*International Journal of Engineering & Technology*, 7 (3.14) (2018) 98-106

**International Journal of Engineering & Technology**

*Website: www.sciencepubco.com/index.php/IJET*

***Research paper***



# DMAIC Six Sigma Methodology in Petroleum Hydrocarbon Oil Classification

**Azimah Ismail[1,2], Saiful Bahri Mohamed[1*], Hafizan Juahir[1*], Mohd Ekhwan Toriman[5], Azlina Md. Kassim[3], Sharifuddin Md Zain[4], Wan Kamaruzaman Wan Ahmad[3], Wong Kok Fah[3], Ananthy Retnam[3], Mazlin Mokhtar[5], Munirah Abdul Zali[3], Mohd. Zaki Mohd. Taib[6], Chun Yang[7]**

*[1]East Coast Environmental Research Institute (ESERI), Universiti Sultan Zainal Abidin, Gong Badak Campus, 21300 Kuala Terengganu, Malaysia*
*[2]Faculty of Design Innovative and Technology, Universiti Sultan Zainal Abidin, Gong Badak Campus, 21300 Kuala Terengganu, Malaysia*
*[3]Environmental Health Division, Department of Chemistry Malaysia, Ministry of Science, Technology and Innovation, Petaling Jaya, Selangor, Malaysia*
*[4]Chemistry Department, University of Malaya, 50603 Kuala Lumpur, Malaysia*
*[5]Faculty of Social Sciences and Humanities, National University of Malaysia, Bangi, Selangor, Malaysia*
*[6]Integrated Envirotech Sdn. Bhd.*
*[7]Environment Change and Climate Change Canada, 335 River Road, Ottawa, Ontario, Canada*
**Corresponding authors E-mail: saifulbahri@unisza.edu.my, hafizanjuahir@unisza.edu.my*

**Abstract**

This research focuses on the use of the DMAIC method (Define, Measure, Analyze, Improve and Control) as a Six Sigma approach in studying oil spill fingerprint of samples recovered from Peninsular Malaysia and Sabah (East Malaysia). The DMAIC approach in this study was used as a way to classify oil types based on data obtained from GC-FID and GC-MS measurements. The cause-effect diagram was used to define the factors leading to the failure of the oil spill fingerprinting based on inaccurate oil type clustering. Discriminant Analysis (DA) was also applied to quantify the root-cause of the failure. An Ishikawa diagram obtained in the analysis phase identifies the potential failure causal. Principal component analysis (PCA) was applied and was successful in discriminating four clusters of oil types, namely diesel, heavy fuel oil (HFO), mixture oil lube and fuel oil (MOLFO) and waste oil (WO) with a total variance of 85.3%. In the control phase, the use of a Pareto chart indicated 100% cumulative percentage of oil type clustering with a 95% confidence level. The DMAIC approach to be effective in solving oil spill fingerprinting problems and results in quality improvement in the clustering of oil spills into the different hydrocarbon types.

***Keywords***: *Cause effect-diagram; chemometrics; DMAIC; oil spill classification; oil spill fingerprint; Six Sigma.*

## 1. Introduction

Six Sigma is commonly known as a methodology to improve steps in many business activities. It has evolved in practice where statistical tools derived from the Deming's PDCA cycle and the seven steps of Juran's function to improve quality in order to achieve defect reduction [1]. The application of DMAIC in six sigma for oil spill fingerprinting, particularly in problem analysis has been impressive but the methodology serves more than that and this can be seen in the improvement of the processes in dealing with a complex mixture of hydrocarbon with a high level of uncertainty. This requires accurate statistics and validation analysis of a broader range of hydrocarbon compounds and on this, DMAIC is clearly useful. It is interesting to note that many oil spill fingerprinting projects are initially based on conventional systems, but clearly, improvement of the projects is desired where structured improvement methods of the DMAIC concept may be used to achieve organizational and market objectivity [2]. The methodologies and implementation concept of DMAIC Six Sigma framework in oil fingerprinting is proven to provide value added elements in the analysis and validation of oil spill samples.

The application of DMAIC Six Sigma in Information Technology systems have improved the flexibility in supply chain related issues [3]. It has been proven that the application of DMAIC Six Sigma in the supply chain performance system has improved the efficiency of delivering finished product with the availability of up-to-date data and appropriate information [4]. DMAIC Six Sigma has also benefitted supply chain objectives, in terms of slate delivery, reaction time and preventive measures through the information imparted based on real-time forecasting and prerequisite data [5]. In many organizations, logistic cost impacts the annual fiscal significantly prior to the establishment of DMAIC Six Sigma, especially in transportation [6]. Clearly, continuous improvement is always sought by many researchers for accurate analysis methodologies and statistical approach such as DMAIC Six Sigma for oil spill fingerprinting, particularly in problem analysis has been impressive but the methodology serves more than that and this can be seen in the improvement of the processes in dealing with a complex mixture of hydrocarbon with a high level of uncertainty. This requires accurate statistics and validation analysis of a

**Copyright © 2018 Authors. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.**

broader range of hydrocarbon compounds and on this, DMAIC is clearly useful. It is interesting to note that many oil spill fingerprinting projects are initially based on conventional systems, but clearly, improvement of the projects is desired where structured improvement methods of the DMAIC concept may be used to achieve organizational and market objectivity [2]. The methodologies and implementation concept of DMAIC Six Sigma framework in oil fingerprinting is proven to provide value added elements in the analysis and validation of oil spill samples.

The application of DMAIC Six Sigma in Information Technology systems have improved the flexibility in supply chain related issues [3]. It has been proven that the application of DMAIC Six Sigma in the supply chain performance system has improved the efficiency of delivering finished product with the availability of up-to-date data and appropriate information [4]. DMAIC Six Sigma has also benefitted supply chain objectives, in terms of slate delivery, reaction time and preventive measures through the information imparted based on real-time forecasting and prerequisite data [5]. In many organizations, logistic cost impacts the annual fiscal significantly prior to the establishment of DMAIC Six Sigma, especially in transportation [6]. Clearly, continuous improvement is always sought by many researchers for accurate analysis methodologies and statistical approach such as DMAIC Six Sigma

Oil spill fingerprinting requires a reliable methodology of sampling, analysis (laboratory and statistical) and validation in order to achieve the objectives of environmental forensics. DMAIC Six Sigma approach in oil spill fingerprinting is clearly required in order to achieve quality improvement and substantial elimination of errors. The initiatives on Lean Six Sigma approach have positively affected many organizational performances [8]. DMAIC Six Sigma has brought relevant improvements in manufacturing system of automotive component in Connecticut [9]. The approach has also contributed to significant development and defect reduction in many industries [10]. The adoption of Six Sigma in many companies results in significant improvements in the production, delivery and management processes [11-12]. Many applications of Six Sigma are focused on improving organizational performance, however, in oil spill fingerprinting, the application of Six Sigma protocols from sampling, up to the interpretation of the results is a new endeavor data [13].

DMAIC Six Sigma is a data-driven improvement cycle used for improving, optimizing and stabilizing business processes and designs. In oil spill fingerprinting, DMAIC can be viewed as a problem-solving approach in data mining and exploratory data analyses which involves applying the steps defined by the framework to data obtained from chromatographic measurements (obtained from gas chromatography flame ionization detector (GC-FID) and gas chromatography mass spectrometry (GC-MS)) to enable an improved classification of hydrocarbon compounds from collected oil spill samples. In this study, we endeavor to apply the DMAIC Six Sigma framework as a problem-solving tool in order to improve the quality of oil type classification from the oil spills recovered from Peninsular Malaysia and East Malaysia (Sabah).

## 2. Methodology

### 2.1. Samples Collection

Spilled oil samples were collected from thirty-two distinct locations in the Peninsular of Malaysia and Sabah (East Malaysia). All the samples were collected by the Department of Environment (DOE), within a one-year period from 2013 until 2014 (Figure 1). The descriptions and the sampling locations are given in Table 1. The samples were immediately registered upon arrival after transporting to the laboratory and kept at 4$^{o}$ C prior to analysis.

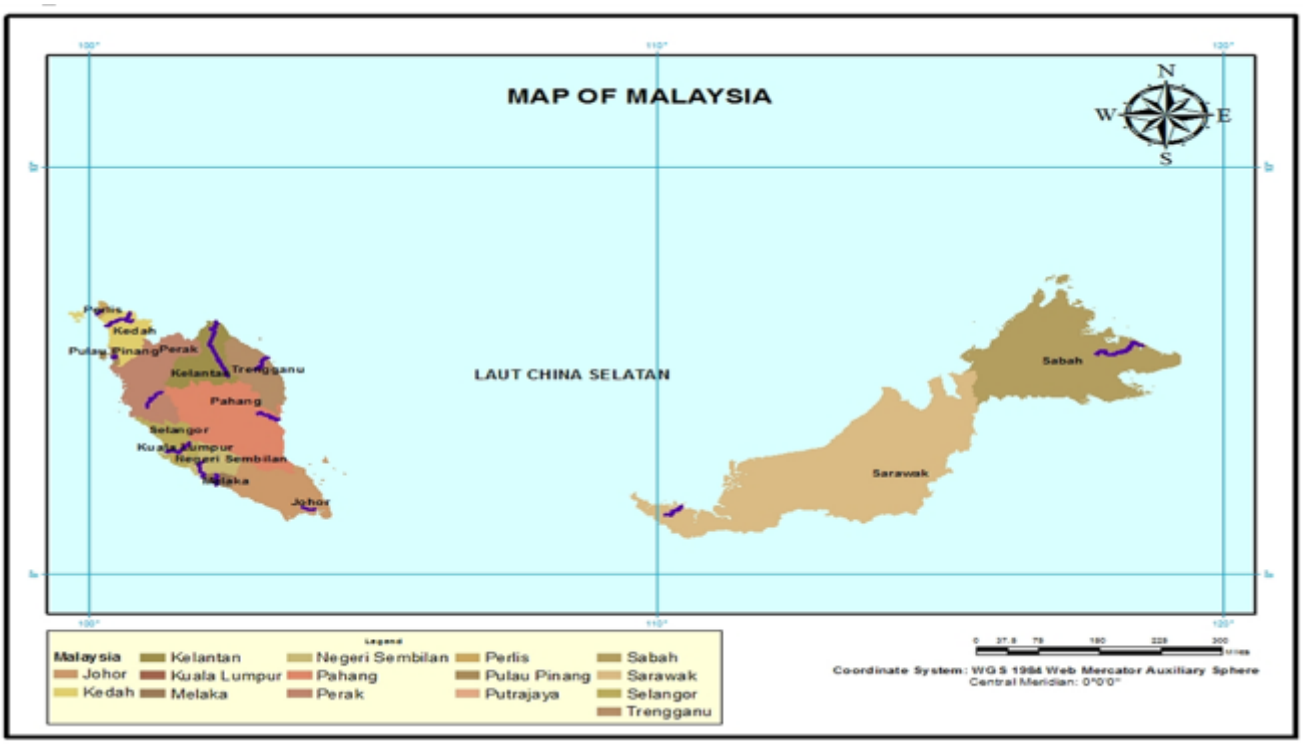



**Fig. 1:** Map showing the oil spill sampling locations within the Peninsular of Malaysia and Sabah (East Malaysia)

### 2.2. Solvents and Sample Preparation for Laboratory Analysis

The chemicals used in the laboratory analysis were Dichloromethane (DCM) and hexane. Both are solvents commonly used in spilled oil classification. All the internal reference standards for polycyclic aromatic hydrocarbon (PAHs), biomarker and *n*-alkane were obtained from Chiron. These include terphenyl-$d_{14}$; $C_{30}17\ \beta$ (H), 21β(H)-hopane and 5α-androstane. Surrogate standards used for this research (phenanthrene-$d_{10}$, perylene-$d_{12}$, acenapthene-$d_{10}$, benz (a) anthracene-$d_{12}$ and o-terphenyl) were procured from AccuStandard Inc.

**Table 1:** Locations of spilled oil samples recovered by the Department of Environment, Malaysia (DOE)

| No. | Samples Code | Samples Description | Sampling Location | Case of Spilled oil |
|---|---|---|---|---|
| 1 | BAS303A | Blackish, spilled (oil + water) | Selangor | Source from a premise drain outlet |
| 2 | BAS304B | Blackish, source (oil + water) | Selangor | Waste oil from public drain |
| 3 | BAS497A | Blackish, spilled (oil + water) | Selangor | Source from premise drain |
| 4 | BAS497B | Blackish, spilled (oil + water) | Selangor | Source from premise drain |
| 5 | BAS497C | Blackish, source oil | Selangor | Source from inside premise |
| 6 | BAS498B | Blackish, spilled (oil + water) | Selangor | Source from river |
| 7 | BAS498C | Blackish, source oil | Selangor | Source from inside premise |
| 8 | BAS498D | Blackish, source (oil + water) | Selangor | Source from inside premise |
| 9 | BAS498E | Blackish, source (oil + water) | Selangor | Source from inside premise |
| 10 | BAS498F | Blackish, source (oil + water) | Selangor | Source from inside premise |
| 11 | BAS741A | Yellowish, source oil | Selangor | Source from lorry tanker |
| 12 | BAS741B | Yellowish, spilled oil | Selangor | Source from public drain |
| 13 | BAS1092-1 | Blackish, oil | Selangor | |
| 14 | BAS1092-2 | Blackish, oil | Selangor | |
| 15 | BAS1092A | Blackish, oil | Selangor | |
| 16 | BAS1092B | Blackish, oil | Selangor | |
| 17 | BAS1383E1 | Blackish, oil | Johor | Sample from Johor Straits (illegal oil transferring from vessels) |
| 18 | BAS1383E2 | Blackish, oil | Johor | Sample from Johor Straits (illegal oil transferring from vessels) |
| 19 | AAS106A | Blackish, source oil | Perak | Waste oil from lorry tanker |
| 20 | AAS106B | Blackish, spilled (oil + water) | Perak | Waste oil collected from the drain |
| 21 | BAS156A | Blackish, source oil | Selangor | Sample from factory drum container |
| 22 | BAS156B | Blackish, spilled (oil + water) | Selangor | Source from public drain and river |
| 23 | BAS156C | Blackish, spilled (oil + water) | Selangor | Source from public drain and river |
| 24 | JAS294A | Yellowish, oil | Johor | Sample from Johor Straits (illegal oil transferring from vessels) |
| 25 | JAS294B | Blackish, oil | Johor | Source spilled from Johor Straits (illegal oil transferring from vessels) |
| 26 | JAS294C | Yellowish, oil | Johor | |
| 27 | JAS294D | Blackish, oil | Johor | |
| 28 | MAS356A | Blackish, spilled (oil sludge) | Melaka | Source spilled on earth |
| 29 | MAS356B | Blackish, spilled (oil + water) | Melaka | Source from outlet factory drainage |
| 30 | SAS141A | Blackish, source oil | Sabah | Sample from lorry tanker |
| 31 | SAS141B | Blackish, spilled (oil + water) | Sabah | Sample from public drain |
| 32 | SAS141C | Blackish, spilled (oil + water) | Sabah | Sample from public drain |

## 2.3. Laboratory Analysis

The laboratory analysis carried out in this work involved the fingerprinting process using GC-FID and GC-MS. In this process, molecular features and compositional parameters in petroleum was specifically targeted. Uncontaminated oil and oil-water are the two types of oil samples used in this work. This analysis enabled us to classify the types of oil and ascertain the quantitative value of the samples which can then be statistically analyzed. Appropriate interpretation of the presence of various oil types allow unambiguous identification of the oil sources. The sample preparation process is illustrated in the flow chart in Figure 2.

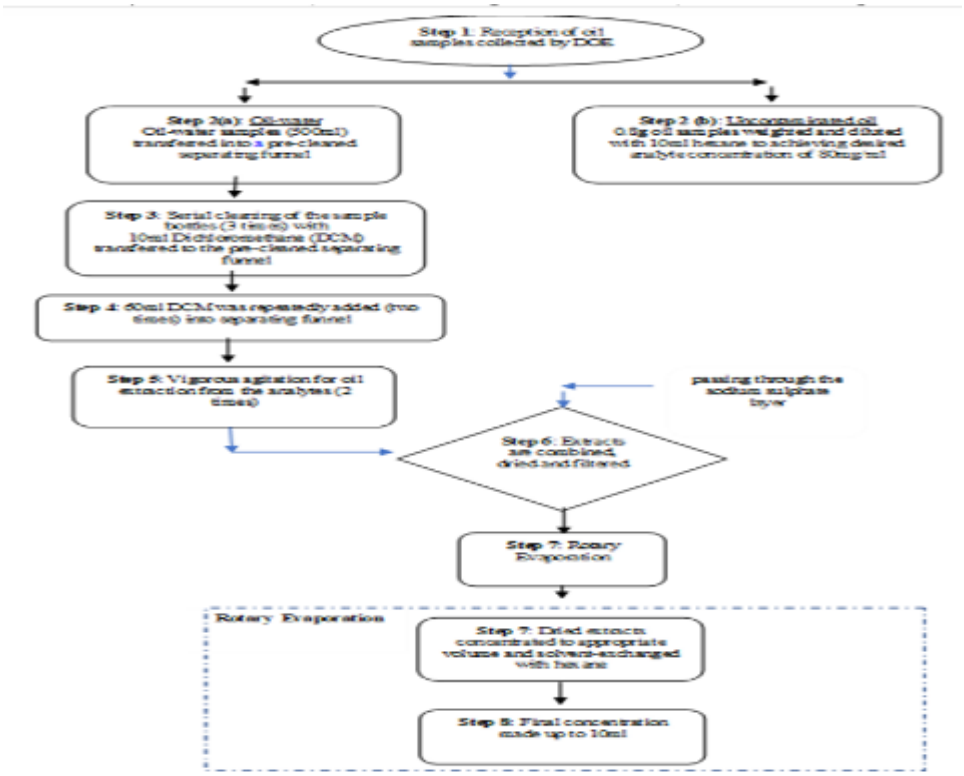



**Fig. 2:** Flow chart of sample preparation for oil-water and uncontaminated sample analysis

### 2.3.1. Column Clean-Up Analysis

Pre-cleaned glass column (30cm x 10.5mm I.D.) was used to perform the column clean-up analysis. The glass column contained approximately 6.0g pre-cleaned silica gel (100-200 mesh, Davisil grade 923) topped with 0.5cm pre-cleaned sodium sulphate with conditioning using of 20ml hexane. Further, the extracted oil was spiked with surrogates (4 mix PAHs compounds and o-terphenyl) and was quantitatively transferred into the column. 3.0ml hexane was quantitatively added to complete the oil transferring process into the column. Prior to exposure of sodium sulphate, 12.0ml hexane standard approximately was added into the column to elute the aliphatic compounds. The aliphatic compounds consist of *n*-alkanes and biomarkers are addressed as F1. Subsequently, approximately 15ml of 50% dichloromethane in hexane used to elute the aromatic compounds addressed as F2 which was spiked with 1ppm of $d_{14}$-terphenyl. Here, the fractions (F1 and F2) were concentrated under a gentle stream of $N_2$ (flow rate of 1.0 ml/s) until the volume reached the final volume of less than 0.5ml. The final volume of injection prior to the analysis were made up, with the combined fractions to 1.0 ml.
Each sample was incorporated with a procedural blank, a matrix spike and a duplicate sample in compliance with the quality control procedure in oil spill fingerprinting. The procedural blank was carried out for all samples using GC-MS. The concentration detection limit for the procedural blank for each analyte was set at 0.03µg/ml and for each analyte or sample the detected concentration was well below the concentration detection limit [14-17].

### 2.3.2. GC-FID and GC-MS Analysis of Oil Spill Fingerprints

The tiered analytical use of GC-FID and GC-MS enable pattern recognition and identification of the types of oil spills. The analyses of n-alkanes and PAHs in highly complex oil samples are performed using the methods reported by [17]. The oil spill analysis was performed for both *n*-alkane distributions and total petroleum hydrocarbons (TPHs) on a Perkin Elmer, Clarus 680 fitted with a flame ionization detection (FID) and PE AutoSystem GC with built-in Autosampler. A column with dimension of 30 m x 0.25 mm ID DB-5 was used to perform the analysis. An Agilent, 7890A GC System equipped with mass-selective detector and CTC PAL ALS Autosampler was used to perform the analysis or identification of target PAH compounds (including five alkylated PAH homologous groups and other EPA priority PAHs) and biomarkers such as terpanes and steranes. The column used was a 30 m x 0.25 mm HP-5MS fused silica column. The approach applied in this study is illustrated in Figure 3. The results from the analysis then underwent statistical analyses to ascertain the types of oil within the samples obtained.

## 2.4. DMAIC Six Sigma Approaches in Oil Spill Classification

DMAIC is the acronym for five phases; define, measure, analyze, improve and control. These are data - driven methodologies to solve problems related to organized process.

**Step** 1 - Define phase
In the define phase of this study, the performance of the oil spill fingerprinting process from sampling en route to the data interpretation is evaluated. This step is necessary in order to improve the oil classification obtained through the oil spill fingerprinting process. The phase states focus on the oil fingerprints, the root-cause leading to failure of oil fingerprinting, analysis constraints and risks. Cause-effect (Fishbone) or Ishikawa diagram is introduced to achieve the solutions [10].

**Step** 2 - Measure Phase
In the measure phase, determination of the appropriate measures to quantify the root-cause for the failure of oil classification from the oil spill fingerprinting is evaluated. Since this phase is a map process of the oil spill fingerprinting process, thus it involves the data collection of oil spill samples and proceeds with the evaluation of the defects along the process using discriminant analysis (DA).
Statistically, DA allows the classification of oil with similar intrinsic chemical properties. Discriminant functions (DFs) perform the validation of raw data sets or set of variables known as predictors. DA enables the determination of the oil spill variables and discriminate them into two or more groups (clusters) [19-26]. Discriminant Factor (DF) of DA is constructed based on the original dataset that have similar discriminant ability of each group towards the original data set with or without standardization [20, 33]. Thus, standardization to the original data set was not necessarily significant and the DA was performed.
Statistically, DA allows the classification of oil with similar intrinsic chemical properties. Discriminant functions (DFs) perform the validation of raw data sets or set of variables known as predictors. DA enables the determination of the oil spill variables and discriminate them into two or more groups (clusters) [19-26].

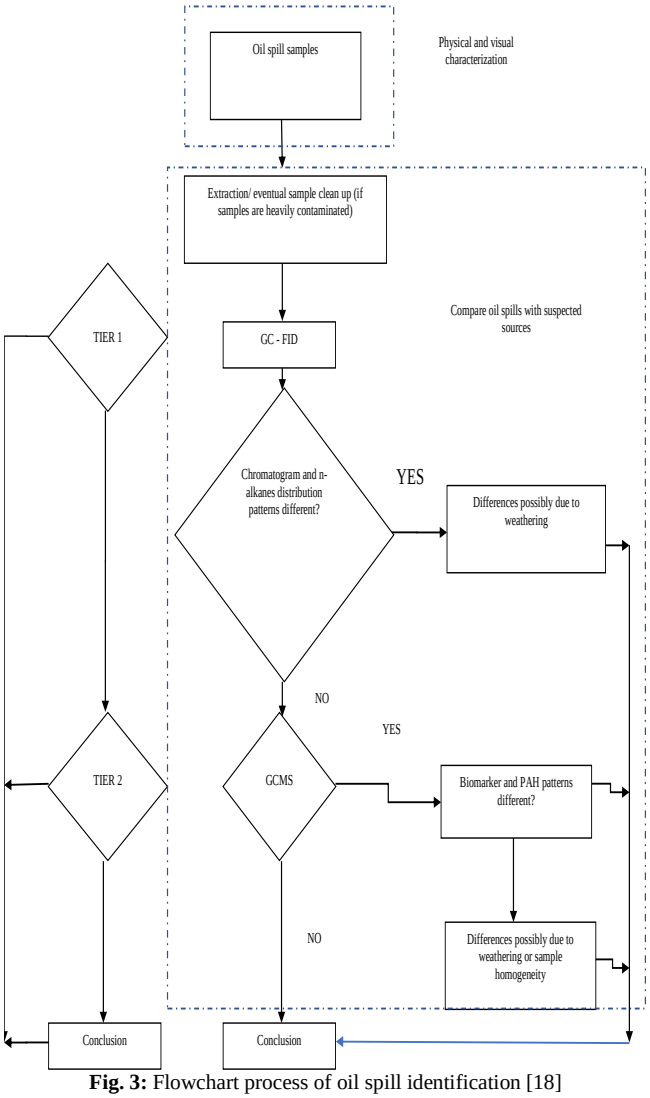



**Fig. 3:** Flowchart process of oil spill identification [18]

Discriminant Factor (DF) of DA is constructed based on the original dataset that have similar discriminant ability of each group towards the original data set with or without standardization [20, 33]. Thus, standardization to the original data set was not necessarily significant and the DA was performed by following function (Equation (1));

$$f(Gi) = ki + \sum_{i=1}^{n} wij \,.pij \quad (1)$$

[20]

Where i is 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 given group and $w_{ij}$ is the weight coefficient which was assigned by DA to a given parameter ($p_{ij}$) [19].
Discriminant Factors (DF) matrix is constructed based on three modes; standard, forward stepwise and backward stepwise. The forward stepwise mode involves a one-by-one addition of the oil spill variables on the matrix until no significant changes are achieved. In contrary, the backward stepwise mode performs the exclusion leave-one-out method of oil spill variables from the least significant one until no changes are observed [10, 20, 22].

**Step** 3 - Analyze Phase of DMAIC
The comprehensive characterization of oil spill fingerprinting, in particular the effectiveness of the analyzing phase of DMAIC Six Sigma is driven to identify the potential causes of the oil spill. Through brainstorming sessions, the root-cause problems that lead addressed via the Cause and Effects Diagram. The potential cause of problems is validated through the cause-effect diagrams.

**Step** 4 - Implement change for improvement
Once the potential causes or cause-root effects have been identified and validated, the process of defect elimination is implemented [10]. Principal Component Analysis (PCA) is a reliable statistical tool in the improvement phase of oil spill fingerprinting. As an unsupervised pattern recognition method, PCA serves to identify the most significant oil spill variables via dimensionality reduction of the dataset. Principal components (PCs) which represent the underlying factors while sustaining the information of the original dataset are employed [20, 34]. PCA/PCR is one of the effective tools for a meaningful data reduction and interpretation and can be effective in detecting correlation between oil spill variables from different pollution sources [22, 25-26]. PCs with eigenvalue > 1.0 are considered significant. Upon varimax rotation, the PCs with factor loadings greater than 0.7 can be considered for further discussion. A PC can be expressed in the form of equation below:

$$z_{ij} = a_{i1}X_{1j} + a_{i2}X_{2j} + a_{i3}X_{3j} + a_{im}X_{mj} \quad (2)$$

where *z* is the component score, *a* is the component loading, X is the measured original variables, *i* is the component number and *m* represents the number of the original variables. The loadings ascertain the significance of the oil spill parameters for each component.

Step 5 - Demonstrate Process Control
In the control phase, the relevant procedures are prepared for implementation and sustained and monitored by those directly involved in the process. Pareto charts are used to demonstrate the process control of the accurately classified oil spills.

# 3. Outcome Analysis and Discussion

## 3.1. Oil Spill Fingerprinting

Oil spill fingerprinting employs a high precision setting on the identification or tracking of the spilled oil candidates in order to achieve desired precision on the oil classification of the unknown sources recovered from various locations. DMAIC Six Sigma methods are applied in this study for improvement and assessment of the common techniques in oil spill-fingerprinting [10]. The assignment of the various oil spill fingerprints was successfully accomplished through pattern matching of mass spectra, retention time data comparison of GC with reference standards, and empirical data from previous research [17, 27]. In general, the source identification of oil spills has been successfully conducted through the application of the tiered methods [10, 17]. Figure 4 shows the GC-FID chromatographic profiling of selected oil samples comprising diesel, heavy fuel oil, lubricating oil, waste oil and waste oil mixed with light fuel oil.

The ninety-four hydrocarbon compounds in this study consists of n-alkanes ($C_{10}$ – $C_{40}$), biomarkers (sesquiterpanes, steranes and terpanes) and PAHs (alkylated PAHs and EPA Priority PAHs). To determine the biomarkers and PAHs the analysis was accomplished through GC-MS in selected ion monitoring (SIM) mode, while the quantification n-alkanes was performed using GC-FID.

## 3.2. Chromatogram Profiling

Chromatogram profiling is important in determining the distribution of *n*-alkanes in the hydrocarbon composition of the spilled oil samples.

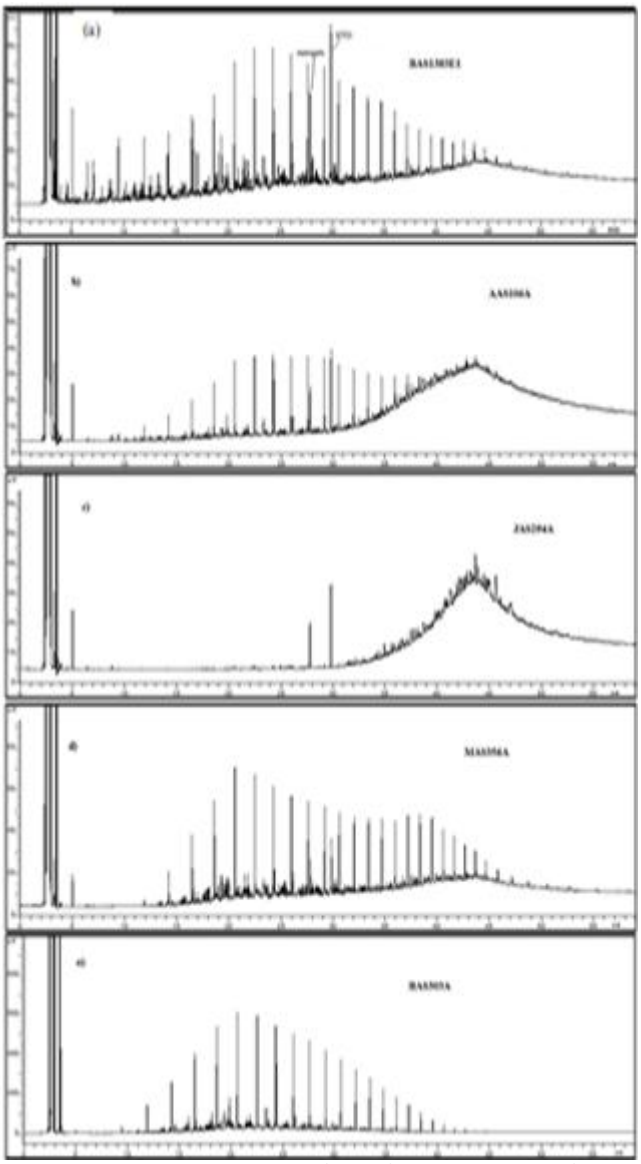


**Fig. 4:** GC-FID Chromatography profiling of the selected types of spilled oil distinguishes by n-alkanes pattern distribution of different oil source a) HFO, b) MOLFO c) lube oil, d) WO and e) Diesel

Each sample extracted from the study area shows a different pattern of chromatogram or hydrocarbon composition due to the different sources of the spilled oil samples. The characteristic of the chromatograms reveals that there are broader unresolved complex mixtures (UCM) as shown in Figures 4 (a) and 4 (d) indicating that these are due to illegal oil transfers from various vessels. The UCM present in the chromatogram could be possibly due to a heavy contribution of petrogenic [28-31]. Significantly, no UCM has been identified in diesel (Figure 4e), and apparently narrow UCM patterns are detected in lube oil and MOLFO (Figure 4b and 4c).

### 3.3. Defining Oil Spill Cases or Problems

The 'define' phase of this study covers the entire step of the oil spill fingerprinting where the whole process from sampling en route to data interpretation was evaluated. Figure 5 outlines the DMAIC approach in identifying the source of the problems and the impacts of the failure of oil classification from the oil spill fingerprinting analysis. This step is necessary because it allows one to investigate the problems in the process of oil spill fingerprinting. Furthermore, the efficiency of the oil classification process can be improved. In the 'define' phase, project team members from the relevant departments were appointed and following a brainstorm session the Ishikawa or Fishbone diagram was generated. From this study majority of the members suggested that the assessment of the potential causes and effects contributes to the failure of the oil spill fingerprinting process.

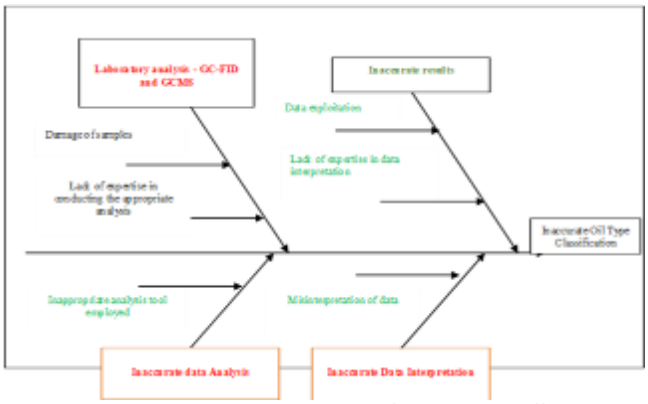



**Fig. 5:** Ishikawa (Fishbone) diagram to define the cause-effect and assessment towards the potential cause and effect determination which contributes to the failure of the oil spill fingerprinting process.

Inappropriate methods in handling oil spill samples and lack of the expertise in conducting the analysis up to weaknesses in data interpretation are seen as common dilemmas of oil spill fingerprinting.

### 3.4. 'Measure' Phase: Field Sampling and Laboratory Analysis of Oil Samples

In the 'measure' phase, the failures found in laboratory analysis en route to data interpretation are due to the lack of adherence in guidelines involving the exchange of oil samples or data between countries on oil spill identification. The use of the statistical - based approach clearly assists in allowing efficient oil classification and eliminates the factors of failure. Furthermore, the lack of expertise or competency leads to errors in data interpretation which results in inaccurate oil spill type classification. DA techniques facilitate in the grouping of the oil spill samples. In this study, the dependent variables are the clusters (types of oil) and the independent variables are the measured oil spill compounds. DA shows four significant clusters of spilled oil samples. DA standard mode discriminant function (DF's) discriminated oil samples to 97.67% classification matrix accuracy (Figure 6).

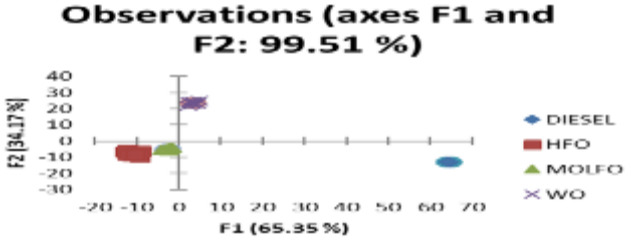



**Fig. 6:** Plot of Discriminant Functions for hydrocarbon oil from the actual sites of oil spill event based on intrinsic chemical properties.

Through DA forward stepwise mode, the oil spill samples are discriminated into 4 discriminant variables ($p < 0.0001$) viz. Diesel, WFO, MOLFO and WO (Table 2) with a 100% classification matrix accuracy. Output of DA forward stepwise mode identifies the discriminating variables in the spatial variance of the oil spill samples.

**Table 2:** Variance factors of PCA after varimax rotation for various of spilled oil hydrocarbon compounds

| Statistics | VF1(FL) | VF2(FL) | VF3(FL) | VF4(FL) |
|---|---|---|---|---|
| Chemical compound | C33(0.703); C39(0.741); C2-N(0.754); C3-N(0.891); C4-N(0.834); C0-P(0.906); C1-P(0.932); C2-P(0.949); C3-P(0.939); C4-P(0.949); C1-D(0.972); C2-D(0.981); C3-D(0.984); C4-D(0.983); C1-F(0.718); C2-F(0.716); C3-F(0.805); C1-C(0.982); C2-C(0.981); C3-C(0.971); An(0.945); FI(0.825); Py(0.949); BaA(0.942); Pe(0.907); BeP(0.982); BP(0.923); BbF(0.983); BkF (0.721); DA(0.937); IP(0.713) | C23Tr (0.944); C24Tr (0.924); 27Ts (0.943); 27Tm (0.978); 29ab (0.889); 30ab (0.989); 31s(0.870); 32r(0.952); 33s(0.988); 33r(0.987); 34s( 0.984); 34r(0.977); 35s(0.984); 35r(0.786); 27bbR(0.975); 27bbS(0.972); 28bbR(0.965); 28bbS(0.958); 29bbR(0.979); 29bbS(0.977) | C10(0.720); C11(0.952); C12(0.989); C13(0.985); C14(0.897); C15(0.965); C16(0.970); C17(0.988); C18(0.982); C19(0.982); C20(0.952); C21(0.898); C22(0.738); CO-F (0.737); BpH(0.735) | C24(0.900); C25( 0.970 ); C26( 0.985 ); C27( 0.989); C28(0.992); C29(0.992); C30( 0.992); C31(0.990); C32(0.989) |
| Eigenvalue | 34.462 | 22.071 | 13.926 | 8.909 |
| Variance (%) | 37.056 | 23.732 | 14.974 | 9.580 |
| Cumulative Variance (%) | 37.056 | 60.788 | 75.762 | 85.342 |

Note: Eigenvalue > 5; Strong factor loadings > 0.7

### 3.5. 'Analyze' Phase of DMAIC

The comprehensive characterization of oil spill fingerprinting, in particular the effectiveness of the analyze phase of DMAIC Six Sigma leads to the identification of the potential causes of failures. Through the brainstorming session, the root-causes that lead to the

failure of oil spill fingerprinting from GC-FID and GC-MS were fairly addressed. The cause-effect diagram is illustrated in Figure 7.

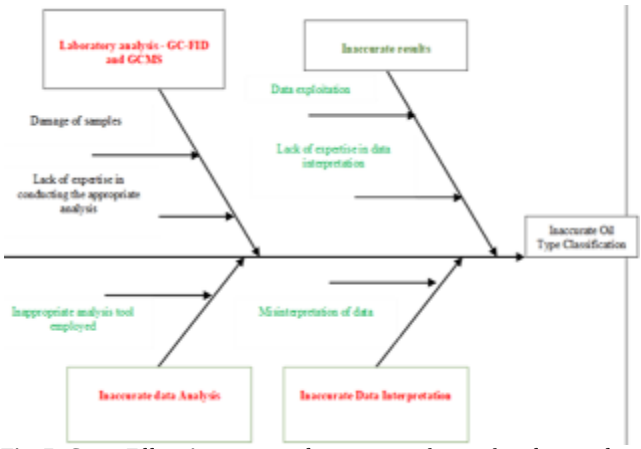



**Fig. 7:** Cause-Effect Assessment of inaccurate oil type classification from oil spill fingerprints using the Ishikawa Diagram

In this study, the classification of the types of oil through a structured process system was implemented based on a series of identification and measures which leads to the main or critical causative factors in order to solve the oil spill identification dilemma. First, the potential cause of the accuracy of oil type classification in oil spill identification was determined to be due to a lack of expertise in conducting oil spill fingerprinting. This may be due to the possibility that the GC-FID and GC-MS measurements were on data interpretation of oil spill as this leads to the causes of inaccurate classification. The use of inappropriate tools in conducting the analysis of the results obtained from GC-FID and GC-MS and lack of knowledge in the various methodologies within a statistical software can be associated with misinterpretation of oil spill data. After an interdisciplinary team had considered the findings, improvement to the oil spill fingerprinting can performed in the DMAIC improve phase.

### 3.6. Implementing Changes to Demonstrate Improvement of Process

In the improvement phase of DMAIC, PCA was applied in this study to reduce the oil spill complex mixture data and ascertain the most significant discriminating variables. The PCA technique enables oil type classification from the complex mixture of spilled oil data extracted from GC-FID and GC-MS. The PCs after varimax rotation yielded four variance factors containing 85.34% of the total variance (with mean data). Table 3 shows the variance factors of the PCA for the various compounds of hydrocarbon spilled oil.

**Table 3:** Characteristics of oil classification with DMAIC Six Sigma before and after the application of the specific domain problem-solving approach

| Number of problem | Prior to DMAIC application | After DMAIC application | Savings |
|---|---|---|---|
| Outsource data analyst | RM 85,000.00 | RM 10,000.00 | RM75,000.00 |
| Inappropriate analysis tool employed | RM 55,000.00 | RM 5,000.00 | RM50,000.00 |
| Misinterpretation of data | RM 45,000.00 | RM 5,000.00 | RM 40,000.00 |

Varimax Factor (VF1) yielded 37.06% of the total oil spill compounds (Figure 6) and 23.73% in VF2 from the total variance of spilled oil in Peninsular Malaysia and Sabah (East Malaysia). The main criterion considered in the varimax rotation is that the number of data to be retained for rotation are eigenvalues, which were greater than one. The factors (VF1) have strong correlations with; the individual saturated hydrocarbon oil variables of *n*-alkanes (C10 - C40), isoprenoids pristane, phytane, naphthalene ($C_0$N - $C_4$N), phenantherene ($C_0$P - $C_4$P), dibenzothiophene ($C_0$D - $C_3$D), flourene ($C_0$F - $C_3$F), chrysene ($C_0$C - $C_3$C) and other EPA priority PAH pollutants (ACL, Ace, BpH, A, FI, Py, BaA, Bph, PE. BeP, BP, BbF, BkF, BaP, DA and IP) as tabled in Table 3. On the other hand, the factors (VF1) are uncorrelated with other compounds of EPA priority PAHs such as $C_{23}$Tr - 29bbS. The activities of illegal transferring of oil from one vessel to another within Johor and Sabah waters and oil discharge from used tyre factories are the main sources of *n*-alkanes (C10 - C40), isoprenoids Pristane, phytane, naphthalene ($C_0$-N - $C_4$-N), phenantherene ($C_0$P - $C_4$P), dibenzothiophene ($C_0$D - $C_3$D), flourenes ($C_0$F - $C_3$F), chrysene ($C_0$C - $C_3$C) and other EPA priority PAH pollutants.

The factor (VF2) exhibits significant positive correlation and communality among other EPA priority PAH pollutants which have a relatively high contribution from waste oil discharge which is either illegal or unintentional as well as factory petroleum discharge. The complex oil mixture dominated by hydrocarbon PAH compounds with high-molecular-weight (four to six rings) [26]. In contrary, VF3 and VF4 exhibit factors with strong collinearity with *n*-alkanes compound from diesel and lubricating oil. The *n*-alkane compounds differentiate the multiple oil spills through its profiles of stable carbon isotopic composition and distribution [32].

Figure 8 shows the PCA score plot in which Diesel, HFO, MOLFO and WO are significantly discriminated from hydrocarbon oil. Outliers 24 and 12 were detected in the scores plot, possibly due to contribution from various factors such as anthropogenic activities, industries and artificial mixed oils. The inliers comprise of thirty observations in four observations in the upper layer of the centroid line and the remainder are under the centroid line. In this study, a substantial dimensional reduction of the dataset and the accurate validation of the oil type classification are achieved. It is clear that PCA performed well in the accurate classification of spilled oil based on GC-FID and GC-MS fingerprints.

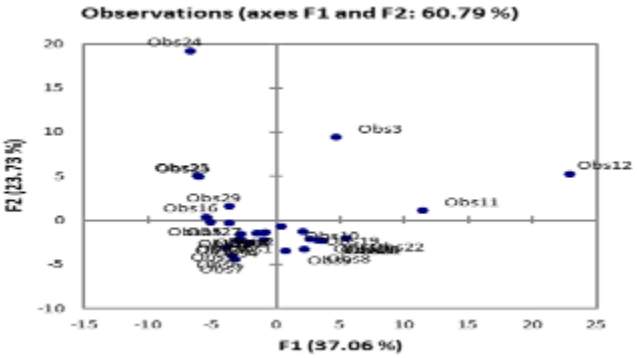



**Fig. 8:** Score Plot of PCA, the oil samples in the analysis replicates the Diesel, HFO, MOLFO and WO

### 3.7. Process Control

Process control is necessary in improving the efficiency of the oil classification process. The Pareto chart (Figure 9) indicates that the oil type classification of hydrocarbon compounds from GC-FID and GC-MS oil spill fingerprinting performed well. Taking the example of hydrocarbon oil C14 in this study, the Pareto chart (Figure 9) indicates that the oil type classification of the hydrocarbon compound from GC-FID and GC-MS oil spill fingerprinting is accurate.

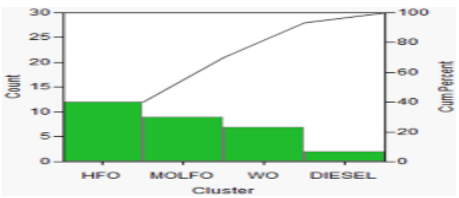



**Fig. 9:** Pareto Chart demonstrate the process control to improve the oil classification through cumulative percentage into four discriminated oil types, HFO, MOLFO, WO and Diesel.

The Pareto chart shows a classification of 100% cumulative performance into four significant oil types viz. Diesel, MOLFO, HFO and WO ($P < 0.05$). An illustration of the Pareto chart clearly reveals the effectiveness of separation, as critical-to-quality (COQ, in oil classification performance based on unique hydrocarbon physicochemical properties.

## 4. Conclusion

DMAIC Six Sigma methodology in this study was applied in the oil spill fingerprinting process as a structured problem-solving approach to improve the conventional process in obtaining better management and determination of fingerprints. The DMAIC approach in this work signifies the importance in the Quality Management System (QMS) in oil spill fingerprint problem-solving which includes the process of chemical analysis using GC-MS and GC-FID en route to data interpretation and the accuracy of data validation. The DMAIC Six Sigma process is a reliable, cost effective method for the overall quality improvement and enhancement of research development that enables efficient and effective classification of oil spill fingerprints obtained from various sources in Peninsular Malaysia and Sabah (East Malaysia). The problem-solving techniques of DMAIC demonstrates the well-defined step-by-step methodology leading to solutions via the various phases of DMAIC. Each phase of DMAIC defines the root-cause of problems for improvement. The Cause-Effect diagram isolates the primary causes of oil type classification from oil spill fingerprints for improvement. Output of DA reveals four significant variables; diesel, hydrocarbon fuel oil (HFO), mixture oil lubricant and fuel Oil (MOLFO) and waste oil (WO) discriminated from the spilled oil samples. PCA output in the Improvement phase showed classification of four (4) oil types that is responsible for 60.79% of the total spilled oil variation. The application of various multivariate techniques allows the oil classification to be optimized and performed at a lower cost. Through the Pareto chart, the Process Control results in a 100% cumulative oil type classification with 95% confidence level. With this work standing as an example, the well-defined step-by-step DMAIC method can act as an important problem-solving methodology in all areas of science.

## Acknowledgement

The Authors wish thank the East Coast Environmental Research Institute (ESERI), UNISZA, Department of Environment (DOE) Malaysia, Department of Chemistry Malaysia, Department of Chemistry, University of Malaya and Integrated Invirotech Sdn. Bhd. for their assistance, provision of data and financial help.

## References

[1] C. L. Leaphart, T. Gonwa, M. L Mai, M. B. Prendergast, H. M. Wadei, J. J. Tepas & C. B. Taner, Formal quality improvement curriculum and DMAIC method results in interdisciplinary collaboration and process improvement in renal transplant patients. Journal of Surgical Research, 177(1) (2012) 7–13.

[2] S. Koziolek and D. Derlukiewicz, 2012. Method of assessing the quality of the design process of construction equipment with the use of DFSS (design for Six Sigma). Automation in Construction, 22 (2012) 223-232.

[3] M.P.J. Pepper, T.A. Spedding, The evolution of Lean Six Sigma. Int.J.Qual. Reliab.Manag. 27(2) (2010) 138–155.

[4] I. Manuj & F. Sahin, A model of supply chain and supply chain decision-making complexity. International Journal of Physical Distribution and Logistics Management, 41(5) (2011) 511–549.

[5] R. Liu & A. Kumar, Leveraging information sharing to configure supply chains. Information Systems Frontiers, 13(1) (2011) 139–151.

[6] T. J. Goldsby & R. Martichenko, Lean Six Sigma logistics: Strategic development to operational success. Plantation, FL: J. Ross Publishing, 2005.

[7] B.W. Jacobs, M. Swink & K. Linderman (2015). Performance effects of early and late six sigma adoption. Journal of Operations Management 36 (2015) 244–257.

[8] K. Jayaraman, T.L. Kee, K.L. Soh, The perceptions and perspectives of Lean Six Sigma (LSS) practitioners: anempirical study in Malaysia. TQMJ 24(5) (2012) 433–446.

[9] S. Kumar and S. Schmitz, Managing recalls in a consumer product supply chain – root cause analysis and measures to mitigate risks. International Journal of Production Research, 49 (1) (2011).

[10] H. Juahir, A. Ismail, S.B. Mohamed, M.E. Toriman, A.M. Kassim, S.M. Zin, W.K.W. Ahmad, K.F. Wong, A. Retnam, M.A. Zali, M.Z.M. Taib, M. Mokhtar, 2017. Improving Oil Classification Quality from Oil Spill Fingerprint beyond Six Sigma Approach. Marine Pollution Bulletin 120 (1-2) (2017) 322 – 332.

[11] S. J. Harjac, A. Atrens & C.J. Moss (2008). Six Sigma reviews of root causes of corrosion incidents in hot potassium carbonate acid gas removal plant. Engineering Failure Analysis 15 (2008) 480-496.

[12] M. Kashif, S. Z. Melatu Shamsi, Z. Awang and M. Mohamad, EXQ Measurement of healthcare experience quality in Malaysian settings: A contextualist perspective. International Journal of Pharmaceutical and Healthcare Marketing, 10 (1) (2016) 27-47.

[13] J.D. Westphal, R. Gulati, S.M. Shortell, Customization or conformity? An institutional and network perspective on the content and consequences of TQM adoption. Adm. Sci. Q. 42 (2) (1997) 366–394.

[14] EPA Methods for Chemical Analysis of Water and Wastes 1986. US EPA, Washington, DC, 1986.

[15] EPA test methods for evaluating solid waste (SW-846). Update III, US EPA, Office of Solid Waste and Emergency Response, Washington, DC, 1997.

[16] EPA Method, 1664, Revision A: n-Hexane extractable material and silica gel treated n-hexane extractable material by extraction and gravimetry. EPA-821-R-98-002, US EPA, Office of Water, Washington, DC, 1999. Etkin, D.S. (Ed.), 1998; N614.

[17] Z. Wang, C. Yang, M. Fingas, B. Hollebone, X. Peng, A. Hansen, and J. Christensen, Characterization, Weathering, and Application of Sesquiterpanes to Source Identification of Spilled Lighter Petroleum Products, Environ. Sci. Technol. 39 (2005) 8700-8707.

[18] P.S. Daling, L.G. Faksness, A.B. Hansen, S.A. Stout, Improved and standardized methodology for oil spill fingerprinting. In: Proceedings of the 25th Arctic and Marine Oil Spill Program (AMOP) Technical Seminar, Environment Canada, Ottawa, 2002; 429–454

[19] F. Zhou, H. Guo, Y. Liu & Y. Jiang, Chemometrics data analysis of marine water quality and source identification in Southern Hong Kong. Marine Pollution Bulletin 54 (2007) 745-756.

[20] P. R. Kannel, S. Lee, S. R. Kanel & S. P. Khan, Chemometric application in classification and assessment of monitoring locations of an urban river system. Analytical Chimica Acta 582 (2007) 390-399.

[21] A. Blundell, J.A. Hannam, J.A. Dearing & J.F. Boyle, Detecting atmospheric pollution in surface soils using magnetic measurements: A reappraisal using an England and Wales database. Environmental Pollution 157(10) (2009) 2878–2890.

[22] H. Juahir, S.M. Zain, M. K. Yusoff, T. I. T. Hanidza, A.S. M. Armi, M.E. Toriman & M. Mokhtar, (2011). Spatial water quality assessment of Langat River Basin (Malaysia) using environmetric techniques. Environment Monitoring Assessment 173 (2011) 625-641.

[23] A. Retnam, M. P. Zakaria, H. Juahir, A.Z. Aris, M.A. Zali, & M.F. Kasim, Chemometric techniques in distribution, characterisation and source apportionment of polycyclic aromatic hydrocarbons (PAHS) in aquaculture sediments in Malaysia. Marine Pollution Bulletin 69 (1-2) (2013) 55–66.

[24] P. Fermo, G. Beretta, R. M. Facino, F. Gelmini, & A. Piazzalunga Ionic profile of honey as a potential indicator of botanical origin and global environmental pollution. Environmental Pollution 178 (2013) 173–181.

[25] A. Ismail, M.E. Toriman, H. Juahir, M.Z. Sharifudin, N.L. Abdul Habir, A. Retnam, M.K.A. Kamarudin, R. Omar and A. Azid, Spatial assessment and source identification of heavy metals pollut pollution in surface water using several chemometric techniques. Marine Pollution Bulletin 106 (1-2) (2016a) 292-300.

[26] A. Ismail, M.E. Toriman, H. Juahir, M.A. Kassim, M.Z. Sharifudin, W.K.W. Ahmad, K.F. Wong, A. Retnam, M.A. Zali, M. Mokhtar and M.A. Yusri, Chemometric techniques in oil classification from oil spill fingerprinting. Marine Pollution Bulletin 111(1-2) (2016b) 339-346.

[27] K.E. Peters, J.M. Moldowan, The Biomarker Guide. Interpreting Molecular Fossils in Petroleum and Ancient Sediments, Prentice Hall, Englewood Cliffs, N. J., 1993; 363.

[28] I. Bouloubassi, A. Saliot, Sources and transport of hydrocarbons in the Rhone delta sediments (Northwestern Mediterranean). Fresenius Journal of Analytical Chemistry 339 (1991) 765 – 771
[29] J.K. Volkman, D.G. Holdsworth, G.P. Neill, H.J.J. Bavor, Identification of natural, anthropogenic and petroleum hydrocarbons in aquatic sediments. Science of the Total Environment 112 (1992) 203-219.
[30] S.A. Stout, A.D. Uhler, K.J. McCarthy, S. Emsbo-Mattingly, Chemical fingerprinting of hydrocarbons. In: Murphy, B.L., Morrison, R.D. (Eds.), Introduction to Environmental Forensics. Academic Press, London, 2002; 137–260. Chapter 6.
[31] M. Kim, S. H. Hong, J. Wona, U. H. Yim, J.-H. Jung, S. Y. Ha, J. G. Ana, C. Joo, E. Kim, G. M. Han, S. Baek, H.W. Choi, W. J. Shim, Petroleum hydrocarbon contaminations in the intertidal seawater after the Hebei Spirit oil spill - Effect of tidal cycle on the TPH concentrations and the chromatographic characterization of seawater extracts. Water Research, 47 (2013) 758 -768.
[32] Y. Li; Y. Xiong; W. Yang; Y. Xie; S. Li; Y. Sun. Compound-specific stable carbon isotopic composition of petroleum hydrocarbons as a tool for tracing the source of oil spills. Marine Pollution Bulletin 58(1) (2009) 114-117.
[33] Ismail, A., Toriman, M. E., Juahir, H., Zain, S. M., Habir, N. L. A., Retnam, A., ... & Azid, A. (2016). Spatial assessment and source identification of heavy metals pollution in surface water using several chemometric techniques. *Marine pollution bulletin*, *106*(1), 292-300.
[34] Al-Odaini, N. A., Zakaria, M. P., Zali, M. A., Juahir, H., Yaziz, M. I., & Surif, S. (2012). Application of chemometrics in understanding the spatial distribution of human pharmaceuticals in surface water. *Environmental monitoring and assessment*, *184*(11), 6735-6748.