natural resource, can be utilized for agriculture with the integration of monitoring systems. Several studies have revealed that there are multiple benefits of RWH (Helmreich and Horn 2009) as the collected water can be used for non-potable use domestically (Adugna et al. 2018; Rowe 2011) and in agriculture (Yousif and Bubenzer 2015). Additionally, the low cost of installation for individual homes for non-potable indoor usage and the associated water conservation potential has further increased the popularity and global usage of this method.

Peninsular Malaysia receives an abundant rainfall throughout the year as it is situated in the equatorial zone (Wan Zin et al. 2010), placing this country at an advantage for extensive opportunities of using RWH. In recent years, the Malaysian government has launched an RWH policy, which has increased the interest in using RWH in Peninsular Malaysia, especially among researchers (Lani et al. 2018). The policy was aimed at reducing the dependency on treated water and providing a convenient buffer during periods of shortfall in water supply (Lee et al. 2016). However, although RWH is considered the best water management practice in Peninsular Malaysia (Man et al. 2014), the presence of atmospheric pollutants, overhanging foliage, and bird debris can cause contamination to the otherwise clean freshwater harvested from rainfall (Despins et al. 2009). Additionally, the high acidic nature of harvested rainwater can significantly affect crops and human health.

Several studies have been conducted to investigate rainwater quality with respect to chemical and microbial contaminants (De Kwaadsteniet et al. 2013; Lee et al. 2010). The quality of harvested rainwater is based on the characteristics of the region under consideration, such as topography, environmental conditions (Liaw and Chiang 2014), proximity to sources of contamination (Sazakli et al. 2007), form of the catchment area, and water tank type (Adugna et al. 2018). The harvested rainwater has multiple uses, particularly for non-potable household use and irrigation. Rainwater is commonly used as an alternative source of water for activities such as daily consumption, flushing toilets, and outdoor activities like irrigation and plant watering. Therefore, the assessment of atmospheric rainwater pollution is vital to design RWH systems for any non-potable uses.

Multivariate techniques such as principal component analysis (PCA), factor analysis, cluster analysis, and discriminant analysis (DA) have been applied for a better understanding of environmental issues through the interpretation of sophisticated data metrics (Sârbu and Pop 2005). Previous studies have shown that multivariate techniques can be used to accurately assess the quality of groundwater (Monjerezi et al. 2011; Cloutier et al. 2008; Menció and Mas-Pla 2008), river water (Ogwueleka 2015; Mustapha et al. 2012; Fan et al. 2010; Juahir et al. 2009; Panda et al. 2006), marine water (Ismail et al. 2016; Wu et al. 2010; Zhou et al. 2007), and reservoirs (Juahir et al. 2019; Gu et al. 2016; Varol 2013). A study by Sazakli et al. (2007) and Al-Khashman (2009) suggests that PCA and factor component analysis could be used to detect the factors influencing the properties of the rainwater. Niu et al. (2014) used factor analysis and cluster analysis to explore the chemical characteristics of rainwater. However, a comprehensive application of multivariate techniques to analyse data from rainwater or rainfall in Peninsular Malaysia is lacking. A previous study applied factor analysis to develop an index to accurately measure and describe the quality of rainwater (Fazillah et al. 2018).

In this study, the assessment of purity index was based on the historical rainfall data as obtained from the 'clean' rain gauges. The changes in precipitation can impact the performance of RWH systems. Therefore, this study focused on the assessment of the characteristics of rainwater used for non-potable uses by developing a purity of rainfall index (PRI). The PRI can be used as an indicator to determine the level of rainwater quality in the rain gauge stations. Thus, it will act as a guideline to classify the use of harvested rainwater. Harvested rainwater is vital as a recharge system in the hydrological cycle. This study suggests, based on the PRI values, adequate and necessary treatments that could then be implemented based on the purpose of the water use. The availability of PRI facilitates the relevant agencies in monitoring air quality as well. Hence, this study will help to achieve the goal of a sustainable and efficient method to address the problem of water supply.

Data and procedure

Description of rain gauge stations

The study area encompassed 131,587 km^2, which covers the entire area of Peninsular Malaysia, and comprises of 18 rain gauge stations (Fig. 1). Peninsular Malaysia has a tropical rainforest climate, which is warm and humid throughout the year, as evidenced by high annual precipitation, humidity, and temperature. The yearly rainfall in Peninsular Malaysia is approximately between 2000 and 4000 mm, with an annual number of wet days ranging from 150 to 200 (Suhaila and Jemain 2007). Precipitation is distributed in Peninsular Malaysia between two monsoon and two inter-monsoon seasons. The southwest monsoon season occurs from May to August, which usually brings heavy rainfall to the west coast of Peninsular Malaysia, whereas the northeast monsoon season occurs from November to February, causing heavy rain in the northeastern region of Peninsular Malaysia (Suhaila et al. 2010). However, during the inter-monsoon seasons, heavy rainfall usually occurs in the west coast of the Peninsular Malaysia. Substantial rainfall also occurs sparingly throughout the country during transitional periods (Gazzaz et al. 2012).

Fig. 1 Peninsular Malaysia

Data collection

Daily rainfall data for this study was obtained from the Department of Chemistry, Malaysia. The data covered a period of 3 years (2017 to 2019) and included 18 rain gauge stations in Peninsular Malaysia.

Data treatment by statistical technique

The chemical analysis was conducted according to Standard Methods (APHA 1995). The quality measurements from the rain gauges were continuously monitoring. The cations and anions were analysed by the METROHM 881 Compact Ion Chromatograph. For trace metal analysis, a Perkin Elmer NexION 300Q ICP-MS (Perkin-Elmer Inc., USA) fitted with a Perkin Elmer integrated auto-sampler, cross-flow nebulizer, and quartz torch were used. Additionally, the pH and EC of rainwater were calculated using pH meter (Mettler Toledo) and conductivity meter (Mettler Toledo), respectively. All rainfall data obtained from the Department of Chemistry, Malaysia, were pre-treated for subsequent statistical analysis. A total of 30,498 data points (23 variables × 1326 dataset) had been engaged to run the multivariate analysis, using Excel 2013 (Microsoft Office) and the XLSTAT add-in software. The data encompassed 23 physicochemical and metal variables: NH_4^+, Ca^{2+}, F^-, Mg^+, K, Na^+, NO^{-3}, $SO_4\text{-}^2$, $C_2H_3O_2$, Cl^-, CHO_2^-, CH_4O_3S, $C_2O_4(_2^{-})$, Cu, Fe, Mn, Hg, Ni, Cd, conductivity (EC), Pb, pH, and Zn. Multivariate techniques consisted of the PCA, DA, and hierarchical agglomerative cluster analysis and were used to analyse the rainfall data to assess the characteristics of the harvested rainwater.

Principal component analysis

PCA is one of the most efficient multivariate statistical techniques used to reveal a simple underlying structure of a dataset (Ismail et al. 2018; Tahir et al. 2018; Halim et al. 2010). PCA is designed to transform the original variables into new uncorrelated variables called the principal components (PCs), which are linear combinations of the original variables. The new axes lie along with the directions of maximum variance (Shrestha and Kazama 2007). PCs provide information on the entire data without losing the original information (Helena et al. 2000). The PCs can be expressed as

$$z_{ij} = a_{i1}x_{1j} + a_{i2}x_{2j} + \ldots + a_{im}x_{mj} \tag{1}$$

Based on the expression mentioned above, z is the component score, and *a* is the component loading, *x* is the measured value of the variable, *i* is the component number, *j* is the sample number, and *m* is the total number of variables. These PCs were subjected to the varimax rotation generating varactors (VFs) to minimize the complexity of the components by making the significant loadings larger and the minor loadings smaller, within each component. The varimax rotation method is a powerful method to ensure that each variable is only correlated with one component and has almost no correlation with other components (Dominick et al. 2012; Mustapha et al. 2013). In the present study, all principal factors with an eigenvalue more than 1 were taken into account (Cloutier et al. 2008). The VF values higher than 0.75 were considered strong factor loadings. The values ranging from 0.50 to 0.75 were deemed to be moderate, and the values ranging from 0.30 to 0.49 were regarded as weak factor loadings (Liu et al. 2003).

In this study, PCA was applied on 23 physicochemical and metal variables to identify the latent factor contributing to the existing parameters of rainfall data. Each factor can be viewed as one aspect of rainwater. Therefore, factor scores were used as a single index that indicated the element that the factor was associated with. PRI was considered a composite of different variables and was computed by weighting each factor score with the respective variance using the equation mentioned below:

$$PRI = \sum_{i}^{n} F_I w_i \tag{2}$$

where *n* is the number of factors selected, F_I is the factor *i* score, and w_i is the percentage of the variance factor that *i* explains. In this study, univariate clustering was used to execute PRI, which was divided into three categories: good, moderate, and bad. The lowest value of PRI indicates good quality of harvested rainwater, while the highest value of PRI indicates the poorest quality of harvested rainwater. The index categories of PRI were based on physicochemical compounds in the atmosphere depending on the weightage value of each pollutant, as explained below:

> GPRI: The weightage value for this category lies from −18.272 to 35.1851. The lowest value of PRI indicates that the quality of harvested rainwater is good and the atmospheric pollutants do not have a significant effect on the rainwater quality. No treatment is needed for this harvested rainwater.
>
> MPRI: The weightage value for this category lies from 36.661 to 186.169. The rain gauge stations which fall in this category were classified as moderately polluted, though the quality of the harvested rainwater did not vary significantly. Simple treatment methods are required before this water can be used in households.
>
> BPRI: The weightage value for this category lies from 331.140 to 397.997. The high weightage values indicate that atmospheric pollution in the rain gauge station sites resulted in in poor quality of the harvested rainwater. Thus, extensive treatment is needed before this water can be used.

The index values comprise of both negative and positive values, which were rescaled to *z* values (to formulate on a new smaller scale to ensure that variance for each variable would be equal to unity) using the following equation:

$$Rescaling\ (1\ to\ 100) = a + [(xi{-}A)\mathrm{X}(b{-}a)/(B{-}A)] \tag{3}$$

where *a* is equal to 1; *xi* is the actual observation; A and B are the lowest and highest factor scores, respectively; and *b* is a constant with a value of 100.

Discriminant analysis

DA is a multicomponent statistical technique that uses pattern recognition on a set of independent variables to classify objects into exclusive or exhaustive groups (Azemin et al. 2018; Khairuddin et al. 2018; Zolkipli et al. 2018; Gazzaz et al. 2012). The linear combination of the independent variables found through these techniques discriminates the groups in such a way that misclassification error rates are minimized (Lambrakis et al. 2004). The variance-covariance between the classes would be maximized, while the variance-covariance within the classes would be minimized under simultaneous consideration of all types analysed (Kowalkowski et al. 2006; Vaselli et al. 1997). Discriminant factors (DFs) were constructed for each cluster using the following equation:

$$f\ (\mathrm{G}i) = \mathrm{k}i + \sum_{j=1}^{n} w_{ij} p_{ij} \tag{4}$$

where *i* represents the number of groups (G), k*i* is the constant inherent to each group, *n* is the number of parameters used to classify a set of data into a given group, and w_j is the weight coefficient assigned by DF analysis (DFA) to a given parameter (p_j) (Kannel et al. 2007).

In this study, DA was conducted on raw data based on three different modes, standard, forward, and backward stepwise, to construct DFs to evaluate variations in physicochemical properties and metal concentrations in rainfall. The data input to DA constituted 1326 samples. The standard DA mode of constructed DFs contained all parameters. In the stepwise forward mode, variables were added step by step, beginning with the most significant changes. On the other hand, in the stepwise backward mode, the variables were removed step by step, beginning with the least significant or no significant changes.

Hierarchical agglomerative cluster analysis

Cluster analysis is an unsupervised pattern recognition method that groups object into classes, defining underlying behaviour or intrinsic structure of the datasets without making prior assumptions about the potential structure of the data (Wu et al. 2010). The most common cluster analysis method used to assemble objects based on the characteristics is called the hierarchical agglomerative cluster analysis (HACA) (Juahir et al. 2011; Mutihac and Mutihac 2008; Shrestha and Kazama 2007). HACA is a statistical reduction method commonly used to classify variables into clusters, which provides intuitive similarity relationships between any single sample and the entire dataset, which is then illustrated by a dendrogram that illustrates the clusters and proximity (Toriman et al. 2015; McKenna 2003). In this study, HACA was conducted using Ward's method that applied Euclidean distance as a measure of similarity (Fazillah et al. 2017; Toriman et al. 2015; Wu et al. 2010).

Findings and argumentations

Data on rainfall properties

Table 1 provides a summary of the characteristics of rainwater in Peninsular Malaysia from 2017 to 2019. The pH of rainwater from Peninsular Malaysia was acidic (pH 3.39–7.36), with the highest acidity recorded was from Muadzam Shah, and the lowest from Bukit Kledang. The neutral pH value recorded in Bukit Kledang, located in Ipoh, Perak, was possibly due to a higher altitude and cleaner air as compared to the areas with low pH values in Muadzam Shah. This well-planned town was still rapidly developing at the point of this study.

High altitude areas are mostly be covered by forestry, which helps increase the O_2 level and decrease the level of CO_2. Thus, rainwater would only be slightly acidic due to the reaction with CO_2 in the atmosphere (Narany et al. 2014). Acidic rain has a higher concentration of CO_2, whose emissions are higher in the urban areas as compared to those in the rural areas (Igboekwe et al. 2011). High EC levels indicate the dissolved ion contents in the rainwater, and as the amount of the ions present in the atmosphere increases, EC also increases as salinity increases (Liu et al. 2003). The Cl^- showed the highest variance among the rain gauge stations (155.5906 mg/L).

Data on the dimensions of rainfall

Before conducting the PCA, the Kaiser-Meyer-Olkin (KMO) and Barlett's sphericity test were conducted on the parameter correlation matrix to examine the validity of the PCA. The KMO result was 0.72, and Bartlett's test was significant ($p <$ 0.0001), which showed the validated use of PCA. The significant PCs were factors that explained the most variation of rainwater in Peninsular Malaysia. Eight significant PCs, with an eigenvalue greater than 1 (Table 1), were extracted from the variables for a total explained cumulative variance of 70.27%. The variances explained by the individual PCs were 23.84% for PC1, 10.81% for PC2, 9.24% for PC3, 6.79% for PC4, 5.68% for PC5, 5.03% for PC6, 4.52% for PC7, and 4.36% for PC8. The first four PCS were the most significant, explaining a total of 50.68% of the variance in the data. Results from this study indicated that six factors in the PCA accounted for 61.39% of the total variations in the physicochemical properties and metal contents in rainwater. These six factors were used to define the dimensions of rainwater in the study area.

Table 1 shows that all 23 physicochemical and metal parameters used in PCA had a satisfactory 0.5-factor loading threshold. PC1 accounted for the most significant amount of variance at a total of 23.84% and was loaded positively with

Table 1 The concentration of physicochemical and metal parameters in rainwater of Peninsular Malaysia

Parameter	Unit	Minimum	Maximum	Median	Mean	SD
NH_4^+	mg/L	0.050	758.290	2.905	14.863	49.117
Ca^{2+}	mg/L	0.050	289.630	3.085	8.403	21.489
F^-	mg/L	0.025	30.350	0.100	0.563	1.677
Mg^+	mg/L	0.100	255.830	1.220	4.237	15.269
K	mg/L	0.040	122.300	1.910	4.431	9.855
Na^+	mg/L	0.160	2397.060	10.160	35.368	135.111
NO^{-3}	mg/L	0.010	1037.190	8.155	18.178	45.587
$SO_4\text{-}^{2}$	mg/L	0.005	218.560	7.025	11.962	19.086
$C_2h_3O_2$	mg/L	0.025	45.700	0.025	0.495	2.348
Cl^-	mg/L	0.040	2645.630	13.120	42.577	155.591
CHO_2^-	mg/L	0.020	29.740	0.050	0.230	1.153
CH_4O_3S	mg/L	0.015	4.350	0.015	0.071	0.291
$C_2O_4(_{2}-)$	mg/L	0.010	2.450	0.035	0.100	0.188
Cu	mg/L	0.003	3.479	0.014	0.050	0.175
Fe	mg/L	0.035	2.985	0.035	0.113	0.205
Mn	mg/L	0.002	1.352	0.046	0.086	0.121
Hg	mg/L	0.000	0.006	0.000	0.000	0.000
Ni	mg/L	0.015	1.908	0.015	0.041	0.149
Cd	mg/L	0.001	0.292	0.001	0.002	0.009
EC	-	0.009	37.000	1.150	1.850	2.893
Pb	mg/L	0.001	0.592	0.001	0.005	0.023
pH	-	3.390	7.360	4.990	5.053	0.609
Zn	mg/L	0.001	26.545	0.135	0.464	1.231

Abbreviations for physicochemical and metal compounds in this study: NH_4^+ ammonium, Ca^{2+} calcium, F^- fluoride, Mg^+ magnesium, *K* potassium, Na^+ sodium, NO^{-3} nitrate, $SO_4\text{-}^{2}$ sulphate, $C_2h_3O_2$ acetate, Cl^- chloride, CHO_2^- formate, CH_4O_3S methane sulfonic acid, $C_2O_4(_{2}-)$ oxalate, *Cu* copper, *Fe* iron, *Mn* manganese, *Hg* mercury, *Ni* nickel, *Cd* cadmium, *EC* conductivity, *Pb* lead, pH, *Zn* zinc; *SD* standard deviation

Mg^+, Na, Cl^-, and EC, characterized by four variables loaded on a single factor and one variable simultaneously loaded in factor 2. This factor was interpreted to be water-soluble ions, considering the nature of the four rainwater variables. These chemicals originate from natural and anthropogenic aerosols (Bellouin et al. 2005). Natural aerosols were mostly soil dust and marine aerosols, whereas anthropogenic aerosols were emitted from agricultural and industrial activities such as vehicle exhaust, road dust, and biomass burning (Farren et al. 2019). The highest loading value of these ions in factor 1 was in line with the composition of these ions as significant components of atmospheric aerosol (Zhang et al. 2011). There was a significant correlation between the appearance of the ions and EC, where the presence of ions in rainwater increased the electrical conductivity of water (Kazi et al. 2009). The presence of a high concentration of ions reduced the pH value (Al-Badaii et al. 2012).

PC2 was responsible for 10.81% of the total variation in physicochemical properties and metal contents and had the highest number of variable loadings on a single factor. The variables were NH_4^+, K, NO^-_3, SO_4^{2-}, Ni, and EC. This factor may be considered inorganic particles except for K and Ni, which were metals. There was a strong correlation among NH_4^+, NO^-_3, and SO_4^{2-} in the atmosphere, which originated from NH_3^+. The emission of NH_3^+ into the atmosphere has increased due to anthropogenic activities (Ianniello et al. 2011). The NH_3^+ emissions are generally caused by livestock farming, human excrements, synthetic fertilizer, crops, biomass burning in agriculture and biofuel use, fossil fuel combustion, and industrial and anthropogenic emissions (Bouwman et al. 1997). Based on the records of 2010, combustion and industrial sources accounted for the highest proportion of NH_3^+ emissions, with the total emissions of 11.8%, mainly from urban areas (Meng et al. 2017).

Anthropogenic activities may influence the acidity levels in rainwater, resulting in the emissions of acidic gasses like SO_2, NO_X, and NH_3. These gasses form sulphuric and nitric acids when dissolved in the atmosphere (Kulshrestha et al. 2003) and significantly influence the acidity of the atmosphere (Zhang et al. 2011). The emission of NH_3^+ occurs when exposed to air. The reaction of NH_3^+ with SO_2 and NO_X contribute to the formation of secondary inorganic aerosols, including NH_4^+, NO^-_3, and SO_4^{2-} (Zhao et al. 2019a, b). The K^+ ion content was also a positively loaded factor, which suggested significant emissions from urban and industrial sources such as wood smoke, combustion of solid fuels, construction and demolition, soil dust, sea salt, and coal fire (Pachon et al. 2013).

Ni was loaded simultaneously in PC6 with other heavy metals. Ni is emitted into the ambient air at a high temperature through natural and anthropogenic sources (Harasim and Filipek 2015). The release of Ni into the atmosphere can be both in a particulate matter or gaseous forms that attach to other fine particles (Tian et al. 2012). The emissions of Ni increased due to transportation and industrialization (Cempel and Nikel 2006). However, natural sources of Ni emissions include windborne dust particles, vegetation, sea salt, and forest fires (Tian et al. 2012). The deposition of this metal in the atmosphere neutralizes the effect of acidity from sulphur and nitrate (Simmons et al. 2001).

The third group (PC3), with a total variance of 9.24%, received high loadings from $C_2H_3O_2$ and CHO_2^-, with a moderate load on CH_4O_3S in the rainwater, which was interpreted to be organic acid. Both $C_2H_3O_2$ and CHO_2^- anions had strong factor loadings in factor 3 from acetic acid and formic acid, which was known as a carboxylic acid. Generally, the existence of carboxylic acids such as $C_2H_3O_2$ and CHO_2^- were found in abundance within the environment, both in urban and rural areas (Kerminen et al. 2000). The sources of these acids may vary, based on the exposure of biomass burning fossil fuel combustion, sea spray, and transportation and industrial emissions as well as photochemical oxidation of precursors of anthropogenic and biogenic origins (Alwe et al. 2019). The acidity of the rainwater was because of the presence of water-soluble carboxylic acids.

On the other hand, methanesulfonic acid (MSA) had a moderate loading in factor 3 (0.58), with a strong correlation with the atmospheric sulphur cycle in the marine environment (Henriques and De Marco 2015). MSA was the primary atmospheric photo-oxidation product from the gaseous dimethylsulfide (DMS) (Hodshire et al. 2019). The sulphur emitted from the oceans primarily in the form of DMS was produced through the sulphur cycle during the biological process in the oceans and in agricultural, industrial, and domestic activities (Dawson et al. 2012). Once in the atmosphere, DMS was oxidized by OH to produce SO_2 and MSA.

The fourth factor accounted for 6.79% of the variance in the rainfall, characterized by mineral ions (Ca^{2+} and F^- with factor loadings of 0.79 and 0.57, respectively) and metal (factor loading of 0.52 for Zn). Sources of these compounds were, however, varied. The Ca^{2+} ions originated from crustal source, in which some anthropogenic activities such as construction and heavy traffic might have contributed to the presence of Ca^{2+} in the atmosphere (Wang et al. 2006a, b). The F^- ions could enter the atmosphere from both natural and anthropogenic sources (Chakraborty and Gupta 2010). Nonetheless, natural sources such as volcanic eruption, rock dust, and marine aerosols contributed a small portion of this compound to the atmosphere (Walna et al. 2007). Among anthropogenic processes, F^- was mainly emitted to the atmosphere from phosphate fertilizer production, aluminium reduction, metal smelting, and other industrial activities including brick, tile, pottery, cement works, ceramic industry, and glass manufacturing (Walna et al. 2013). The existence of Zn in the harvested rainwater also varied as Zn can be sourced from road dust, tire wear of vehicles, galvanizing activities, incineration, and fossil fuel combustion (Wang et al.

2006a). Additionally, past studies have revealed that a possible reason for the concentration of Zn might be the result of contamination from using metal rooftops to collect rainwater (Van Metre and Mahler 2003).

The fifth factor explained 5.68% of the variance in the dataset, distinguished by the high loading from Fe (0.84) and moderate loading from oxalate (0.50), and Mn^+ (0.65) and negative loading from pH (0.51). The pH level was negatively correlated as the heavy metals (Fe and Mn^+) decreased the acidity of rainwater (Simmons et al. 2001). Fe and Mn^+ had crustal origins and were emitted primarily from industrial metallurgical processes (Wang et al. 2006a, b). Interestingly, Fe had the highest loading in this factor, mainly due to exhaust emissions of gasoline and other fuels in transportation. On the other hand, Mn^+ had a moderate load in this factor, which was usually emitted from the combustion of fossil fuels and incinerator ash (Moreno et al. 2011; Allen et al. 2001). The significant correlation between Fe and Mn^+ suggested that these metals had two common sources, motor vehicles, and industrial activities (Kabadayi and Cesur 2010). Besides, oxalate was an organic ion produced through photochemical reactions that occurred in the atmosphere, which was strongly associated with sea salt particles in the atmosphere (Kerminen et al. 2000). The interaction of organic ion with metals may be an essential parameter in controlling the acidity of rainwater, as supported through the negative correlations between Fe and Mn^+ with pH.

The sixth factor accounted for approximately 5.0% of the variance in the data, which showed high loading from Pb (0.80) but moderate loadings from Cu (0.59) and Ni (0.57), and were interpreted as toxic trace metals. These metals were found to be tightly bonded and usually originated from natural and anthropogenic sources. In the natural process, trace metals were emitted from crustal minerals, including erosion, surface winds, volcanism, forest fire, and the oceans (Wang et al. 2006a, b). Among anthropogenic activities, trace metals were emitted through industrial processes such as combustion of fossil fuels and wood, vehicular exhaust, and waste incineration (Allen et al. 2001). These metals could also be emitted from the exhaust of gasoline and diesel-fuelled road vehicles (Allen et al. 2001). The highest loading for Pb indicated that there were diverse sources of emission, such as from manufacture of paint and batteries (Zhang et al. 2002), whereas the lower factor loadings of Cu and Ni demonstrated that emissions from vehicles were not significant (Munim et al. 2014).

Lastly, both of the seventh and eighth factors receive only single digit loadings from Cd and Hg. These factors indicated the presence of heavy metals in the rainwater, which might be due to the interaction between the type of material used for the rooftops and the collected rainwater. The contamination of heavy metals occurred when rainwater that was collected on the roof dissolved the metal component from the surface of the material used for the rooftops (Chang et al. 2004). However, as the value of variance was less than 5.0%, it can be assumed that the underlying construct of rainfall data in Peninsular Malaysia was associated with six, rather than eight, factors, accounting approximately for 61.4% of the variations in the dataset. Therefore, the heavy metals (Cd and Hg) that were detected in PC7 to PC8 were not considered to be significant in this study (Reimann et al. 2002).

In this study, PCA did not result in significant data reduction, as 21 parameters (about 91% of the 23 parameters) explained 61% of the data variance (Table 2). However, PCA served as a means to identify these parameters, which contributed significantly to the spatial variation in the rainwater. The findings showed that the chemical property of rainwater within the study area was not significantly affected by heavy metals. Hence, the rainwater was considered safe for non-potable human consumption.

Determination of the PRI

Apart from determining the variation in the dataset, PCA was used to develop an index for the rainfall data using factor scores generated through PCA. Based on the value of rescaling, the lowest value indicated good PRI, which contained a small number of chemical properties, while the higher values indicated high loads of chemical properties. Of a total of 1366 rainfall datasets, 1249 were found to be in the 'good' category, 73 were in the 'moderate' category, and 4 were under the 'bad' category. Generally, the PRI values obtained in this study indicated that the rainwater could be harvested and is safe for non-potable uses.

Index variation of rainfall

Index variations of physicochemical and metal parameters of the rainfall from different rain gauge stations were further evaluated through DA. The significance of discriminant function and the most significant variables associated with the differences between scale indices were assessed using the dataset based on the PRI developed using the PCA. The indices were treated as dependent variables, and the measured physicochemical and metal parameters were treated as independent variables. The classification functions (DFs) and classification matrices (CMs) obtained from standard, forward, and backward modes are listed in Table 3 and 4.

The standard DA mode was used to construct DFs that contained all parameters. The coefficient of sulphate and chloride were close to zero (Table 3). As shown in Table 4, all modes of DA correctly assigned more than 98% of the cases with different variables. In a standard mode that uses 23 discriminant variables, CM correctly assigned 98.87% of the cases. Through the forward and backward stepwise mode on 17 and 19 discriminant variables, 98.04% and 98.11% of the

Table 2 Eigenvalues from the principal component analysis illustrating variance, cumulative variance and factor loading

Parameter	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
NH_4^+	0.132	0.850	-0.045	0.048	0.056	-0.016	0.018	0.050
Ca^{2+}	0.197	0.194	0.070	0.791	0.013	0.010	0.000	-0.001
F^-	-0.033	0.067	0.143	0.568	0.169	0.029	-0.250	-0.228
Mg^+	0.985	0.073	0.007	0.066	-0.016	0.030	0.018	0.005
K	0.476	0.601	0.047	0.298	-0.060	-0.007	0.050	0.033
Na^+	0.989	0.016	-0.002	0.004	-0.015	0.035	0.012	0.007
NO^{-3}	0.072	0.828	0.009	0.047	0.017	0.160	0.079	-0.003
$SO_4{-}^2$	0.442	0.551	-0.051	0.184	0.308	-0.031	-0.158	-0.096
$C_2h_3O_2$	0.000	-0.038	0.791	0.032	0.044	0.013	0.037	0.035
Cl^-	0.988	0.025	-0.001	0.022	-0.012	0.030	0.004	0.000
CHO_2^-	0.019	0.017	0.834	-0.046	0.044	0.017	-0.013	0.006
CH_4O_3S	0.017	0.001	0.581	0.243	0.146	-0.077	-0.056	-0.019
$C_2O_4(_2-)$	-0.009	-0.120	0.330	0.214	0.502	-0.049	0.150	-0.074
Cu	0.287	0.181	-0.026	0.228	0.099	0.589	0.126	0.047
Fe	-0.015	0.023	0.070	-0.023	0.836	0.082	0.006	0.025
Mn	0.157	0.154	-0.017	0.188	0.650	0.094	0.189	0.162
Hg	-0.016	0.025	0.025	-0.009	0.031	-0.006	-0.073	0.919
Ni	-0.022	0.565	0.018	-0.023	-0.068	0.575	0.185	0.002
Cd	0.002	0.074	0.009	-0.002	0.066	-0.009	0.886	-0.080
EC	0.788	0.520	0.025	0.077	0.199	0.076	-0.047	-0.040
Pb	0.074	-0.002	0.010	-0.034	0.126	0.808	-0.128	-0.035
pH	0.035	-0.173	-0.140	0.497	-0.507	-0.137	0.149	0.277
Zn	0.028	0.205	-0.078	0.523	0.162	0.356	0.149	0.042
Eigenvalue	5.483	2.486	2.124	1.563	1.306	1.158	1.040	1.002
Variability (%)	23.841	10.810	9.236	6.794	5.678	5.033	4.521	4.358
Cumulative %	23.841	34.651	43.887	50.682	56.360	61.392	65.913	70.271

*Red colour showed a significant positive loading; yellow indicated moderate positive loading and green represented moderate negative loading

CMs were reported as correct, respectively. However, CM produced 98.04% of correct assignments in the stepwise forward mode that used 17 discriminant parameters. This result was similar to those obtained from standard and backward modes, but with fewer parameters. The stepwise backward mode that included Na^{2+}, EC, and pH as additional variables, with the exclusion of Ca^{2+}, had a higher spatial variation.

The Pillai's Trace test for standard, forward, and backward modes revealed a trace value of 1.4323 ($p < 0.0001$), 1.4291 ($p < 0.0001$), and 1.4300 ($p < 0.0001$), respectively. The null hypothesis (H_0) stated that the mean vectors of the three classes were equal. The alternative hypothesis (H_a) stated that at least one of the mean vectors was different from the others. As the computed p value was lower than the significance level alpha (0.05), the null hypothesis was rejected. The risk to reject H_0 was true only when the result was lower than 0.01%. Thus, the DA suggested that NH_4^+, Ca^{2+}, Mg^+, K, NO^{-3}, $SO_4{-}^2$, $C_2H_3O_2$, Cl^-, CHO_2^-, CH_4O_3S, Cu, Mn, Hg, Ni, Cd, Pb, and Zn were sufficient to discriminate the three groups and accounted for most of the index scale variations in rainwater.

Spatial classification of rain gauge stations

HACA was conducted on the rainfall dataset to evaluate spatial variation among 18 rain gauge stations. The clustering procedure generated three clusters, and the stations in these clusters had similar characteristics and natural backgrounds. This analysis resulted in the grouping of rain gauge stations in three clusters within a hierarchical dendrogram at $(D_{ink}/D_{max}) \times 100 < 15$ (Fig. 2). Cluster 1 (Alor Setar, Batu Embun, Bayan Lepas, Bukit Kledang, Chuping, Kluang, Kota Bharu, Kuala Terengganu, Kuantan, Melaka, Mersing, Petaling Jaya, Senai, Sitiawan, and Tanah Rata), cluster 2 (Bachok), and cluster 3

Table 3 Classification function coefficients for discriminant analysis

Parameters	Standard mode			Forward stepwise mode			Backward stepwise mode		
	1	2	3	1	2	3	1	2	3
NH_4^+	-0.283	-0.278	-0.379	-0.004	0.014	-0.006	-0.198	-0.197	-0.313
Ca^{2+}	-0.465	-0.439	-0.316	0.013	0.061	0.280			
F^-	0.621	0.427	-1.174						
Mg^+	-0.131	0.277	-1.152	-0.169	0.126	-1.913	-1.964	-1.411	-2.182
K	-0.505	-0.245	-0.973	0.065	0.365	-0.154	-0.296	-0.047	-0.860
Na^+	-0.088	-0.136	-0.426				0.108	0.051	-0.259
NO^{-3}	0.100	0.081	0.139	0.003	-0.018	0.035	0.078	0.060	0.120
$SO_4\text{-}^2$	-0.019	0.036	-0.480	0.033	0.110	-0.231	-0.018	0.034	-0.505
$C_2h_3O_2$	0.549	1.709	4.104	0.204	1.360	3.732	0.314	1.497	3.952
Cl^-	-0.022	0.020	1.088	0.026	0.044	0.964	0.024	0.058	1.072
CHO_2^-	0.487	1.847	3.146	0.232	1.586	2.821	1.243	2.553	3.577
CH_4O_3S	5.123	13.745	20.545	1.294	9.942	15.587	3.299	12.272	19.054
$C_2O_4({}_2-)$	10.635	13.400	17.404						
Cu	-1.212	10.432	59.207	1.812	13.406	62.571	-3.317	8.348	56.453
Fe	1.199	-0.543	-3.636						
Mn	10.444	23.549	62.046	6.047	18.331	54.901	11.811	24.489	62.292
Hg	2508.997	12560.595	31239.984	6408.967	16454.392	35093.370	3532.876	13469.855	31841.362
Ni	-2.513	4.045	-8.256	0.694	7.992	0.386	2.432	8.763	-4.112
Cd	83.717	267.938	597.796	37.273	221.248	535.377	101.754	288.482	622.065
EC	11.716	12.502	17.492				8.096	9.013	14.802
Pb	26.394	70.908	267.251	8.318	49.309	229.189	33.250	76.309	266.787
pH	32.696	33.844	39.590				24.643	26.168	33.710
Zn	0.345	0.864	0.502	0.091	0.586	0.098	-0.634	-0.087	-0.377

Table 4 Classification matrix for discriminant analysis of Index variation

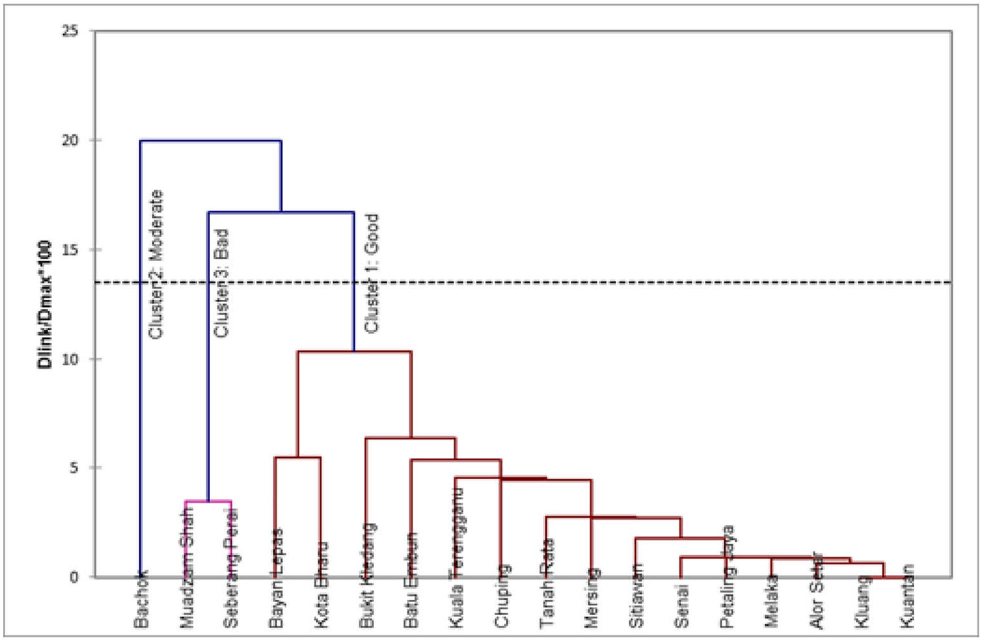
Index scale	% Correct	Predicted index scale		
		Good	Moderate	Bad
Standard mode				
Good	99.84	1247	2	0
Moderate	82.19	13	60	0
Bad	100	0	0	4
Total	98.87	1260	62	4
Forward stepwise mode				
Good	99.28	1240	9	0
Moderate	76.71	17	56	0
Bad	100	0	0	4
Total	98.04	1257	65	4
Backward stepwise mode				
Good	99.60	1244	5	0
Moderate	72.60	20	53	0
Good	100	0	0	4
Total	98.11	1264	58	4

(Muadzam Shah and Seberang Perai) corresponded to good, moderate, and bad PRI, respectively (Fig. 3). The results from HACA implied that rainwater samples could be collected regularly from only one station within each cluster, as the selected station in the cluster represents the entire cluster.

Cluster 1 was classified as having a good PRI, as the results showed these regions having the least contribution of physicochemical and metal parameters as compared to the other clusters. This cluster comprised of 15 rain gauge stations which were scattered across the regions of Peninsular Malaysia. These regions were characterized by the residential areas, city centre with traffic congestions, as well as the airport, commercial, and industrial activities, which were considered under good PRI.

Cluster 2 represented moderate PRI, with only Bachok being placed in this cluster. Bachok is a rural area, with paddy cultivation being the main activity for the residents (Shamsudin 2013). There was a significant correlation between the location of rain gauge stations and the condensation of the ionic compounds of the atmosphere. The composition of water-soluble ions in the atmosphere was considered to be the most significant factor influencing the variation of PRI in this study. As Bachok in

Fig. 2 The plot of discriminant functions for rainwater based on the index value of 1 = good, 2 = moderate, and 3 = bad



Peninsular Malaysia was located at the border of the southern edge of the South China Sea, it was believed that the continental sources contributed to the emission of the ionic compound in the atmosphere (Farren et al. 2019). Moreover, undersea activities that included photo-oxidation and biogenic process introduced highly ionic compounds into the atmosphere (Miranda and Tomaz 2008). Bachok was also exposed to the marine atmosphere, whereby anthropogenic sources highly affected the atmospheric composition of the area. Other sources of anthropogenic activities such as agriculture and dust from the road also contributed to the emission of the pollutant to the atmosphere (Dominick et al. 2015). Bachok is dominantly an agricultural area, with limited industrialization. However, this area was found to be contaminated with heavy metals from industrial activities that surrounded the rain gauge stations, such as vehicular emissions, biomass burning, and fossil fuel combustion. Additionally, the natural sources of heavy metals suspended in the surface dust were also emitted to the atmosphere (Allen et al. 2001).

Cluster 3, which consisted of Seberang Perai and Muadzam Shah, was reported to have bad PRI, possibly due to emissions

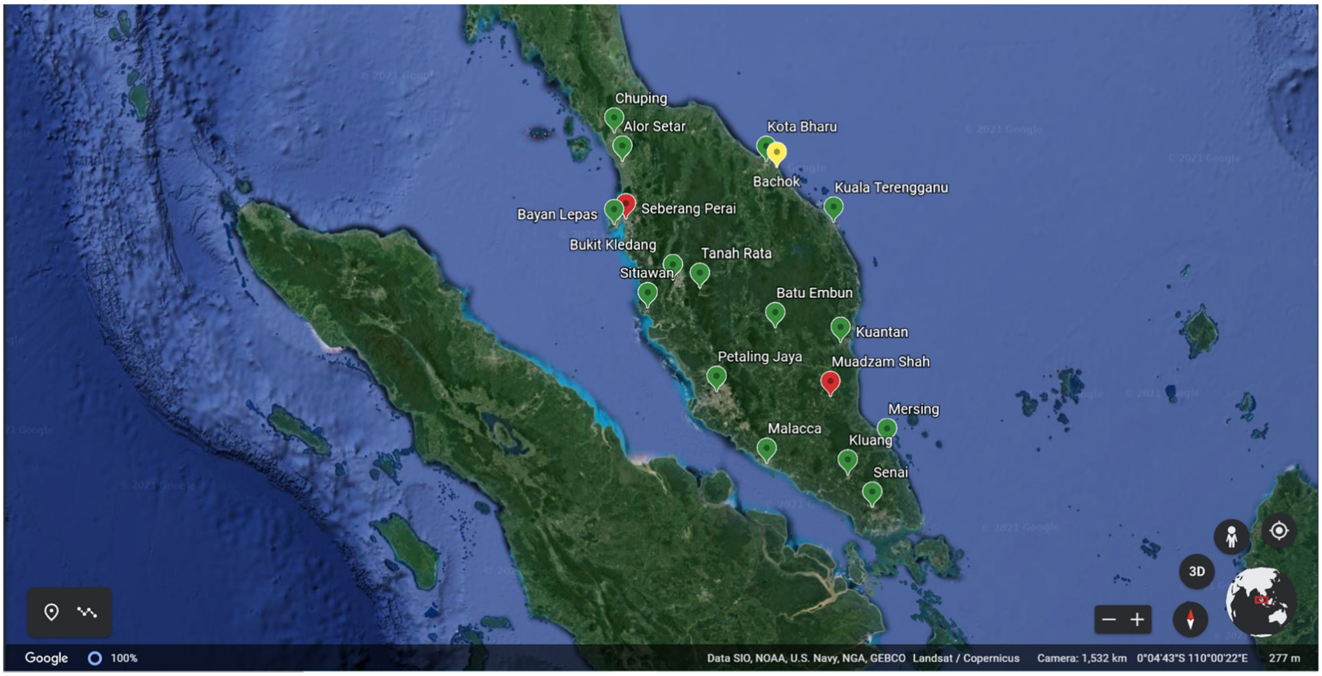


Fig. 3 Location of rain gauge station based on clusters: cluster 1 (green), cluster 2 (yellow), and cluster 3 (red). Source: Google earth

from both marine and non-marine sources (Farren et al. 2019). In contrary to cluster 2, which was dominantly exposed to the marine atmosphere, Seberang Perai and Muadzam Shah were more urbanized even though both regions were located near the coasts. Hence, both the land and marine atmospheres had contributed to the deposited chemical compounds in the atmosphere. Seberang Perai, located on the northwest coast of Peninsular Malaysia, has been undergoing rapid industrialization (Samat 2007). Because of the strategic location in the coastal area, this region had experienced dense-development and high commercial concentration as a centre of both economic and cultural activities (Zhao et al. 2019a, b). Seberang Perai has also been the hub for marine transportation after replacing Georgetown for port activities (Norsukhairin et al. 2013). Thus, this region had various types of motor vehicles and water transportation that could emit high amounts of heavy metals. On the other hand, Muadzam Shah is located in Pahang, a rural area. However, this is a well-planned region that is rapidly growing with industrial activities with the aim of becoming a satellite city (Rose et al. 2016). Additionally, Muadzam Shah is also known as an education hub in Pahang, whereby many schools and higher education institutions have been established (Shamsuddin et al. 2014). The region was therefore dense with vehicles that contributed to the emission of heavy metals from vehicle exhaust, as well as other chemical compounds from road dust.

The primary land uses in Seberang Perai are agriculture, forestry, industry, and residence. Past studies have reported that activities such as factories and agriculture accounted for a significant proportion of pollutants in this region (Mohd Nasir et al. 2011). On the other hand, the main agricultural activities within Seberang Perai included palm oil cultivation, market gardening, paddy, and agro-based business activities (Eltayeb Elhadary et al. 2013), which mainly contributed to ammonia emissions (Bauer et al. 2016). Livestock farming was also identified to be one of the significant sources of ammonia in the atmosphere (Paoli et al. 2010) due to the emissions resulting from animal houses and associated storage systems, animal manure, and animal grazing (Beusen et al. 2008). The industrialization of poultry and swine industry in Seberang Perai, as well as the largest cattle farm in Pahang located in Muadzam Shah, can be linked to the emission of nitrogen, phosphate, and ammonia (Mallin and Cahoon 2003).

Conclusion

This study has demonstrated the reliability and efficiency of harvested rainwater to support sustainable water supply in Peninsular Malaysia. The implementation of RWH was found to be reliable and efficient in most local and urban regions. Analysis of physicochemical properties of and metals in the rainwater collected from 18 gauge stations was conducted using PCA, where six PFs were found to be significant, accounting for a total variance of 61.39%. Furthermore, the emissions varied between multi-source factors, which include both natural and anthropogenic activities. The contaminations were possibly derived from industrial emissions, transportation, agricultural activities, fuel combustion, and marine atmosphere. This study also identified the atmospheric particulates containing organic and inorganic compounds that contributed to the variations in physicochemical properties and metal contents. These compounds were deposited in the atmosphere and diluted in the rainwater during rainfall. The PRI was computed from the factor scores generated in PCA within three categories: good, moderate, and bad. This was confirmed by the DA, which showed a 98.04% accuracy of spatial variation for stepwise forward mode, which yielded the 17 most significant variables. The HACA clustered the 18 rain gauge stations into three significant cluster regions: good, moderate, and bad PRI. The coastal and urbanized areas showed high contamination of the physicochemical and metal parameters in rainwater. Additionally, the rain gauge stations situated near coastal areas were also exposed to marine. However, RWH can be implemented in larger scales due to the characteristics of rainwater in Peninsular Malaysia and was found safe for non-potable uses.

Acknowledgement The authors would like to express their appreciation to the Department of Chemistry, Malaysia and Meteorology Department, Malaysia, for providing the data required for the completion of this study. The authors wish to thank the East Coast Environmental Research Institute (ESERI), UniSZA, for the use of their research facilities.

Authors' contributions Abdullah S. N. F. wrote the manuscript.
Ismail A. revised the manuscript.
Juahir H. revised the manuscript.
Lananan F. revised the manuscript.
Hashim N. M. edited the manuscript.
Ariffin N. edited the manuscript.
Mohd T. A. T. edited the manuscript.

Funding This study is a part of the Special Research Grant Scheme UniSZA/2017/SRGS/18.

Data availability The datasets generated and analysed during the current study are not publicly available due to the protection of confidential information but are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate Not applicable.

Consent for publication Not applicable.

Competing interests The authors declare no competing interests.

References

Adugna D, Jensen MB, Lemma B, Gebrie GS (2018) Assessing the potential for rooftop rainwater harvesting from large public institutions. Int J Environ Res Public Health 15(2):336. https://doi.org/10.3390/ijerph15020336

Al-Badaii F, Gasim MB, Mokhtar M, Toriman ME, Rahim SA (2012) Water-pollution study based on the physico-chemical and microbiological parameters of the Semenyih River, Selangor, Malaysia. Arab World Geographer 15(4):318–334. https://doi.org/10.5555/arwg.15.4.hn3076665801k728

Al-Khashman OA (2009) Chemical characteristics of rainwater collected at a western site of Jordan. Atmos Res 91(1):53–61. https://doi.org/10.1016/j.atmosres.2008.05.007

Allen AG, Nemitz E, Shi JP, Harrison RM, Greenwood JC (2001) Size distributions of trace metals in atmospheric aerosols in the United Kingdom. Atmos Environ 35(27):4581–4591. https://doi.org/10.1016/S1352-2310(01)00190-X

Alwe HD, Millet DB, Chen X, Raff JD, Payne ZC, Fledderman K (2019) Oxidation of volatile organic compounds as the major source of formic acid in a mixed forest canopy. Geophys Res Lett 46(5):2940–2948. https://doi.org/10.1029/2018GL081526

APHA (1995) Standard methods for the examination of water and wastewater, 19th edn. American Public Health Association Inc., New York

Azemin A, Juahir H, Hamdan MR, Ismail Z, Mohd KS, Ghazali A, Abdullah MH, Ismail A, Elfithri R (2018) Pattern recognition for varieties of Malaysian Herb, Ficus deltoidea Jack through chemometric applications from GC-MS fingerprinting. Int J Eng Technol (UAE) 7(3.14 Special Issue 14):348–355

Bauer SE, Tsigaridis K, Miller R (2016) Significant atmospheric aerosol pollution caused by world food cultivation. Geophys Res Lett 43(10):5394–5400. https://doi.org/10.1002/2016GL068354

Bellouin N, Boucher O, Haywood J, Reddy MS (2005) Global estimate of aerosol direct radiative forcing from satellite measurements. Nature 438(7071):1138–1141. https://doi.org/10.1038/nature04348

Beusen AHW, Bouwman AF, Heuberger PSC, Van Drecht G, Van Der Hoek KW (2008) Bottom-up uncertainty estimates of global ammonia emissions from global agricultural production systems. Atmos Environ 42(24):6067–6077. https://doi.org/10.1016/j.atmosenv.2008.03.044

Bouwman AF, Lee DS, Asman WAH, Dentener FJ, Van Der Hoek KW, Olivier JGJ, Tg N (1997) A global high-resolution emission inventory for ammonia. Glob Biogeochem Cycles 11(4):561–587

Cempel M, Nikel G (2006) Nickel: A review of its sources and environmental toxicology. Pol J Environ Stud 15(3):375–382

Chakraborty A, Gupta T (2010) Chemical characterization and source apportionment of submicron (PM 1) aerosol in Kanpur Region, India. Aerosol Air Qual Res 10(5):433–445. https://doi.org/10.4209/aaqr.2009.11.0071

Chang M, McBroom MW, Scott Beasley R (2004) Roofing as a source of nonpoint water pollution. J Environ Manag 73(4):307–315. https://doi.org/10.1016/j.jenvman.2004.06.014

Cloutier V, Lefebvre R, Therrien R, Savard MM (2008) Multivariate statistical analysis of geochemical data as indicative of the hydrogeochemical evolution of groundwater in a sedimentary rock aquifer system. J Hydrol 353:294–313. https://doi.org/10.1016/j.jhydrol.2008.02.015

Dawson ML, Varner ME, Perraud V, Ezell MJ, Gerber RB, Finlayson-Pitts BJ (2012) Simplified mechanism for new particle formation from methanesulfonic acid, amines, and water via experiments and ab initio calculations. Proc Natl Acad Sci U S A 109(46):18719–18724. https://doi.org/10.1073/pnas.1211878109

De Kwaadsteniet M, Dobrowsky PH, Van Deventer A, Khan W, Cloete TE (2013) Domestic rainwater harvesting: microbial and chemical water quality and point-of-use treatment systems. Water Air Soil Pollut 224(7). https://doi.org/10.1007/s11270-013-1629-7

Despins C, Farahbakhsh K, Leidl C (2009) Assessment of rainwater quality from rainwater harvesting systems in Ontario, Canada. J Water Supply: Res Technol - AQUA 58(2):117–134. https://doi.org/10.2166/aqua.2009.013

Dominick D, Juahir H, Latif MT, Zain SM, Aris AZ (2012) Spatial assessment of air quality patterns in Malaysia using multivariate analysis. Atmos Environ 60:172–181. https://doi.org/10.1016/j.atmosenv.2012.06.021

Dominick D, Latif MT, Juneng L, Khan MF, Amil N, Mead MI, Nadzir MSM, Moi PS, Samah AA, Ashfold MJ, Sturges WT, Harris NRP, Robinson AD, Pyle JA (2015) Characterisation of particle mass and number concentration on the east coast of the Malaysian Peninsula during the northeast monsoon. Atmos Environ 117:187–199. https://doi.org/10.1016/j.atmosenv.2015.07.018

Eltayeb Elhadary YA, Samat N, Obeng-Odoom F (2013) Development at the peri-urban area and its impact on agricultural activities: an example from the Seberang Perai region, Penang state, Malaysia. Agroecol Sustain Food Syst 37(7):834–856. https://doi.org/10.1080/21683565.2013.797950

Fan X, Cui B, Zhao H, Zhang Z, Zhang H (2010) Assessment of river water quality in Pearl River Delta using multivariate statistical techniques. Procedia Environ Sci 2(5):1220–1234. https://doi.org/10.1016/j.proenv.2010.10.133

Farren NJ, Dunmore RE, Mead MI, Nadzir MSM, Samah AA, Phang SM, Bandy BJ, Sturges WT, Hamilton JF (2019) Chemical characterisation of water-soluble ions in atmospheric particulate matter on the east coast of Peninsular Malaysia. Atmos Chem Phys 19(3):1537–1553. https://doi.org/10.5194/acp-19-1537-2019

Fazillah A, Toriman E, Juahir H, Izwan A, Mohamad N, Ismail A (2017) Evaluation of socioeconomic status on drug addicts. Int J Adv Sci, Eng, Inform Technol 7(6):2122–2130

Fazillah A, Juahir H, Toriman E, Mohamad N, Kamarudin MKA, Fairuz A, Omar SMH, Adi-ana G, Mahadzirah M (2018) Family support indices for substance abuser in Terengganu. Int J Eng Technol 7(4):2655–2662. https://doi.org/10.14419/ijet.v7i4.10452

Gazzaz NM, Yusoff MK, Ramli MF, Aris AZ, Juahir H (2012) Characterization of spatial patterns in river water quality using chemometric pattern recognition techniques. Mar Pollut Bull 64(4):688–698. https://doi.org/10.1016/j.marpolbul.2012.01.032

GhaffarianHoseini A, Tookey J, GhaffarianHoseini A, Yusoff SM, Hassan NB (2016) State of the art of rainwater harvesting systems towards promoting green built environments: a review. Desalin Water Treat 57(1):95–104. https://doi.org/10.1080/19443994.2015.1021097

Ghimire SR, Johnston JM, Garland J, Edelen A, Ma X(C), Jahne M (2019) Life cycle assessment of a rainwater harvesting system compared with an AC condensate harvesting system. Resour Conserv Recycl 146(April):536–548. https://doi.org/10.1016/j.resconrec.2019.01.043

Ghisi E, Bressan DL, Martini M (2007) Rainwater tank capacity and potential for potable water savings by using rainwater in the residential sector of southeastern Brazil. Build Environ 42(4):1654–1666. https://doi.org/10.1016/j.buildenv.2006.02.007

Gu Q, Zhang Y, Ma L, Li J, Wang K, Wang K, Zheng K, Zhang X, Sheng L (2016) Assessment of reservoir water quality using multivariate statistical techniques : a case study. Sustainability 8(243):1–17. https://doi.org/10.3390/su8030243

Halim MA, Majumder RK, Nessa SA, Oda K, Hiroshiro Y, Jinno K (2010) Arsenic in shallow aquifer in the eastern region of Bangladesh: insights from principal component analysis of groundwater compositions. Environ Monit Assess 161:453–472. https://doi.org/10.1007/s10661-009-0760-9

Harasim P, Filipek T (2015) Nickel in the environment. J Elem 20(2):525–534. https://doi.org/10.5601/jelem.2014.19.3.651

Helena B, Pardo R, Vega M, Barrado E, Fernandez JM, Fernandez L (2000) Temporal evolution of groundwater composition in an alluvial aquifer (Pisuerga River, Spain) by principal component analysis. Water Res 34(3):807–816. https://doi.org/10.1016/S0043-1354(99)00225-0

Helmreich B, Horn H (2009) Opportunities in rainwater harvesting. Desalination 248(1–3):118–124. https://doi.org/10.1016/j.desal.2008.05.046

Henriques AC, De Marco P (2015) Methanesulfonate (MSA) catabolic genes from marine and estuarine bacteria. PLoS ONE 10(5):1–15. https://doi.org/10.1371/journal.pone.0125735

Hodshire AL, Campuzano-Jost P, Kodros JK, Croft B, Nault BA, Schroder JC, Jimenez JL, Pierce JR (2019) The potential role of methanesulfonic acid (MSA) in aerosol formation and growth and the associated radiative forcings. Atmos Chem Phys 19(5):3137–3160. https://doi.org/10.5194/acp-19-3137-2019

Ianniello A, Spataro F, Esposito G, Allegrini I, Hu M, Zhu T (2011) Chemical characteristics of inorganic ammonium salts in PM2.5 in the atmosphere of Beijing (China). Atmos Chem Phys 11(21):10803–10822. https://doi.org/10.5194/acp-11-10803-2011

Igboekwe MU, Akankpo AO, Udoinyang IE (2011) Hydrochemical evaluation of groundwater quality in Michael Okpara University of Agriculture, Umudike and its environs, Southeastern Nigeria. J Water Res Protect 03(12):925–929. https://doi.org/10.4236/jwarp.2011.312103

Ismail A, Toriman MEME, Juahir H, Kassim AMAM, Zain SMSM, Ahmad WKWWKW, Wong KFKF, Retnam A, Zali MAMA, Mokhtar M, Yusri MAMA (2016) Chemometric techniques in oil classification from oil spill fingerprinting. Mar Pollut Bull 111(1–2):339–346. https://doi.org/10.1016/j.marpolbul.2016.06.089

Ismail A, Mohamed SB, Juahir H, Toriman ME, Kassim A, Zain SM, Ahmad WKW, Fah WK, Retnam A, Mokhtar M, Zali MA, Taib MZM, Yang C (2018) DMAIC six sigma methodology in petroleum hydrocarbon oil classification. Int J Eng Technol 7:98–106. https://doi.org/10.14419/ijet.v7i3.14.16868

Jasrotia AS, Majhi A, Singh S (2009) Water balance approach for rainwater harvesting using remote sensing and GIS techniques, Jammu Himalaya, India. Water Resour Manag 23(14):3035–3055. https://doi.org/10.1007/s11269-009-9422-5

Juahir H, Zain SM, Khan R a, Yusoff MK, Mokhtar MB, Toriman ME (2009) Using chemometrics in assessing langat river water quality and designing a cost-effective water sampling strategy. Maejo Int J Sci Technol 3(1):26–42

Juahir H, Zali MA, Retnam A (2011) Spatial characterization of water quality using principal component analysis approach at Juru river basin, Malaysia. World Appl Sci J 14(SPL ISS 1):55–59

Juahir H, Ismail A, Mohamad M, Hamzah FM, Sudianto S, Lasim MLM, Shahriz MA (2019) The assessment of Danau Kota Lake water quality using chemometrics approach. IOP Conf Ser: Mater Sci Eng 621:012019

Kabadayi F, Cesur H (2010) Determination of Cu, Pb, Zn, Ni, Co, Cd, and Mn in road dusts of Samsun City. Environ Monit Assess 168(1–4):241–253. https://doi.org/10.1007/s10661-009-1108-1

Kamarudin MKA, Nalado AM, Toriman ME, Juahir H, Umar R, Ismail A, Abd Wahab N, Md Saad MH, Maulud KN, Hanafiah MM, Mohd Saudi AS, Harith H (2019) Evolution of river geomorphology to water quality impact using remotesensing and GIS technique. Desalin Water Treat 149:258–273. https://doi.org/10.5004/dwt.2019.23838

Kamarudin MKA, Abd Wahab N, Abu Samah MA, Saudi ASM, Ismail A, Toriman ME, Hamzah FM, Saad MHM, Hoe LI, Bati LIHSNAM (2020) Assessing of water quality and sedimentation problems in Lata Sungai Limau, Malaysia. Desalin Water Treat 187(October):1–10. https://doi.org/10.5004/dwt.2020.25269

Kannel PR, Lee S, Kanel SR, Khan SP (2007) Chemometric application in classification and assessment of monitoring locations of an urban river system. Anal Chim Acta 582:390–399. https://doi.org/10.1016/j.aca.2006.09.006

Kazi TG, Arain MB, Jamali MK, Jalbani N, Afridi HI, Sarfraz R a, Baig J a, Shah AQ (2009) Assessment of water quality of polluted lake using multivariate statistical techniques: a case study. Ecotoxicol Environ Saf 72(2):301–309. https://doi.org/10.1016/j.ecoenv.2008.02.024

Kerminen V, Ojanen C, Pakkanen T, Hillamo R, Aurela M, Merila J (2000) Low-molecular-weight dicarboxylic acids in an urban and rural atmosphere. J Aerosol Sci 31(3):349–362

Khairuddin Z, Juahir H, Endut A, Ismail A, Kamarudin MKA, Ghazali A, Mokhtar M, Toriman ME, Islam MS, Lananan F, Hassan AR (2018) Moringa Oleifera dosage clustering for remediation process of Batik effluents using chemometric technique. Int J Eng Technol (UAE) 7(3.14 Special Issue 14):85–89. https://doi.org/10.14419/ijet.v7i3.14.16867

Kowalkowski T, Zbytniewski R, Szpejna J, Buszewski B (2006) Application of chemometrics in river water classification. Water Res 40(4):744–752. https://doi.org/10.1016/j.watres.2005.11.042

Kulshrestha UC, Kulshrestha MJ, Sekar R, Sastry GSR, Vairamani M (2003) Chemical characteristics of rainwater at an urban site of south-central India. Atmos Environ 37(21):3019–3026. https://doi.org/10.1016/S1352-2310(03)00266-8

Lambrakis N, Antonakos A, Panagopoulos G (2004) The use of multicomponent statistical analysis in hydrogeological environmental research. Water Res 38(7):1862–1872. https://doi.org/10.1016/j.watres.2004.01.009

Lani NHM, Yusop Z, Syafiuddin A (2018) A review of rainwater harvesting in Malaysia: prospects and challenges. Water 10(4):1–21. https://doi.org/10.3390/w10040506

Lee JY, Yang JS, Han M, Choi J (2010) Comparison of the microbiological and chemical characterization of harvested rainwater and reservoir water as alternative water resources. Sci Total Environ 408(4):896–905. https://doi.org/10.1016/j.scitotenv.2009.11.001

Lee KE, Mokhtar M, Mohd Hanafiah M, Abdul Halim A, Badusah J (2016) Rainwater harvesting as an alternative water resource in Malaysia: potential, policies and development. J Clean Prod 126:218–222. https://doi.org/10.1016/j.jclepro.2016.03.060

Liaw CH, Chiang YC (2014) Framework for assessing the rainwater harvesting potential of residential buildings at a national level as an alternative water resource for domestic water supply in Taiwan. Water (Switzerland) 6(10):3224–3246. https://doi.org/10.3390/w6103224

Liu C, Lin K, Kuo Y (2003) Application of factor analysis in the assessment of groundwater quality in a blackfoot disease area in Taiwan. Sci Total Environ 313:77–89. https://doi.org/10.1016/S0048-9697(02)00683-6

Mallin MA, Cahoon LB (2003) Industrialized animal production - a major source of nutrient and microbial pollution to aquatic ecosystems. Popul Environ 24(5):369–386. https://doi.org/10.1023/A:1023690824045

Man S, Hashim NM, Ahmad AH, Khin MT, Sidek NS (2014) Kebolehupayaan sistem penuaian hujan sebagai bekalan air alternatif di Malaysia: Suatu penelitian awal. Geografia: Malaysian J Soc Space 10(6):97–104

McKenna JE (2003) An enhanced cluster analysis program with bootstrap significance testing for ecological community analysis. Environ Model Softw 18(3):205–220. https://doi.org/10.1016/S1364-8152(02)00094-4

Menció A, Mas-Pla J (2008) Assessment by multivariate analysis of groundwater-surface water interactions in urbanized Mediterranean streams. J Hydrol 352:355–366. https://doi.org/10.1016/j.jhydrol.2008.01.014

Meng W, Zhong Q, Yun X, Zhu X, Huang T, Shen H, Chen Y, Chen H, Zhou F, Liu J, Wang X, Zeng EY, Tao S (2017) Improvement of a global high-resolution ammonia emission inventory for combustion and industrial sources with new data from the residential and transportation sectors. Environ Sci Technol 51(5):2821–2829. https://doi.org/10.1021/acs.est.6b03694

Miranda R, Tomaz E (2008) Characterization of urban aerosol in Campinas, São Paulo, Brazil. Atmos Res 87(2):147–157. https://doi.org/10.1016/j.atmosres.2007.08.002

Mohd Nasir MF, Saiful Samsudin M, Mohamad I, Amir Awaluddin MR, Ariffin Mansor M, Juahir H, Ramli N (2011) River water quality modeling using combined principle component analysis (PCA) and multiple linear regressions (MLR): A case study at Klang River, Malaysia. World Appl Sci J 14(SPL ISS 1):73–82

Monjerezi M, Vogt RD, Aagaard P, Saka JDK (2011) Hydrogeochemical processes in an area with saline groundwater in lower Shire River valley, Malawi: an integrated application of hierarchical cluster and principal component analyses. Appl Geochem 26:1399–1413. https://doi.org/10.1016/j.apgeochem.2011.05.013

Moreno T, Pandolfi M, Querol X, Lavín J, Alastuey A, Viana M, Gibbons W (2011) Manganese in the urban atmosphere: identifying anomalous concentrations and sources. Environ Sci Pollut Res 18(2):173–183. https://doi.org/10.1007/s11356-010-0353-8

Munim N, Talib M, Othman M, Dominick D, Mohamad N, Juahir H, Tahir NM (2014) Composition of selected heavy metals in road dust from Kuala Lumpur city centre. Environ Earth Sci 72:849–859. https://doi.org/10.1007/s12665-013-3008-5

Mustapha A, Aris AZ, Ramli MF, Juahir H (2012) Spatial-temporal variation of surface water quality in the downstream region of the Jakara River, north-western Nigeria: a statistical approach. J Environ Sci Health - A Toxic/Hazard Subst Environ Eng 47(11):1551–1560. https://doi.org/10.1080/10934529.2012.680415

Mustapha A, Aris AZ, Juahir H, Ramli MF (2013) River water quality assessment using environmentric techniques: case study of Jakara River Basin. Environ Sci Pollut Res 20:5630–5644. https://doi.org/10.1007/s11356-013-1542-z

Mutihac L, Mutihac R (2008) Mining in chemometrics. Anal Chim Acta 612:1–18. https://doi.org/10.1016/j.aca.2008.02.025

Narany TS, Ramli MF, Aris AZ, Nor W, Sulaiman A, Juahir H, Fakharian K, Sheikhy Narany T, Ramli MF, Aris AZ, Sulaiman WNA, Juahir H, Fakharian K, Narany TS, Ramli MF, Aris AZ, Nor W, Sulaiman A, Juahir H, Fakharian K (2014) Identification of the hydrogeochemical processes in groundwater using classic integrated geochemical methods and geostatistical techniques , in Amol-Babol Plain, Iran. Sci World J 2014:1–15. https://doi.org/10.1155/2014/419058

Niu H, He Y, Lu XX, Shen J, Du J, Zhang T, Pu T, Xin H, Chang L (2014) Chemical composition of rainwater in the Yulong Snow Mountain region, Southwestern China. Atmos Res 144:195–206. https://doi.org/10.1016/j.atmosres.2014.03.010

Norsukhairin, S., Abdul, S., Azid, A., Sharif, S. M., Latif, T., Aris, Z., Zain, M., Dominick, D., Mutalib, S. N. S. A., Juahir, H., Azid, A., Mohd Sharif, S., Latif, M. T., Aris, A. Z., Zain, S. M., Dominick, D., Norsukhairin, S., Abdul, S., Azid, A., … Dominick, D. (2013). Spatial and temporal air quality pattern recognition using environmetric techniques: a case study in Malaysia. Environ Sci Process Impacts 15:1717–1728. doi:https://doi.org/10.1039/c3em00161j

Ogwueleka TC (2015) Use of multivariate statistical techniques for the evaluation of temporal and spatial variations in water quality of the Kaduna River, Nigeria. Environ Monit Assess 187(3):137. https://doi.org/10.1007/s10661-015-4354-4

Pachon JE, Weber RJ, Zhang X, Mulholland JA, Russell AG (2013) Revising the use of potassium (K) in the source apportionment of PM 2.5. Atmos Pollut Res 4(1):14–21. https://doi.org/10.5094/APR.2013.002

Panda UC, Sundaray SK, Rath P, Nayak BB, Bhatta D (2006) Application of factor and cluster analysis for characterization of river and estuarine water systems - a case study: Mahanadi River (India). J Hydrol 331(3–4):434–445. https://doi.org/10.1016/j.jhydrol.2006.05.029

Paoli L, Pirintsos SA, Kotzabasis K, Pisani T, Navakoudis E, Loppi S (2010) Effects of ammonia from livestock farming on lichen photosynthesis. Environ Pollut 158(6):2258–2265. https://doi.org/10.1016/j.envpol.2010.02.008

Reimann C, Filzmoser P, Garrett RG (2002) Factor analysis applied to regional geochemical data: problems and possibilities. Appl Geochem 17(3):185–206. https://doi.org/10.1016/S0883-2927(01)00066-X

Rose RAC, Samsuddin NM, Rosmiz MZ, Lyndon N (2016) Implikasi Perubahan Pembangunan Bandar Terhadap Keselamatan Awam: Kajian Kes di Bandar Muadzam Shah, Pahang. J Soc Sci Human 1(1):107–120

Rowe MP (2011) Rain water harvesting in Bermuda. J Am Water Resour Assoc 47(6):1219–1227. https://doi.org/10.1111/j.1752-1688.2011.00563.x

Samat N (2007) Integrating Gis and cellular automata spatial model in evaluating urban growth : prospects and challenges. Jurnal Alam Bina 9:15

Santos C, Taveira-Pinto F (2013) Analysis of different criteria to size rainwater storage tanks using detailed methods. Resour Conserv Recycl 71:1–6. https://doi.org/10.1016/j.resconrec.2012.11.004

Sârbu C, Pop HF (2005) Principal component analysis versus fuzzy principal component analysis: a case study: the quality of danube water (1985-1996). Talanta 65(5):1215–1220. https://doi.org/10.1016/j.talanta.2004.08.047

Sazakli E, Alexopoulos A, Leotsinidis M (2007) Rainwater harvesting, quality assessment and utilization in Kefalonia Island, Greece. Water Res 41(9):2039–2047. https://doi.org/10.1016/j.watres.2007.01.037

Shamsuddin A, Ishfaq M, Ruslan M, Halim AA, Zahari NF, Farhana N, Fazi M (2014) Educators' awareness and acceptance towards goods and services tax (GST) implementation in Malaysia: a study in Bandar Muadzam Shah, Pahang. Int J Bus, Econ Law 4(1):1–10

Shamsudin SA (2013) Merungkai Beberapa Faktor Jangkitan HIV / AIDS dalam Kalangan Penagih Dadah Mengikut Perspektif Sosiobudaya *6*:49–63

Shrestha S, Kazama F (2007) Assessment of surface water quality using multivariate statistical techniques: A case study of the Fuji river basin, Japan. Environ Model Softw 22:464–475. https://doi.org/10.1016/j.envsoft.2006.02.001

Simmons G, Hope V, Lewis G, Whitmore J, Gao W (2001) Contamination of potable roof-collected rainwater in Auckland, New Zealand. Water Res 35(6):1518–1524. https://doi.org/10.1016/S0043-1354(00)00420-6

Suhaila J, Jemain AA (2007) Fitting daily rainfall amount in Malaysia using the normal transform distribution. J Appl Sci 7(14):1880–1886. https://doi.org/10.3923/jas.2007.1880.1886

Suhaila J, Deni SM, Zawiah Zin WAN, Jemain AA (2010) Trends in Peninsular Malaysia rainfall data during the southwest monsoon and northeast monsoon seasons: 1975-2004. Sains Malaysiana 39(4):533–542

Sunardi S, Ariyani M, Agustian M, Withaningsih S, Parikesit P, Juahir H, Ismail A, Abdoellah OS (2020) Water corrosivity of polluted reservoir and hydropower sustainability. Sci Rep 10(1):1–8. https://doi.org/10.1038/s41598-020-68026-x

Tahir NAM, Juahir H, Zakaria MP, Zain SM, Retnam A, Sing TS, Manan N, Toriman ME, Ismail A, Kamarudin MKA, Ghazali A, Wahab NA (2018) Environmental forensic techniques in identifying the sources of 16 USEPA priority pollutants polycyclic aromatic hydrocarbons (PAHs) at the Langat River, Selangor, Malaysia. Int J Eng Technol (UAE) 7(3.14 Special Issue 14):123–131. https://doi.org/10.14419/ijet.v7i3.14.16872

Tian HZ, Lu L, Cheng K, Hao JM, Zhao D, Wang Y, Jia WX, Qiu PP (2012) Anthropogenic atmospheric nickel emissions and its distribution characteristics in China. Sci Total Environ 417–418: 148–157. https://doi.org/10.1016/j.scitotenv.2011.11.069

Toriman ME, Abdullah SNF, Azizan IA, Kamarudin MKA, Umar R, Mohamad N (2015) Spatial and temporal assessment on drug addiction using multivariate analysis and GIS. Malaysian J Anal Sci 19(6):1361–1373

Van Metre PC, Mahler BJ (2003) The contribution of particles washed from rooftops to contaminant loading to urban streams. Chemosphere 52(10):1727–1741. https://doi.org/10.1016/S0045-6535(03)00454-5

Varol M (2013) Dissolved heavy metal concentrations of the kralkizi, dicle and batman dam reservoirs in the tigris river basin, turkey. Chemosphere 93(6):954–962. https://doi.org/10.1016/j.chemosphere.2013.05.061

Vaselli O, Buccianti A, De Siena C, Bini C, Coradossi N, Angelone M (1997) Geochemical characterization of ophiolitic soils in a temperate climate: a multivariate statistical approach. Geoderma 75:117–133. https://doi.org/10.1016/S0016-7061(96)00084-5

Walna B, Kurzyca I, Siepak J (2007) Variations in the fluoride level in precipitation in a region of human impact. Water, Air, Soil Pollut: Focus 7(1–3):33–40. https://doi.org/10.1007/s11267-006-9108-4

Walna B, Kurzyca I, Bednorz E, Kolendowicz L (2013) Fluoride pollution of atmospheric precipitation and its relationship with air circulation and weather patterns (Wielkopolski National Park, Poland). Environ Monit Assess 185(7):5497–5514. https://doi.org/10.1007/s10661-012-2962-9

Wan Zin WZ, Jamaludin S, Deni SM, Jemain AA (2010) Recent changes in extreme rainfall events in Peninsular Malaysia: 1971-2005. Theor Appl Climatol 99(3–4):303–314. https://doi.org/10.1007/s00704-009-0141-x

Wang X, Sato T, Xing B (2006a) Size distribution and anthropogenic sources apportionment of airborne trace metals in Kanazawa, Japan. Chemosphere 65(11):2440–2448. https://doi.org/10.1016/j.chemosphere.2006.04.050

Wang Y, Zhuang G, Zhang X, Huang K, Xu C, Tang A, Chen J, An Z (2006b) The ion chemistry, seasonal cycle, and sources of PM 2.5 and TSP aerosol in Shanghai. Atmos Environ 40(16):2935–2952. https://doi.org/10.1016/j.atmosenv.2005.12.051

Wu M-L, Wang Y-S, Sun C-C, Wang H, Dong J-D, Yin J-P, Han S-H (2010) Identification of coastal water quality by statistical analysis methods in Daya Bay, South China Sea. Mar Pollut Bull 60(6):852–860. https://doi.org/10.1016/j.marpolbul.2010.01.007

Yousif M, Bubenzer O (2015) Geoinformatics application for assessing the potential of rainwater harvesting in arid regions. Case study: El Daba'a area, Northwestern Coast of Egypt. Arab J Geosci 8(11): 9169–9191. https://doi.org/10.1007/s12517-015-1837-0

Zhang FS, Yamasaki S, Kimura K (2002) Waste ashes for use in agricultural production: II. Contents of minor and trace metals. Sci Total Environ 286(1–3):111–118. https://doi.org/10.1016/S0048-9697(01)00968-8

Zhang T, Cao JJ, Tie XX, Shen ZX, Liu SX, Ding H, Han YM, Wang GH, Ho KF, Qiang J, Li WT (2011) Water-soluble ions in atmospheric aerosols measured in Xi'an, China: seasonal variations and sources. Atmos Res 102(1–2):110–119. https://doi.org/10.1016/j.atmosres.2011.06.014

Zhao L, Wong WB, Hanafi ZB (2019a) The evolution of George Town's urban morphology in the Straits of Malacca, late 18th century-early 21st century. Front Architect Res 8:1–22. https://doi.org/10.1016/j.foar.2019.09.001

Zhao Y, Yuan M, Huang X, Chen F, Zhang J (2019b) Quantification and evaluation of atmospheric pollutant emissions from open biomass burning with multiple methods: a case study for the Yangtze River Delta region, China. Atmos Chem Phys 19(1):327–348. https://doi.org/10.5194/acp-19-327-2019

Zhou F, Guo H, Liu Y, Jiang Y (2007) Chemometrics data analysis of marine water quality and source identification in Southern Hong Kong. Mar Pollut Bull 54(6):745–756. https://doi.org/10.1016/j.marpolbul.2007.01.006

Zolkipli HM, Juahir H, Adiana G, Zainuddin N, Ismail A, Maliki ABHM, Hussain NI, Kamarudin MKA, Toriman ME, Mokhtar M (2018) Spatial assessment and the most significant parameters for drinking water quality using chemometric technique: a case study at Malaysia water treatment plants. Int J Eng Technol (UAE) 7(3.14 Special Issue 14):115–122. https://doi.org/10.14419/ijet.v7i3.14.16871