CRANFIELD UNIVERSITY

Reuben N Okparanma, B.Tech., M.Tech.

RAPID MEASUREMENT OF POLYCYCLIC AROMATIC HYDROCARBON CONTAMINATION IN SOILS BY VISIBLE AND NEAR-INFRARED SPECTROSCOPY

SCHOOL OF APPLIED SCIENCES

ENVIRONMENTAL SCIENCE AND TECHNOLOGY

PhD THESIS

CRANFIELD UNIVERSITY

SCHOOL OF APPLIED SCIENCES

ENVIRONMENTAL SCIENCE AND TECHNOLOGY

PhD THESIS

Academic Year 2010–2013

Reuben N. Okparanma, B.Tech., M.Tech.

Rapid measurement of polycyclic aromatic hydrocarbon contamination in soils by visible and near-infrared spectroscopy

Main supervisor: Dr Abdul M. Mouazen
Co-supervisor: Dr Thomas Mayr

October 2013

This thesis is submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

Abstract

Polycyclic aromatic hydrocarbons (PAHs) are widely distributed organic pollutants. At petroleum contaminated sites, PAHs are often the key risk drivers because of their carcinogenicity. Assessing the risk of PAH at contaminated sites by conventional soil sampling, solvent extraction and gas chromatography–mass spectrometry (GC–MS) analysis is expensive and time-consuming. Employing a rapid and cheap measurement technique for PAH would be beneficial to risk assessment by eliminating costs and time associated with the conventional method. The literature has shown that visible and near infrared (vis-NIR) spectroscopy is a rapid and cheap technique for acquiring information about key soil properties. In this study, models based on vis-NIR spectroscopy (350–2500 nm) were developed to predict and map PAH in contaminated soils for the ultimate aim of informing risk assessment and/or remediation. The reference chemical analytical method used was GC–MS while the multivariate analytical technique used for model development was partial least squares (PLS) regression analysis with full cross-validation. A total of 150 soil samples from the UK were used for the laboratory-scale study while 137 samples were used for the near-onsite adaptive trials at three oil spill sites in Ogoniland, Niger Delta province of Nigeria. Both laboratory- and field-scale results showed that soil diffuse reflectance decreased with increasing PAH concentration. Hydrocarbon absorption features observed around 1647 nm in the first overtone region of the NIR spectrum showed a positive link to PAH. Laboratory-scale study showed that both individual and combined effects of oil concentration, and moisture and clay contents on soil spectral characteristics and calibration models were significant ($p<0.05$). For the field-scale study, inverse distance weighting soil maps of PAH developed with chemically-measured and vis-NIR-predicted data were comparable with a fair to good agreement between them (Kappa coefficient = 0.19–0.56). Hazard assessment of the oil spill sites using both measurement methods showed that the impact of the contamination varied distinctly across the management zones. The type of action required for site-specific risk assessment and/or remediation also varied among the different zones. This result shows promise that vis-NIR can be a good screening tool for petroleum release sites.

Acknowledgements

First, I would like to express my sincere gratitude to my supervisor, Dr. Abdul M. Mouazen, for his resolve and astuteness in seeing this work through to its logical conclusion right on schedule. But for his doggedness and tutelage, this work would not have achieved the much it did in such a short time. I also thank my course adviser, Dr. Frédéric Coulon, for bringing his wealth of experience on environmental chemistry and bioremediation to bear on this work. Other members of the thesis committee that I would like to acknowledge are the Chair, Dr. Sue Impey, and my co-supervisor, Dr. Thomas Mayr.

I also would like to acknowledge the technical support team within the Environmental Analytical Facility for their kind cooperation and guidance during practical sessions in the laboratories, especially Dr. Keith Richards for his invaluable help on the GC–MS system.

I am grateful to the Petroleum Technology Development Fund (PTDF), Nigeria, for fully funding this research. The same gratitude goes to Rivers State University of Science and Technology, Port Harcourt, Nigeria, for granting me the leave of absence to undertake this research. I thank my colleagues, Assoc. Prof. AJ Akor, Prof. MJ Ayotamuno, Dr. AH Igoni, Dr. I Fubara-Manuel, Engr. S Nkakini, and Engr. D Davidson, for their cooperation while the study lasted.

I would also like to thank my friends in the research group: Dr Boyan Kuang (China), Mohammed Aldhumayri (Iraq), Dr Zaka Quraishi (Pakistan), Raed Al-Asadi (Iraq), Dr Yucil Tekin (Turkey), and Graham Halcro (Scotland) for their contributions during the study. I also would acknowledge the contributions of anonymous reviewers, which did help to shape this work.

Finally, I thank Glory, my wife and my sons, Emmanuel, Justin, Divine, and David for their constant love, support and understanding throughout my PhD. My bosom friends, Engr. Perez Araka, Mr. Ebenezer Jumbo, Mr. Chigozi Wali, and Mr. Dike Donatus are all highly appreciated.

Table of Contents

List of Figures

List of Tables

List of Equations

Nomenclature

$\bar{x}$ = the sample mean,

$\hat{y}_{CV.i}$ = Estimate for y_i based on the calibration equation with *i* deleted

$\hat{y}$ and y_i = Predicted and measured reference values respectively

a_1 = PAH concentrations in the first duplicate sample

a_2 = PAH concentrations in the second duplicate sample

C = Concentration

E = Measured mean PAH value from current study

n = Number of measurements

N = Number of samples in the set

R = Spectral reflectance

s = Sample standard deviation

SD, Standard deviation of the measured reference values

$E(.)$ = Statistical expectation (average) over the population of future samples

T = “Known” PAH value from commercial laboratory

t = Student’s *t* for a desired level of confidence

Abbreviations

ACE, Acenaphthylene

ANOVA, Analysis of variance

ANT, Anthracene

ASTM, American Society for Testing and Materials

BSI, British Standard Institution

CCD, Charge Coupled Detector

COPC, Constituents of potential concern

CSB, Compound-specific biomarkers

DPR, Department of Petroleum Resources

DRIFT, Diffuse Reflectance Infrared Fourier Transform

DTGS, Deuterated triglycine sulphate

ECIA, Electrochemical immunoassay

ELISA, Enzyme-linked immunosorbent assay

FLU, Fluorene

FS, Fluorescence spectroscopy

FTIR, Fourier Transform Infrared

GAC, Generic Assessment Criteria

GC–FID, Gas chromatography–flame ionisation detection

GC–MS, Gas chromatography–mass spectrometry

IDW, Inverse Distance Weighting

IMA, Immunoassay

IR, Infrared

ISO, International Standard Organisation

LDL, Laboratory detection limit

LIF, Laser-induced fluorescence

LOQ, Limit of quantitation

MC, Moisture content

MIR, Mid-infrared

MSD, Mass selective detection

NDES, Niger Delta Environmental Survey

PAH, Polycyclic aromatic hydrocarbons

PCA, Principal component analysis

PHC, Petroleum hydrocarbons

PHE, Phenanthrene

PHS, Priority hazardous substances

PLS, Partial least squares

PYR, Pyrene

RMSECV, Root mean square error of cross-validation

RMSEP, Root mean square error of prediction

ROST, Rapid Optical Screening Tool

RPD, Ratio of prediction deviation

SEP, Standard error of prediction

SPDC, Shell Petroleum Development Company of Nigeria

TEC, Toxicity equivalent concentration

TEF, Toxicity equivalency factor

TPH, Total petroleum hydrocarbons

TTEC, Total toxicity equivalent concentration

UNEP, United Nations Environment Programme

US EPA, United States Environmental Protection Agency

USDA, United States Department of Agriculture

UVIF, Ultraviolet-induced fluorescence

Vis-NIR, Visible and near-infrared

WSDE, Washington State Department of Ecology

Chapter 1: General Introduction

1.1 Oil pollution in Nigeria

The history of crude oil production, and causes and effects of crude oil pollution in Nigeria have been widely reported in the open literature (e.g., Van-Dissel and Omuka, 1994; Benka-Coker and Ekundayo, 1995; Ekundayo and Obuekwe, 1997; Ogri, 2001; Ayotamuno et al., 2002; Aroh et al., 2010; Kadafa, 2012). In Nigeria, crude oil was first discovered in 1956 at Oloibiri in Bayelsa State in the Niger Delta province and commercial production began in 1958 (Van-Dissel and Omuka, 1994). Activities involved in crude oil production include exploration, drilling, processing, transportation, and storage (DPR, 2002). Most of these activities are known to cause various environmental problems in operational areas due to oil spills brought about by equipment failure and/or wilful damage of production facilities by vandals (SPDC, 2013). Available records show staggering levels of oil spill incidence in Nigeria since records began; even though the accuracy of some of the data has often come under strong criticisms. Table 1-1 shows the record of oil spill incidence in the Niger Delta province of Nigeria from 1976 to 2005 (DPR, 2006). This reflects the status of the Niger Delta province as host to most of Nigeria's oil production facilities (NDES, 1995). Since oil and gas became the mainstay of Nigeria's economy (NDES, 1995), the province has continued to bear the brunt of oil and gas production activities in the country. For instance, Figure 1-1 shows the latest statistics of oil spill incidence in the Niger Delta province from 2007 to 2013 (SPDC, 2013), which suggests that in the last 37 years (1976 to 2013) oil spills have remained unabated in the province. The cost of petroleum hydrocarbon

(PHC) decontamination in Nigeria is unknown as yet. Nonetheless, with hindsight of an estimated initial cost of $1 billion required for the first 10 years of a 30-year proposal for hydrocarbon decontamination in Ogoniland in the Niger Delta province (UNEP, 2011), PHC decontamination in Nigeria is expected to run into tens of billions of US Dollars.

Currently, dealing with oil pollution in Nigeria has been largely by Remediation by Enhanced Natural Attenuation (RENA) (e.g., Ebuehi et al., 2005), even though the method has recently been reported as unsuitable in the Niger Delta because of the depth (over 5 m) of oil penetration (UNEP, 2011). Usually, RENA is preceded by environmental impact assessment (EIA) involving assessment of the risk of PHCs in the contaminated site. In Nigeria, assessing the risk of PHCs at contaminated sites relies on the measurement of the total petroleum hydrocarbon (TPH) content of the contaminated soil apparently as a risk indicator compound as required by regulation (DPR, 2002). TPH is a collective term describing a complex mixture of hydrocarbons (Weisman, 1998). Complex TPH mixture measured as a single concentration number inherently is not a direct indicator of risk to either the environment or humans but, may give an indication of the degree of hydrocarbon contamination in soil (Weisman, 1998). Thus, for risk-based assessment of petroleum contaminated sites, the literature recommends the determination of the polycyclic aromatic hydrocarbon (PAH) fractions (Askari and Pollard, 2005). This informed the decision in this study to focus on the measurement of PAHs as the key risk drivers at petroleum contaminated sites rather than TPH.

Table 1-1 Oil spill data in the Niger Delta, Nigeria (1976–2005) (Modified after DPR, 2006)

Year	Number of oil spills	Quantity spilled (m^3)[a]	Quantity recovered (m^3)[a]
1976	128	4,158.62	1,134.37
1977	104	5,227.33	270.76
1978	154	77,791.39	62,234.67
1979	157	110,364.01	10,092.69
1980	241	95,473.44	6,743.72
1981	238	6,792.24	869.69
1982	252	6,809.57	345.22
1983	173	7,687.23	1,010.51
1984	151	6,392.71	261.50
1985	187	1,888.23	273.35
1986	155	2,051.73	87.76
1987	129	5,066.28	971.25
1988	208	1,458.23	310.82
1989	195	1,212.78	342.30
1990	160	2,375.40	332.69
1991	201	16,984.26	442.93
1992	378	8,138.22	234.78
1993	428	1,550.48	466.96
1994	515	4,814.55	371.38
1995	417	10,123.84	494.45
1996	430	7,369.54	188.08
1997	339	12,993.67	No data
1998	399	15,880.47	No data
1999	225	2,687.51	No data
2000	637	13,366.34	No data
2001	412	19,233.64	No data
2002	446	38,414.05	No data
2003	609	5,609.77	No data
2004	543	2,719.31	No data
2005	496	1,706.66	No data
Total	9,107	496,343.07	87,479.88

[a] Original unit is barrels (bbl.)

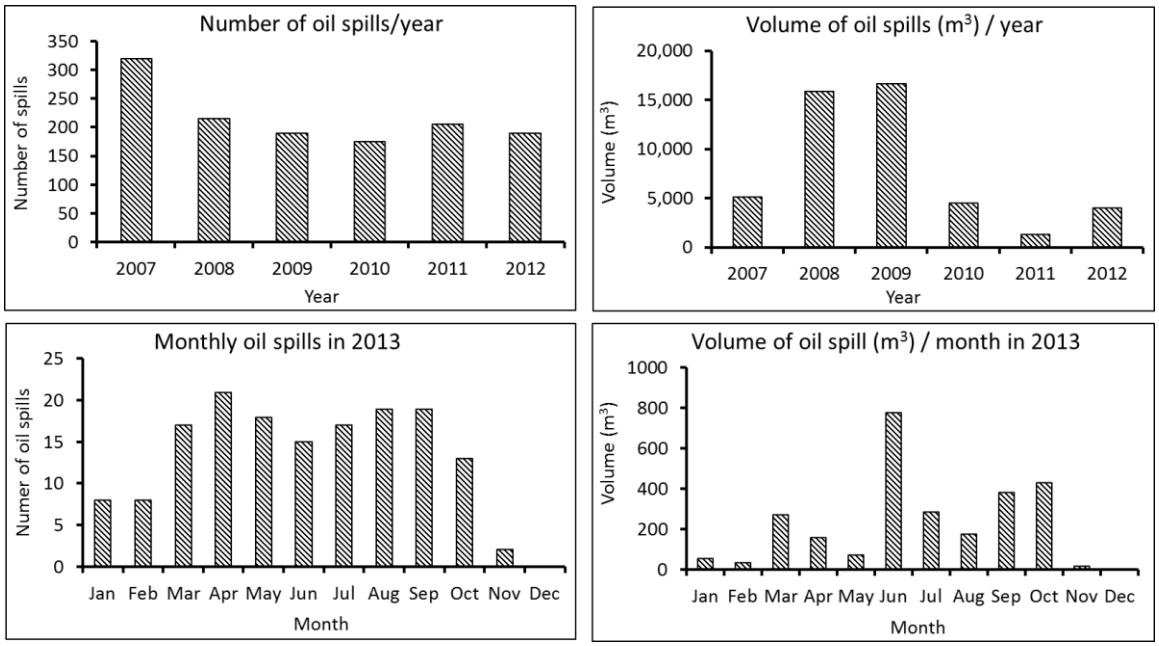


Figure 1-1 Oil spill incidence in the Niger Delta, Nigeria from 2007 to 2013. Volume of oil spill was originally in barrels (bbl.) (Modified after SPDC, 2013).

1.2 Chemical nature of polycyclic aromatic hydrocarbons

Polycyclic aromatic hydrocarbons (PAHs) are the class of hydrocarbons containing two or more fused aromatic hydrocarbons. Figure 1-2 shows the chemical structures of the sixteen United States Environmental Protection Agency (US EPA) priority PAHs.

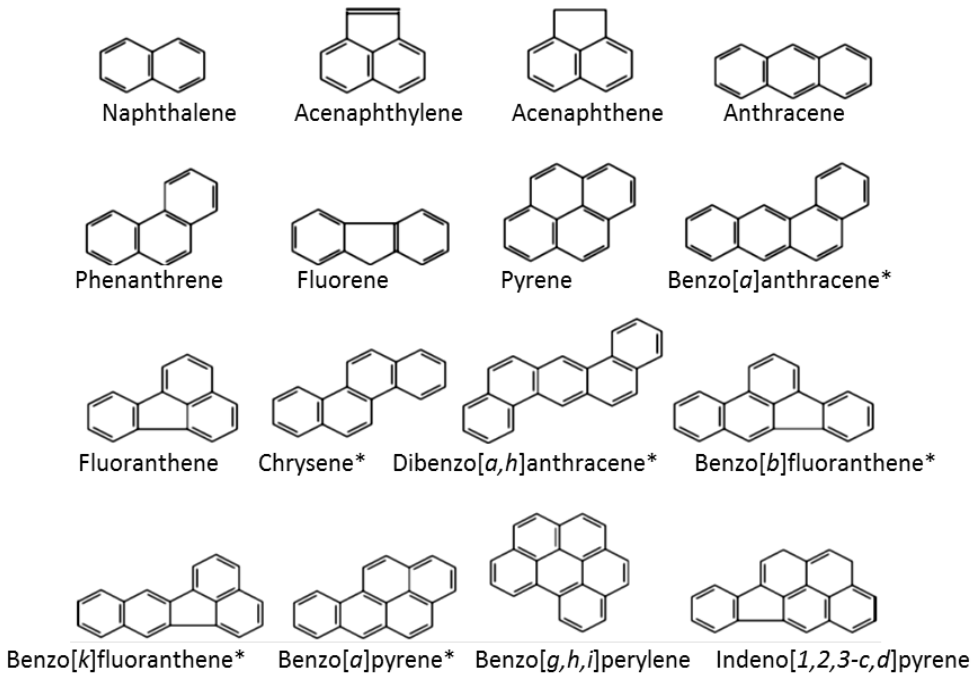


Figure 1-2 Chemical structures of the sixteen United States Environmental Protection Agency priority polycyclic aromatic hydrocarbon compounds.
*Non-threshold indicator compounds, also known to possess some genotoxic carcinogenic potential (Contaminated Land: Applications in Real Environments, 2010).

PAHs are found in large quantities in crude oil and its daughter products including diesel fuel #2 (Chemical Abstract Service [CAS] No. 68476-34-6) – here referred to as diesel. The weathering of crude and mineral oils once spilled on land introduces petroleum hydrocarbons (including PAHs), which negatively

impact soil's biological, chemical, and physical characteristics (Umechuruba, 2005; Adedokun et al., 2006; Daniel-Kalio and Pepple, 2006). PAHs are harmful to the environment because some are potentially carcinogenic. According to the Washington State Department of Ecology (2007), carcinogenic PAHs are those identified as Group A (known human) or B (probable human) carcinogens by the US EPA.

1.3 Physical nature of polycyclic aromatic hydrocarbons

The solubility and sorption characteristics of PAHs are important factors that control their fate and transport in the soil (Huang, et al., 2003). Since PAHs exhibit strong hydrophobicity, they primarily sorb to the organic matter or organic carbon (OC) in the soil (Huang, et al., 2003). Figure 1-3 illustrates the mechanism in which hydrocarbons will be sorbed to OC in the soil. The clay–organic matter interaction has a large influence on the sorption of PAHs in soil, which will vary in part with the amount of organic matter in the soil (Dexter et al., 2008).

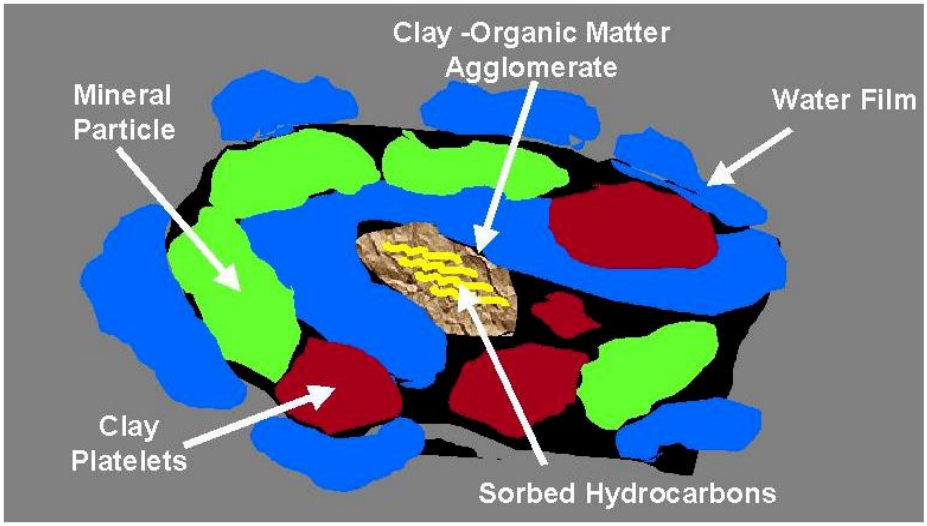


Figure 1-3 Hydrocarbon interactions with soil (Source: Prommer et al., 2003)

1.4 Managing polycyclic aromatic hydrocarbons in soil

Spillage of petroleum products is a common occurrence in both the upstream and downstream sectors of the oil and gas industry (Atlas, 1981; Ebuehi et al., 2005). Since PAHs are major constituents of petroleum products and are harmful to humans and the environment (Adedokun and Ataga, 2006; Daka and Ekweozor, 2004; Jack et al., 2005), it is of utmost importance to rid the soil of these hazardous substances as quickly as they are spilled. Nonetheless, to completely get rid of PAHs from the soil would be a difficult task because of the growing dependence of our daily lives on petroleum products. Thus, best practice requires that a viable environmental management system (EMS) is put in place for effective management of PAHs in soil. One of the major components of EMS for a petroleum release site is risk assessment. In the tiered risk-based assessment framework for a petroleum release site, hazard assessment is crucial. Hazard assessment is aimed at identifying both the chemistry and the areal extent to which the PAHs exceed their applicable soil screening levels, as well as the likely risk to specified targets (Petts et al., 1997). The outcome of this determines whether or not further action is needed to reclaim the contaminated site.

1.5 Challenges of methods for managing polycyclic aromatic hydrocarbons in soil

1.5.1 Conventional methods

Traditional methods of hazard assessment in petroleum release sites involve prior manual soil sampling, extraction of PAH compounds from the soil sample

using various extraction solvents, and analysis of the liquid extract by traditional laboratory methods such as gas chromatography–mass spectrometry (GC–MS) (EPA, 1999a). Soil sampling is tedious, time consuming, and labour-intensive. Solvent extraction of PAHs is time-consuming and hazardous. PAH analysis by GC–MS is uneconomical; particularly when large-scale contamination is involved and dense sampling is required for high-resolution soil mapping of contaminants (Peterson, 2002; Okparanma and Mouazen, 2013a). This has prompted increasing demand for rapid, portable, and cost-effective measurement methods for PAHs capable of collecting high-resolution spatial information for effective management of the contaminants.

1.5.2 Innovative methods

Over the years, several portable techniques have been developed and employed for the rapid measurement of PAHs in soils. These include field-portable GC–MS systems (Barnes, 2009), immunoassay techniques (Zhou et al., 2009; Wei et al., 2009), and visible and near-infrared (vis-NIR) spectroscopy (Bray et al., 2010). Of these, only the vis-NIR spectroscopic method has been successfully used for on-line (tractor-mounted) measurement of soil properties (e.g., Mouazen et al., 2007a; Kuang and Mouazen, 2013), which opens the window of opportunity for on-line measurement of PAHs using the approach. Nevertheless, minimal research can be found on how vis-NIR spectroscopic method will perform in the presence of moisture and mineral oil across a wide range of soil texture due to the interaction effect of these soil variables. Moreover, vis-NIR spectroscopy has not been used yet for hazard assessment

of petroleum release sites for the management of PAHs in soils. Vis-NIR spectroscopy has not also been used to measure PAHs in Nigerian soils.

1.6 Near-infrared analysis

The historical perspectives, fundamental principles, and practical applications of NIR spectroscopy have been widely reported in the open literature (e.g., Williams and Norris, 1987; Osborne et al., 1993; Pasquini, 2003; Adamchuk et al., 2004; Davies, 2005).

1.6.1 Historical perspectives

What is now known as NIR in the electromagnetic spectrum was discovered by William Herschel in 1800. With inputs from early studies by Ampere in 1835 and James Maxwell in 1864, the first NIR spectra were measured in 1881 by Abney and Festing who predicted that spectral absorptions were related to the chemical composition of liquids. In 1905, William W. Coblentz published the spectra and corresponding wavelengths (1000–10,000 nm) of a large list of compounds thus, paving the way for vibrational spectroscopy of molecules in the mid-infrared (2500–50,000 nm) region. After World War II, study of the NIR region (700–2500 nm) was extended to quantitative measurement of a few simple organic compounds. Then in 1960, Karl Norris at the United States Department of Agriculture (USDA) showed the usefulness of the NIR region in quantitative analysis of agricultural samples (Davies, 2005).

1.6.2 Fundamental principles

NIR spectroscopy uses the NIR region of the electromagnetic to identify certain properties of a material. NIR spectroscopy is a type of vibrational spectroscopy in which specific properties of a material are identified by the characteristic energy absorbed by chemical bonds in the NIR spectrum. In vibrational spectroscopy, vibration of chemical bonds between atoms in a molecule due to the absorption of energy is explained by simple harmonic (parabolic) approximation (Atkins and de Paula, 2002). In NIR spectroscopy, absorption of energy by substances is due to overtones and combinations of fundamental vibrations that occur in the mid-infrared region linked to the stretching and bending of chemical bonds involving hydrogen and other atoms such as C–H, O–H, N–H, and S–H (Osborne et al., 1993). Using chemometrics, these overtones and combinations are assigned bands with certain attributes such as intensity and wavelength position, which may indicate the origin of the spectral feature to help in the recognition of certain physicochemical properties of the substance (Pasquini, 2003).

1.6.3 Practical applications

Although various measurement modes of sampling are employed in NIR spectroscopy, the most commonly used for solid materials (e.g., soil) is diffuse reflectance (Murray, 2012). The apparent ‘miscibility’ between chemometrics and NIR spectroscopy stems from the dependence of diffuse reflectance measurement on the highly variable physical properties of a sample (such as particle size). This variability in physical variables causes changes in the signal

intensity due to light scatter effects and absorbance by solid particles. To develop an analytical procedure based on reflectance measurements, the expression in equation 1-1, which does not depart much from the Kubelka-Munk prediction (Pasquini, 2003), is normally employed.

$$f(C) = Log_{10}\frac{1}{R} \quad (1\text{-}1)$$

Where R = spectral reflectance, and C = concentration. Therefore, to extract the relevant information necessary for a given analytical procedure, multivariate analysis is employed (Pasquini, 2003). Various linear multivariate techniques that have been adopted include multiple linear regression, principal component regression, partial least squares (PLS) regression analysis and penalised spline; while the non-linear techniques include artificial neural network, support vector machine, boosted regression trees, random forest and wavelet transform (Malley et al., 1999; Gammoun et al., 2006, 2007; Perez-Caballero et al., 2009; Chakraborty et al., 2010; Dos Santos et al., 2010; Viscarra Rossel and Behrens, 2010; Mouazen et al., 2010).

With the replacement of the first generation add-on (with visible, ultraviolet or mid-infrared) NIR units by the new generation single stand-alone units, the application of vis-NIR spectroscopy in the measurement of soil physicochemical properties both spatially and temporally has become widespread. Although much of the early applications of NIR spectroscopy were for qualitative purposes in the food and beverage industry (Osborne et al., 1993), its earliest

qualitative and quantitative applications in soil science predominantly for agricultural purposes were reported in the 1980s (Adamchuk et al., 2004).

In recent years, the scope of NIR in the analysis of environmental media has transcended soil science. It is now known that organic materials of petroleum origin such as mineral oils have strong spectral features in the NIR region (Aske et al., 2001). In vis-NIR spectroscopy (350–2500 nm), absorption of energy by hydrocarbon derivatives is due to overtones and combinations of fundamental vibrational C–H stretching modes of saturated CH_2 and terminal CH_3, or aromatic C–H (ArCH) functional groups (Aske et al., 2001). The difficulty in interpreting vis-NIR spectra because of broad and overlapping bands (Stenberg, 2010), to a large extent has been overcome by the use of advanced chemometrics and data-processing techniques (Pasquini, 2003). In the analysis of spectroscopic data, multivariate calibration generally solves the problem of interference from compounds closely related to the analyte thereby eliminating the need for selectivity (Naes et al., 2002).

1.7 Research aim and objectives

1.7.1 Aim

The first part of this research, which is laboratory-based, aims to investigate the combined influences of moisture content, texture, and oil concentration on the performance of vis-NIR in the measurement of PAHs in soils. The second part, which is field-based, aims to apply the methodology to hazard assessment of petroleum release sites in Ogoniland in the Niger Delta province of Nigeria using the knowledge from the laboratory investigation.

1.7.2 Objectives

To achieve the research aim, the following specific objectives have to be addressed:

1) To critically review relevant literatures on selected conventional and non-conventional techniques for the analysis of petroleum hydrocarbons in contaminated soils.
2) To analyse the combined effects of diesel concentration, and clay and moisture contents on the spectral characteristics and quality of partial least-squares (PLS) regression models developed with vis-NIR spectroscopy to predict PAHs in diesel-contaminated soils under laboratory controlled conditions.
3) To evaluate the ability of PLS models developed with vis-NIR spectroscopy to predict individual PAH fractions in diesel-contaminated soils under laboratory controlled conditions.

4) To evaluate the performance of PLS models developed with vis-NIR spectroscopy to predict PAHs in contaminated soil samples collected from three petroleum release sites in Ogoniland in the Niger Delta province of Nigeria.
5) To develop soil maps of PAH and total toxicity equivalent concentrations (TTEC) for hazard assessment of the three petroleum release sites in Ogoniland in the Niger Delta province of Nigeria.

1.8 Thesis structure

The schematic of the structure of this thesis showing how the laboratory- and field-level studies are linked is shown in Figure 1-4.

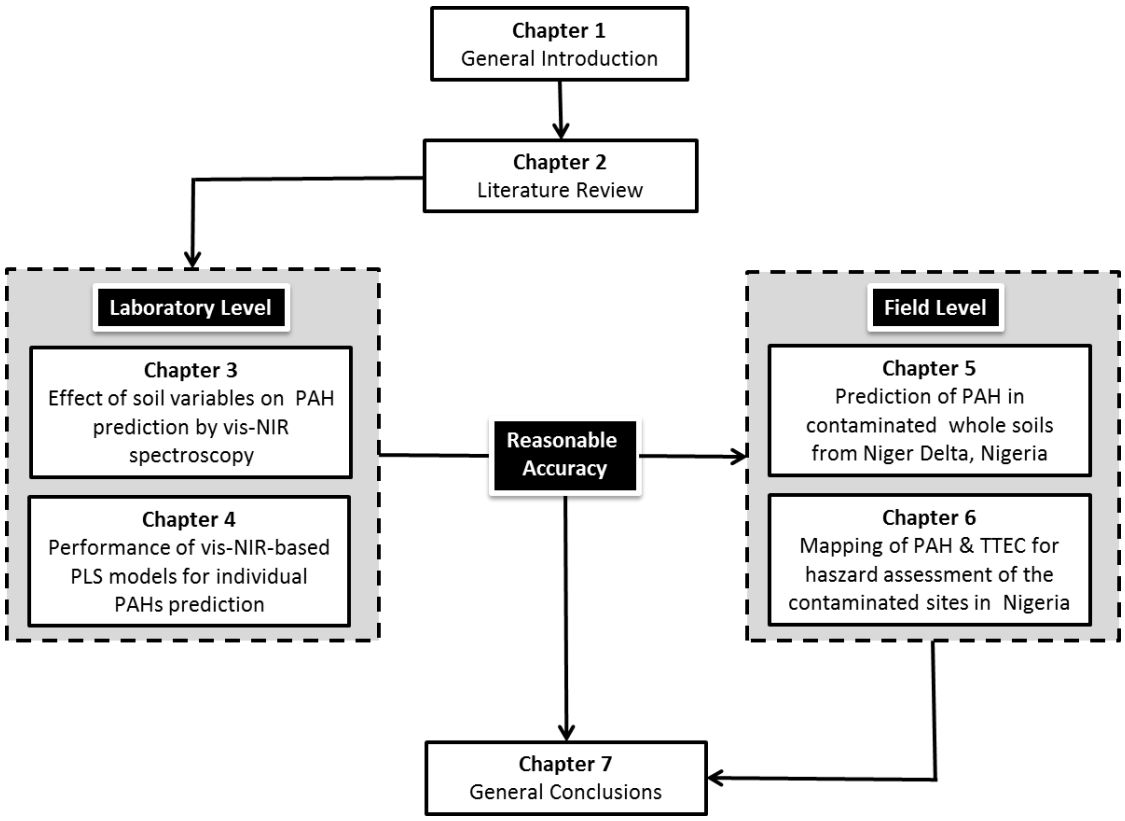


Figure 1-4 Schematic of thesis structure

The contents of the seven chapters of this thesis detailing the entire aspect of the work done in this study are summarised below:

- Chapter 1: This is the general introduction detailing the research aim and objectives, among others;
- Chapter 2: This chapter presents a critical literature review of conventional and innovative techniques for the analysis of petroleum hydrocarbons in contaminated soils, which include both spectroscopic and non-spectroscopic techniques. The content of this chapter is part of

a review paper published in Applied Spectroscopy Reviews (Okparanma and Mouazen, 2013a), which is reused in this thesis after the permission of Taylor and Francis Group (Appendix A);

- Chapter 3: In this chapter, investigations were carried out on: (1) the effects of the interaction of diesel concentration, and clay and moisture contents on the spectral responses of diesel-contaminated soil under laboratory controlled conditions, and (2) the effects of these soil variables on the performance of PLS regression models developed with vis-NIR spectroscopy for the prediction of PAHs in soils. Model quality indicators used include residual prediction deviation (RPD), root-mean-square error (RMSE), and coefficient of determination (r^2). The chapter also presents the unsupervised principal component analysis (PCA) carried out for pattern recognition latent in the entire dataset. Succinctly, a total of 150 artificially contaminated soil samples were used for the investigation. All 150 soil samples were collected from within the UK, separated into six subgroups of 25 samples each, and given various levels of treatment including oven-drying, wetting and contamination with diesel fuel. The sixth subgroup was not subjected to oven-drying and wetting. This set of samples was used as field-moist intact samples. PLS calibration models were then developed for each treatment level to study the effects of the treatments on model quality as well as soil spectral responses. The content of this chapter is part of a research paper published in Water, Air, and Soil Pollution (Okparanma and Mouazen, 2013b), which is reused in this thesis after the permission of Springer (Appendix B);

- Chapter 4: This chapter is the second part of the laboratory investigation mentioned in chapter 3. As such, the experimental setup here is basically the same as discussed in Chapter 3. But, unlike chapter 3, it examines at length the prediction accuracy of PLS models of selected individual PAH fractions including acenaphthylene, fluorene, phenanthrene, anthracene, and pyrene for various clay contents and predetermined levels of diesel and moisture. Central to both chapter 3 and 4 is the fact that the reasonable prediction accuracy of the PLS models in both instances gave an interesting insight into the likely outcome of the field studies. The outcome of the two chapters is the link between the laboratory and field investigations. The content of this chapter is part of a research paper in press in The Scientific World Journal (Okparanma and Mouazen, 2014b);
- Chapter 5: This chapter is the first part of the field investigation. The investigation carried out here is a near-onsite adaptive trial with fresh whole soil samples collected from Nigeria and taken to the UK for analysis. The study describes the development of vis-NIR spectroscopy calibration and general prediction models of PAH based on 137 fresh whole soil samples collected from three petroleum release sites in Ogoniland in the Niger Delta province of Nigeria. PLS regression analysis with full cross-validation was used for model development. Succinctly, this chapter shows part of the practical aspect of this research – moving from laboratory to the field. This follows findings of the laboratory studies in chapter 3 and 4 that vis-NIR spectroscopy can be used to predict PAHs in field-moist intact soils with reasonable accuracy.

The content of this chapter is part of a research paper published in Environmental Pollution (Okparanma et al., 2014a), which is reused in this thesis after the permission of Elsevier (Appendix C);

- Chapter 6: This chapter deals with the development of soil maps of PAH and TTEC for hazard assessment of the three petroleum release sites in Ogoniland in the Niger Delta province of Nigeria using vis-NIR spectroscopy. The experimental setup here is the same as in chapter 5. The soil maps were developed using both chemical and vis-NIR spectral data from chapter 5. Data points were interpolated using ArGIS® (ESRI, USA). In this chapter, error distribution between measured and predicted maps of PAH and TTEC are analysed using histogram plots. For fair comparison between soil maps developed by the two measurement methods, the same data classification method and interval are used. Then the maps are compared using kappa statistic. The chapter also shows how potential management zones were delineated from the TTEC soil maps using both measured and predicted data. Finally, it presents the outcome of the hazard assessment of the three petroleum release sites in Nigeria based on Generic Assessment Criteria (GAC) established in this study for the studied sites. The chapter is the concluding part of the practical application of this research demonstrating the potential of vis-NIR spectroscopy for hazard assessment of the three petroleum release sites in Ogoniland in the Niger Delta province of Nigeria. The content of this chapter is part of a follow up research paper submitted recently to Biosystems Engineering (Okparanma et al., 2014c);

- Chapter 7 is the final chapter of this thesis. This chapter provides generic conclusions for the entire work done in this study. It also highlights prospects and suggested guidelines for future works.

1.9 Significance and justification of the study

This study hopes to improve profitability in risk assessment by reducing costs and time associated with the conventional methods. Apart from costs and time factors, the choice of vis-NIR spectroscopy for this study was based on occupational health and safety considerations.

Chapter 2: Literature Review

2.1 Chapter summary

In the analysis of petroleum hydrocarbon–contaminated soils for total petroleum hydrocarbons (TPHs) and PAHs, the roles of spectroscopic and non-spectroscopic techniques are inseparable. Therefore, spectroscopic techniques cannot be discussed in isolation. In this chapter, spectroscopic techniques including Raman, fluorescence, infrared, mid-infrared and near-infrared spectroscopies, as well as mass spectroscopy (coupled to a gas chromatograph) and non-spectroscopic techniques such as gravimetry, immunoassay, and gas chromatography–flame ionization detection (GC–FID) are reviewed. To bridge the perceived gap in coverage of the quantitative applications of vis-NIR spectroscopy in the rapid determination of TPHs and PAHs in soils, a detailed review of studies from the period 1999–2012 are presented. This chapter also highlights the strengths and limitations of these techniques and evaluates their performance from the perspective of their attributes of general applicability, namely economic, portability, operational time, accuracy, and occupational health and safety considerations. Overall, the fluorescence and MIR spectroscopic techniques had the best performance (85% total score, respectively) in comparison to the others, and the GC–MS technique performed the least (55% total score). Method-specific solutions geared toward performance improvement are also suggested.[1]

[1] Reprinted after the permission of Taylor and Francis Group

2.2 Introduction

Over the years, several spectroscopic and non-spectroscopic techniques have been developed for the analysis of TPH and PAH in soil samples; the most frequently used are immunoassay, general gravimetry, laboratory-based GC–FID, GC–MS, infrared (IR) spectroscopy, mid-infrared (MIR) spectroscopy, Raman spectroscopy, and fluorescence spectroscopy. Recently, a couple of equally important innovative methods have shown reasonable potential for the measurement of TPH and PAH in oil-contaminated soils. These include field-portable GC–MS (Barnes, 2009), new-generation near-infrared analysis by vis-NIR spectroscopy (Graham, 1998; Malle and Fowlie, 1998; Malley et al., 1999; Chakraborty et al., 2010, 2012; Forrester et al., 2010; Bray et al., 2010), and MIR spectroscopy (Forrester et al., 2013). Although it was not until recently that some immunoassay techniques (Zhou et al., 2009; Wei et al., 2009), field-portable GC–MS systems (Barnes, 2009), and vis-NIR spectroscopy (Bray et al., 2010) were used to detect PAHs in soil samples. Previously, the laboratory-based GC–MS systems, fluorescence spectroscopy, and Raman spectroscopy were used for the analysis of PAH in environmental samples, but GC–MS systems are preferred because of their relative selectivity and sensitivity (Wang and Fingas, 1995, 2003; Weisman, 1998; Brassington et al., 2010).

2.3 Analytical techniques for petroleum hydrocarbons in soils

Analytical methods for petroleum hydrocarbons currently in use are numerous and it would be a Herculean task to review all of them in one study. Therefore, a selected number of these techniques are considered. These methods are

distinguishable by the level of analytical details they provide and their method of application as screening techniques, conventional nonspecific methods, and methods for detailed component analysis (Whittaker et al, 1995), which may be field and/or laboratory based. Table 2-1 summarises the most frequently used analytical methods for petroleum hydrocarbons in soils.

Table 2-1 Summary of selected spectroscopic and non-spectroscopic techniques for petroleum hydrocarbon measurement

Measurement technique	Detection device	Measured target
Gas chromatography	Flame ionization detector	TPH
	Mass selective detector	TPH, PAH, and CSB
Infrared spectroscopy	IR spectrometer/ diffuse reflectance	TPH and PAH
General gravimetry	Gravimetric balance	TPH
Immunoassay	Enzyme-linked immunosorbent assay kit	TPH and PAH
	Electrochemical immunoassay kit	TPH and PAH
Raman spectroscopy	Charge coupled detector	TPH and PAH
Fluorescence	Polychromator and CCD camera	TPH and PAH
spectroscopy	Silicon-intensified target camera	TPH and PAH
Mid-infrared spectroscopy	DTGS and InSb detectors	TPH
Visible and near-infrared	High-intensity probe/ mug lamp	TPH and PAH
spectroscopy		

Modified from and reprinted by permission of AEHS Foundation Inc., Amherst, MA. (Appendix D). CCD: charge coupled detector, CSB: compound-specific biomarkers, DTGS: deuterated triglycine sulphate.

The conventional methods include GC–FID (EPA Method 8015), GC–MS (EPA Method 8270 and 625), IR spectroscopy (EPA Method 418.1), immunoassay (EPA Method 4030 and 4035), and gravimetry (EPA Method 1664) (EPA, 1978,

1984, 1996a, b, 1999a, b, 2000). The non-conventional ones are Raman spectroscopy, fluorescence spectroscopy, MIR spectroscopy, and vis-NIR spectroscopy.

2.3.1 Laboratory-based (reference) techniques

2.3.1.1 General gravimetry

Gravimetric methods employ an initial cold solvent extraction step and a final weight-difference step. In between there may be a further clean-up step with silica gel to remove biogenic material. If it does not involve a clean-up step, it is termed an oil and grease method; if it does, it is termed a TPH method (Weisman, 1998). In the general gravimetric TPH method, EPA Method 1664 (EPA, 1999b), soil samples are uniformly graded by sieving, oven-dried at 105 ^{o}C for 12 h, and TPH compounds are eluted with n-hexane. The liquid extract (eluate) is contacted with silica gel to remove biogenic polar materials and then evaporated. The residue is retained and weighed, and the weight difference is reported as a percentage of the total soil sample on a dry weight basis. EPA method 1664 (EPA, 1999b) recommends the use of a 0.45-µm filter because of the presence of suspended solids (Weisman, 1998). Among the earliest methods developed, though obviously one of the fast declining choices (Villalobos et al., 2008), gravimetric methods have been widely used to determine TPHs in contaminated soils (Villalobos et al., 2008). Prior to the study by Villalobos et al. (2008), gravimetric methods were described as quick and inexpensive methods, but, in their recent study, the long time required for complete hexane evaporation, of not less than 60 min, “elevates the energetic

costs of the overall procedure" (p. 156) and analytical losses at higher times that cause negative errors are incurred. The latter limitation corroborates similar findings in earlier studies (Whittaker et al., 1995; Rhodes et al., 1990; White and Irvine, 1994). The extraction efficiency of gravimetric methods, albeit poor, is greatly affected by the type of eluting solvent used (EPA, 1978; Douglas et al., 1992). Hexane has poor extraction efficiency for higher molecular weight petroleum compounds (Weisman, 1998) and low polarity, which causes the co-extraction of natural organic matter containing multiple polar functional groups (EPA, 1996a; Essington, 2004). Consequently, other chlorinated compounds like chloroform (Abu and Atu, 2008) as well as toluene (Adesodun and Mbagwu, 2008) have been used as liquid extractants. It is well known that both chloroform and toluene have serious health implications as evident in the risk phrases published in their respective safety data sheets. Additionally, gravimetric methods are nonspecific because they give no information about the type of hydrocarbon present (Weisman, 1998; Villalobos et al., 2008). As a result, they are not suitable for assessing PAH compounds. Instead, the method is best suited for screening TPHs in very oily sludges or samples containing very heavy molecular-weight hydrocarbons because light hydrocarbons ($<C_{15}$) are easily volatilized at temperatures below 70 to 85 ^{o}C during the evaporation step (Weisman, 1998). Detection limits for TPHs of a ~50 $mgkg^{-1}$ in soils have been reported (Weisman, 1998). Other advantages and disadvantages of the gravimetric method are summarised in Table 2-2.

Table 2-2 Characteristics of selected petroleum hydrocarbon measurement techniques

Technique	Application method	Approx. LDL ($mgkg^{-1}$) [a]	Estimated analysis run time (min)	Sampling type	Strength	Limitations
GC-based	Laboratory and field	10	45 – 2,880 [b, c] (Excluding extraction time, lab-based GC cycle time may be 40 min, while some portable devices take <10 s)	Purge and trap, head-space, solvent extraction	Selectivity and high sensitivity; oil source identification; specific (MSD); quantitative and qualitative applications, portable, rapid	Laboratory-based GCs are not for compounds < C_6; Non-specific (FID); Expensive (capital equipment and analytical costs); Problem of co-elution; requires expertise; produces COPC
IR-based	Laboratory and field	6.32 [d] -15.2 [e]	1 [f] (Excluding sample extraction time)	Solvent extraction	Quick, simple, and inexpensive; portable	Non-specific; low sensitivity; analytical losses; poor extraction efficiency; quantitative application only; produces COPC
Gravimetric	Laboratory	50	–	Solvent extraction	Quick, simple and inexpensive	Non-specific; low sensitivity; analytical losses; poor extraction efficiency; quantitative application only; produces COPC

(Continued on next page)

Table 2-2 Characteristics of selected petroleum hydrocarbon measurement techniques (*Continued).*

Technique	Application method	Approx. LDL ($mgkg^{-1}$) [a]	Estimated analysis run time (min)	Sampling type	Strength	Limitations
Immunoassay	Field	10–500 (under laboratory conditions, 0.1 has been achieved with trained staff) [h]	1.5 – 3 [g]	Optical density	Quick, simple, inexpensive and portable; increasingly reasonable accuracy	Non-specific; low sensitivity; only measures aromatics; quantitative application; screening only, cross-reactivity
NIR-based	Field	1.0 [i]	0.6 [j] – < 2 [i] (depends on the number of scans per sample)	Diffuse reflectance/ transmittance spectra	Rapid, simple, inexpensive, portable, zero-solvent extraction, non-invasive, little or no sample preparation	Non-specific, matrix effect (water, soil & nature of oil), overlapping spectra, long pathlength, indirect correlation, qualitative applications, high-level chemometrics
Fluorescence-based	Laboratory and field	0.05 [e]	2.6 [e] (Excludes extraction time)	Emission spectra	Rapid, portable, specific, inexpensive, quantitative and qualitative applications	Sensitive to non-hydrocarbons, sensitivity is affected by soil matrix, prior sample extraction is required

(Continued on next page)

Table 2-2 Characteristics of selected petroleum hydrocarbon measurement techniques *(Continued).*

Technique	Application method	Approx. LDL ($mgkg^{-1}$) [a]	Estimated analysis run time (min)	Sampling type	Strength	Limitations
Raman Spectroscopy	Laboratory and field	–	0.1 [k] – < 3 [l]	Emission spectra	Rapid, portable, non-invasive, inexpensive, quantitative and qualitative application, specific	Laser alteration of samples, fluorescence contamination
Mid-infrared Spectroscopy	Laboratory and field	–	0.3 [f]	Diffuse reflectance/ transmittance spectra	Rapid, portable, non-invasive, quantitative and qualitative application, specific	Not suitable for wet samples

Modified from and reprinted by permission of AEHS Foundation, Inc., Amherst, MA. (Appendix D); LDL, Laboratory detection limit; COPC, Constituents of potential concern

[a] In soil.

[b] Data from Lambert et al. (2001).

[c] Data from Askari and Pollard (2005).

[d] Data from Billets (2001).

[e] Data from Greason (2009).

[f] Data from Forrester et al. (2010).

[g] Dexsil® Corporation, Hamden, CT, USA.

[h] Data from Chuang et al. (2003).

[I] Data from Malley et al. (1999).

[j] Data from Graham (1998).

[k] DeltaNu® Inc. (a subsidiary of Intevac Inc., Santa Clara, CA, USA).

[l] Data from Puppels et al. (1991).

2.3.1.2 Infrared (IR) spectroscopy

This method harnesses the spectra of the stretching and bending vibration associated with a molecule when it absorbs energy in the IR region of the electromagnetic spectrum for property elucidation (Weisman, 1998). In the IR spectrum, spectra of hydrocarbon derivatives originate from fundamental C–H stretching modes of saturated CH_2 and terminal CH_3 functional chemical groups. In the IR region, these occur around the absorption frequencies of 3000 to 2900 cm^{-1} (~3,333 to 3,448 nm) or at the specific frequency of 2930 cm^{-1} (~3,413 nm) (Weisman, 1998). Usually, as-received samples are first extracted with an eluting solvent containing no C–H bonds and the eluate is contacted with silica gel to remove biogenic polar components before being subjected to IR spectrometry. The absorbance of the eluate is then measured at the specified frequency and compared against the calibration curve developed for the instrument. The instrument calibration standard usually is a petroleum hydrocarbon of known TPH concentration (Weisman, 1998).

The primary advantage of IR-based methods is that they are quick, simple, and inexpensive with common detection limits of a ~10 $mgkg^{-1}$ in soil (Weisman, 1998; Lambert et al., 2001). Before the advent of GC-based methods, IR-based methods were frequently used to detect TPH in soils (Current and Tilotta, 1997) because they were recognized by the USEPA as an official TPH screening method (e.g., EPA Method 418.1) as well as by the International Organization for Standardization (ISO) (e.g., ISO/TR 11046) (EPA, 1978; ISO, 1992; Whittaker et al., 1995; Becker et al., 2002). But, following the ban on the use of

Freon (also known as 1,1,2-trichlorotrifluoroethane, CFE) as an extracting solvent because of its potential to deplete the ozone layer, the use of IR-based methods has plummeted over the years (Weisman, 1998). Despite the ban, a handful of studies can be found in the open literature on the use of IR-based methods (Forrester et al., 2010; Lambert et al., 2001; Becker et al., 2002). However, its use as a TPH measurement method is no longer supported by international standardization; the ISO for instance, has replaced ISO/TR 11046 with ISO/DIS 16703, which recommends the use of GC–FID after extraction with a halogen-free solvent (ISO, 1992, 2001; Douglas et al., 1992). ISO/DIS 16703 has been updated with ISO 16703 since 2004 (ISO, 2001, 2004). In addition to the limitations on its use, a major constraint of the IR-based method, according to the literature (Whittaker et al., 1995; Fan et al., 1994), is the insensitivity of the technique to unsaturated components of weathered hydrocarbons not exhibiting detectable adsorption bands at the monitoring wavelength. Additionally, the use of standard hydrocarbon mixtures different from the contaminating oil for prior equipment calibration invariably does not produce a true contaminant concentration. This is because different hydrocarbons respond differently to IR spectroscopy since single hydrocarbon oil may not be suitable as a universal calibration standard (Whittaker et al., 1995; Fan et al., 1994). The proportion of saturated and unsaturated hydrocarbon groups varies with each oil derivative and produces correspondingly variable IR spectroscopic responses (Lambert et al., 2001). The non-specificity of IR-based methods (Lambert et al., 2001) also limits their suitability for PAH assessments. IR-based methods are prone to interference,

both negative and positive biases, due to the use of dissimilar calibration standards as the spilled oil and from spurious signals due to CH_3 groups associated with nonpetroleum sources (Weisman, 1998). As stated previously, multivariate calibration solves the interference problem in general. The accuracy of IR-based techniques is dependent on the extraction efficiency of the extracting solvent, which in turn is affected by the type of solvent used (Weisman, 1998; Lambert et al., 2001). Sample porosity also has a profound influence on IR signal intensity (Forrester et al., 2010). (See also Table 2-2).

2.3.1.3 Gas chromatography–flame ionization detection (GC–FID)

The origin, principles, and techniques of chromatography have been widely documented (Sherma, 1972). Succinctly, chromatography is a separation method in which a mixture is applied as a narrow initial zone to a stationary, porous sorbent, which causes the components to undergo differential migration by the flow of the mobile phase, a liquid or a gas (Sherma, 1972). In gas chromatography, an inert carrier gas (helium, hydrogen, or nitrogen) carries the gaseous mixture (or, if aqueous, liquids with boiling points <400 ^{o}C) to be analysed through a capillary column onto a detector at the end of the column (Weisman, 1998; Sherma, 1972), which allows better resolution of components in complex mixtures.

In the GC–FID method, as-received samples are first refrigerated at 4 ^{o}C until extraction and dried either chemically (using a suitable drying agent, say anhydrous sodium sulphate) or physically in an oven at 105 ^{o}C for 24 h to remove any residual moisture. TPH compounds in the dried samples are then

extracted employing eluting solvents (e.g., acetone, dichloromethane, hexane, or pentane), and different forms of adsorbents (e.g., silica gel, alumina, or Florisil® [Fisher Scientific Ltd., Loughborough, UK]) are used for the extract clean-up and fractionation into aliphatics and aromatics (Wang and Fingas, 1995) prior to injection into a chromatographic column. Sample extracts are introduced into the capillary column by headspace, purge-and-trap (for volatile compounds in the rage C_6 to C_{25} or C_{36}), or direct injection (for the less volatile fractions) methods. As the temperature of the column is gradually raised, TPH compounds are separated according to their boiling points as they migrate toward the end of the column onto the flame ionization detector. In the detector, the high-concentration effluent eluting the column is trapped and ionized by burning in a hydrogen–air or oxygen flame, causing the gas in the detector to conduct electric current, and the conductivity is measured by a DC-powered collector electrode above the flame. The retention time of a PAH compound prior to elution from the column is typical of the species under a set of conditions and is used to correlate the detector response to the amount of compound present. The detector's responses in a given range are then integrated to give the total concentration of hydrocarbons with reference to external and/or internal hydrocarbon standards (Weisman, 1998; Sherma, 1972).

GC–FID is mostly preferred for laboratory applications because it provides relative selectivity and sensitivity (Weisman, 1998; Wang and Fingas, 1995, 2003; Brassington et al., 2010) and is recognized by the EPA, British Standard

Institution (BSI), and ISO. EPA method 8015 (EPA, 2000) is used to determine TPH, BS ISO 15009:2002 (BSI, 2002) is for volatile aromatic and halogenated hydrocarbons, and BS ISO 16703:2004 (BSI, 2004) is used to determine the content of hydrocarbons in the range C_{10} to C_{40} (n-alkanes) in solids, including soils and wastes. GC–FID is used for both quantitative and qualitative applications, including the screening of environmental samples (Vallejo et al., 2001; Snape et al., 2005; Saari et al., 2007), unravelling the type and identity of fresh to mildly weathered oil in environmental samples for pattern recognition of the petroleum hydrocarbons (Wang and Fingas, 2003; Risdon et al., 2008), and characterizing and resolving the profile of unresolved complex mixtures in petroleum-contaminated sediments (Frysinger et al., 2003). The biodegradation rate constant of petroleum hydrocarbons in a contaminated site is highly variable and difficult to evaluate due to variable site conditions. However, GC–m FID has been used to develop a simple correlation model to estimate the bioventing degradation rate constant of gasoline in several soils without having to conduct lengthy and expensive experiments (Eyvazi and Zytner, 2009). Detection limits for GC–FID depend on the method and sample matrix with typical values of a 10 mg kg^{-1} in soil (Weisman, 1998). However, high analytical costs and operational time (Askari and Pollard, 2005; Creighton and Richards, 1997), instrument calibration problems (Krupcik et al., 2004), effects of sample matrix (Saari et al., 2007), and the impact of GC operating conditions (Saari et al., 2010) are some of the challenges of the method (see also Table 2-2).

2.3.1.4 Gas chromatography–mass spectrometry (GC–MS)

Over the years, a couple of alternatives to FID (see Table 2-1) have been developed for more detailed analysis of a wider range of sample matrixes due to the selectivity of FID for hydrocarbons (Wang and Fingas, 1995). The most prominently used is the mass selective detection (MSD) technique. MSD basically uses the characteristic mass spectra of molecular and/or fragmented ions produced after ion impact to identify compounds in the sample (Masucci and Caldwell, 2004). The mass spectrometer has been described as a universal detector because of its versatility in the measurement of TPHs, PAHs, and CSBs for a wide variety of environmental samples (Poster et al., 2006) and is recommended by the EPA for the determination of both TPHs and PAHs (EPA methods 8270 and 625) (EPA, 1999a, 1978). The popular choice of a MSD technique for most environmental analysis is because of its specificity and discrete monitoring capabilities, particularly when operated in the selective-ion mode (Poster et al., 2006). As part of its wide-reaching applications, GC–MS has been used in environmental monitoring programs to assess sediment quality in terms of concentration of total PAHs (Peterson et al., 2002), to investigate the amount of PAHs in the topsoil of a tar-contaminated industrial site (Lorenzi et al., 2010), for the fingerprinting analysis of some environmental sediments containing unsaturated priority PAHs (Yang et al., 2011), and to monitor the bioremediation of PAH-contaminated soil through in-vessel composting with fresh organic wastes (Zhang et al., 2011). Despite its widespread application, a major drawback of the GC–MS is that it requires volatile and thermally stable analytes; as such, only about 10% of organics are

amenable to GC–MS analysis (Chuang et al., 2003). Moreover, the MSD is reported to have a lower sensitivity than the FID, because in the impact ion mode, the respective detectors collect and measure different proportions of the generated molecular ions (Poster et al., 2006). Quantitative chemical analysis with laboratory-based GC–MS is undoubtedly exhaustive and, like GC–FID, involves lengthy and labour-intensive extraction protocols and costly GC-based analysis, and it is uneconomical in the assessment of large-scale contamination involving dense sampling for mapping of zones requiring remediation (Peterson, 2002). More advantages and disadvantages of the GC–MS method are summarised in Table 2-2.

2.3.2 Field-based (innovative) techniques

2.3.2.1 Immunoassays (IMA)

IMA is a field-based immunochemical method in which antibodies are used to selectively bind specific petroleum constituents (Weisman, 1998). The underlying principle of the IMA methods is variable depending on the linked label used for response detection. The most prominent are enzyme-linked immunosorbent assay (ELISA), fluorescence immunoassay, and electrochemical immunoassay (ECIA).

In the ELISA method, the response of an antibody to sorb the sample analyte in relation to the enzyme-labelled analyte is determined by its optical density at sorption equilibrium. The concentration of the analyte in the sample is inversely related to the optical density of the antibody because the labelled enzyme, with

high antibody affinity, is more sensitive to the colouring agent (Weisman, 1998). Currently, there are a number of commercially available ELISA test kits including Ensys® and RaPID® assay (Strategic Diagnostics, Newark, Delaware, USA). Ensys® is less sensitive to heavier hydrocarbon components usually found in weathered oils (Lambert et al., 2001). The RaPID® assay has been used for the measurement of PAHs in soil (Chuang et al., 2003) and electrical transformer oil (Kim et al., 2001). According to a recent study, RaPID® is prone to the problem of cross-reactivity (Zhang et al., 2010). Cross-reactivity, which is the ability to respond to compounds structurally similar to the analyte, affects the specificity of PAH immunoassays and often results in biased results. This is because PAHs are a class of structurally related compounds (Zhang et al., 2010). As a result, the ELISA test kits are unsuitable for risk-based studies, which involve the assessment of PAHs in the medium. They are, however, recognized by the EPA as official screening methods for TPHs, EPA method 4030 (EPA, 1996a) and PAHs, EPA method 4035 (EPA, 1996b).

Fluorescence immunoassay is based on selective antigen–antibody binding and fluorescence labelling reagents (Weisman, 1998) and its use for the screening of aromatics in water samples is widespread (Lambert et al., 2001). Regardless of the media, it is important to note that fluorescence immunoassays also suffer the same cross-reactivity problems as the ELISA test kits; according to Zhou et al. (2009), about 15% cross-reactivities of the anti-naphthalene antibody bound to seven structurally related compounds were observed during the screening of naphthalene in water samples with fluorescence immunoassay (Zhou et al.,

2009). This problem is even more complex in real-world situations because potential cross-reactants are unlimited in number and most are seldom determinable, so all cross-reactivity values for the cross-reactants for PAH immunoassays are too difficult to determine (Zhang et al., 2010).

Redox-labelled ECIA is a direct competitive immunoassay based on surface-immobilized anti-PAH monoclonal antibody and electrocatalytic redox-labelled tracer recently developed for benzo[a]pyrene measurement (Wei et al., 2009). Although a detection limit of a 2.4 $ngmL^{-1}$ was reported, this is not sufficient for most practical applications. Additionally, cross-reactivity is not peculiar to polyclonal immunoassays but is also a major challenge for monoclonal immunoassays (Zhang et al., 2010), including redox-labelled ECIA, suggesting that ECIA also suffers the same fate as both ELISA test kits and fluorescence immunoassay. In addition, it has been reported that immunoassay test methods are affected by the soil matrix and the age of the weathered oil and that their sensitivity to hydrocarbons decreases with increasing soil clay content (Weisman, 1998). More limitations and strengths of the immunoassay method are summarised in Table 2-2.

2.3.2.2 Fluorescence spectroscopy

In fluorescence spectroscopy, analysis of sample is made possible through excitation of the molecules of the analyte by an incident short wavelength radiation. During the process of relaxation of the excited molecules to a lower energy level, long wavelength radiation is emitted in less than a microsecond. The intensity of the emitted radiation is then used to identify the analyte (State

University of New York at Oswego, 2008). Figure 2-1 shows the electronic transition energy levels. Qualitative and quantitative information about the analyte is provided by the characteristic emission spectrum produced (Aldstadt et al., 2002).

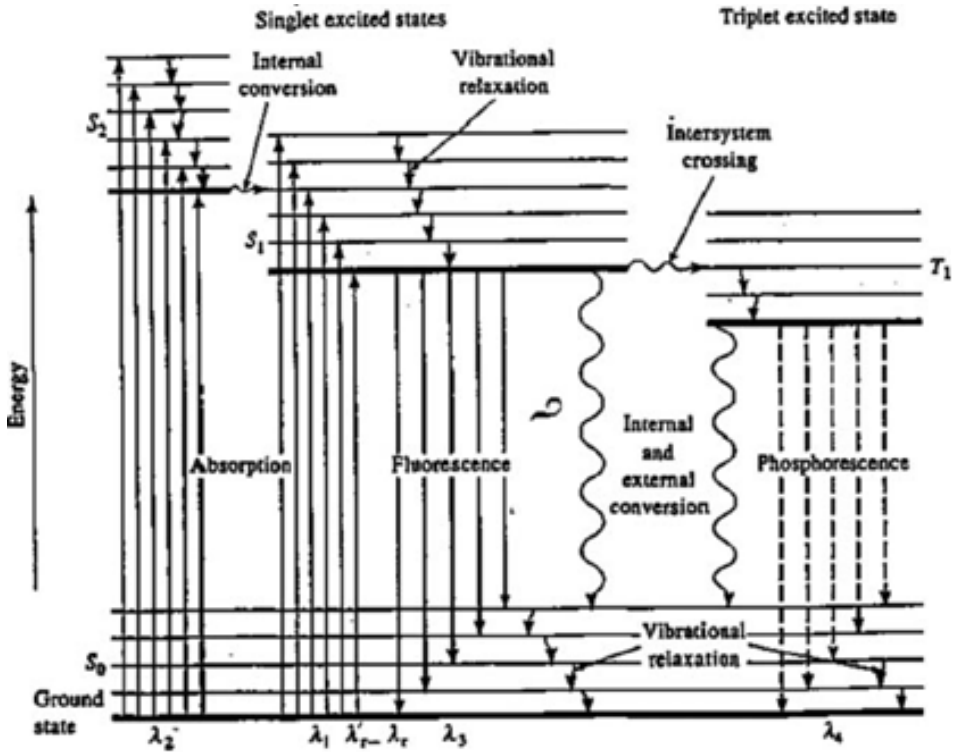


Figure 2-1 Electronic transition energy levels (Source: State University of New York at Oswego [2008])

Fluorescence spectroscopy is a field-portable technique that has been used to detect fluorescent compounds like PAHs based on the principle that the intensity of the emitted radiation is indicative of the relative concentration of the PAH as well as the number of aromatic rings. Petroleum hydrocarbon molecules absorb energy in the wavelength range 200 to 400 nm and fluoresce in the range 280 to 500 nm, with each molecule fluorescing at a specific wavelength, thereby providing the possibility of differentiating between the

various molecular classes (Quick Results on Site, 2012). Monocyclic aromatic hydrocarbons fluoresce at lower wavelengths than PAHs, though lower boiling PAHs such as naphthalene fluoresce at lower wavelengths than the higher boiling PAHs like benzo[a]pyrene (Peterson et al., 2002). Some commonly used fluorescence spectroscopic methods include ultraviolet-induced fluorescence (UVIF) and the Rapid Optical Screening Tool (ROST®, Fugro Consultants, TX, USA) laser-induced fluorescence (LIF) system.

UVIF employs a powerful UV lamp, which energizes the hydrocarbons on illumination, thereby causing them to fluoresce. The fluorescence signal is detected with a charge coupled device camera, a polychromator (or a combination of both), or a silicon-intensified target camera. Most applications utilize cone penetrometer technology (CPT). The CPT enables continuous measurement over the subsurface media of the investigation, providing semi-quantitative measurements (Askari and Pollard, 2005). In one report, the EPA reported an average analysis time of 2.6 min and a detection limit of 3.4 $mgkg^{-1}$ for field TPH measurement with a UVF-3100A device (Sitelab Corp., MA, USA) (Greason, 2009). Using the QED® hydrocarbon analyser designed by QROS (Quick Results On Site Ltd., NC, USA), even lower detection limits of 1 $mgkg^{-1}$ in soils for petroleum fuels and oils and 0.1 $mgkg^{-1}$ for PAHs in soil can be achieved in a single 5-s analysis with a throughput capacity of up to 15 samples per hour (Quick Results On Site, 2012). Currently, in North Carolina, the QED® hydrocarbon analyser has been approved as a replacement for EPA method 8015 (EPA, 2000) (based on GC–FID) for monitoring remediation of fuel spills

from leaking underground storage tanks (Greason, 2009). Despite the quantitative strength of UVIF, analysis of complex samples can be difficult due to overlap of spectra of different luminescent compounds, and prior sample extraction is required (Barnes, 2009).

The ROST® LIF system is a tunable dye laser-induced fluorescence system designed as a field screening tool for detecting petroleum hydrocarbons in the subsurface (Bujewski and Rutherford, 1997). Unlike UVIF, the ROST® LIF system uses a pulsed laser instead of UV light to cause fluorescence in PAH compounds. The laser is transmitted through a truck-mounted CPT probe (housing a sapphire window) via excitation and emission optical fibres that are pushed into the ground (Bujewski and Rutherford, 1997). Available TPH data showed that the ROST® LIF system can achieve an accuracy of 89.2% with false negatives and positives of 5.4% and a limit of detection of a 5 $mgkg^{-1}$ in soil (Bujewski and Rutherford, 1997). However, using partial least squares (PLS) regression analysis, Aldstadt et al. (2002) reported a sufficiently close match between predicted and measured PAHs for the technique to be used as a screening tool for all but 3 (i.e., acenaphthylene, dibenzo[a,h]anthracene, and naphthalene) of the 16 priority PAHs. The ROST® LIF system is designed for qualitative applications, because it can only detect the presence or absence or relative concentration of contaminants and is sensitive to non-hydrocarbon compounds in the soil, and its sensitivity is affected by the soil matrix (Barnes, 2009). Other advantages and disadvantages of the fluorescence method are summarised in Table 2-2.

2.3.2.3 Field-portable GC–MS

Prompted by the need to reduce costly delays associated with laboratory-based GC systems, portable versions are emerging. Currently, a variety of portable GC–MS systems exist, including (among others) CT-1128 GC–MS (Constellation Technology Corp., FL, USA), HAPSITE (INFICON, NY, USA), and EM 640 (Bruker-Franzen Analytical Systems, MA, USA). Currently, the EM 640 system has been replaced with the E^2M GC–MS system (Bruker-Daltonik, Leipzig, Germany). The reported average weight of the portable GC systems is between 16 and 60 kg and typical analysis run time is about 10 min for some models (Harris, 2003). Field-portable GC–MS systems obviously differ from their laboratory-based counterparts in terms of provision of real-time quantification. However, they require prior sample extraction, on-site carrier gas, considerable electrical power, and ancillary equipment, just like laboratory-based GC systems (Barnes, 2009). Some analysts have also observed that the major problem with existing portable GC and GC–MS instruments, especially microchip GCs, is sensitivity (Harris, 2003). A trade-off appears to exist between the size and performance of the GC–MS instruments such that the smaller the portable GC or GC–MS instrument is, the greater the sacrifice in sensitivity, separating power, and identifying power (Harris, 2003). Thus, according to Harris (2003), there will always be a place for the bench top instrument for routine high-throughput, high-volume analysis. More on the merits and demerits of GC–MS systems are summarised in Table 2-2.

2.3.2.4 Raman Spectroscopy

Raman spectroscopy is a vibrational spectroscopy based on the inelastic scattering (anti-Stokes) of a monochromatic light source, usually from a laser source, used in assessing the vibration of Raman-active molecules such as PAHs and for identifying species. Figure 2-2 shows the energy level diagram for Raman scattering process. The Raman signal is detected with a charge coupled detector camera.

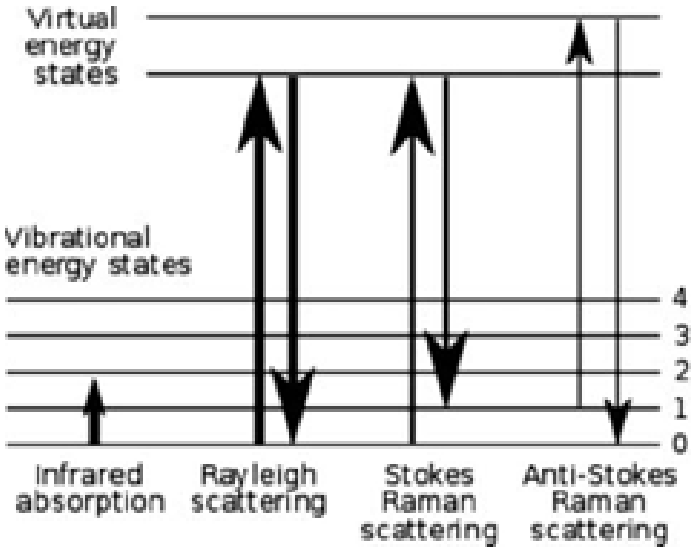


Figure 2-2 Energy level diagram for Raman scattering process (Source: Wikipedia [2012]. The line thickness is roughly proportional to the signal strength from the different transitions).

Applications of Raman spectroscopy have been widely reviewed (Lyon et al., 1998), and Raman spectral data for several PAHs have also been widely reported (Maddams et al., 1990). In general, Raman spectroscopy is a rapid measurement tool that has made it possible to obtain high-quality spectra on a timescale that is economical for analytical work (Maddams et al., 1990). It suffices to say that Raman is both qualitative (Li and Dai, 2012) and quantitative

(Pfannkuche et al., 2012) in its application. However, extreme care is needed to avoid laser alteration of samples (Jehlicka et al., 2005), and fluorescence contamination is often a problem with some Raman-based systems (Pfannkuche et al., 2012). Moreover, in all Raman instruments, noise is present to some extent and is a limiting factor in detection because it defines the detection limit of a particular compound (Brown, 2010). It has also been observed that although the miniaturization of the Raman instrument enhances convenience, this has often come at the price of sensitivity; the spectral range and spectral resolution are sacrificed with possible negative consequences to material identification and verification (Brown, 2010). More information about the strengths and limitations of Raman systems are summarised in Table 2-2.

2.3.2.5 Mid-infrared spectroscopy

The mid-infrared (MIR) region in the electromagnetic spectrum is normally defined as the wavelength ranges of 2,500 to 25,000 nm. MIR spectroscopy is a vibrational spectroscopy that relies on the measurement of light absorption due to fundamental molecular vibrations occurring in the MIR range for identification of substances (Willey, 1976; Fuller and Griffiths, 1978). Generally, measurement of light absorption in the IR range is achieved either by the 'dispersive' method, which uses a monochromatic light source, or Fourier Transform (FT) method, which uses a broadband source and manipulated by an interferometer (Murray, 2012). FTIR method is faster, has greater sensitivity and precision than the 'dispersive' method (Murray, 2012). Today, almost all IR devices are FTIR based. Consequently, MIR spectroscopy is now almost

always referred to as FTIR (Murray, 2012). FTIR devices employ various measurement modes for sample identification including transmission for liquid analysis; attenuated total reflection for solids, liquids, pastes, and gels analysis; and diffuse reflectance for solids such as soils, rocks, and minerals (Murray, 2012). Diffuse reflectance in the IR spectrum came to limelight in 1976, and became known as diffuse reflectance infrared Fourier transform (DRIFT) spectroscopy (Willey, 1976). Early applications of DRIFT have been widely reported (e.g., Fuller and Griffiths, 1978). In the analysis of soil properties, DRIFT has been used both for qualitative (e.g., Nguyen et al., 1991) and quantitative purposes (e.g., Janik et al., 1995; Janik and Skjemstad, 1995).

More recently, DRIFT in conjunction with partial least-squares (PLS) regression analysis has been used to predict TPH in hydrocarbon contaminated soils, demonstrating a cross-validation root-mean-square error of 903 mg kg^{-1} for TPH ranging from 0–11000 mg kg^{-1} (Forrester, 2010). In FTIR, absorption of energy by hydrocarbon derivatives is due to fundamental vibrational C–H stretching modes of terminal CH_3 and saturated CH_2 functional chemical groups linked to TPH (Forrester et al., 2010, 2013). Apart from portability, FTIR spectra are relatively simple to interpret because FTIR has high molecular selectivity and can identify low-level contaminant concentration, which is particularly an advantage in identifying unknown samples (Murray, 2012). Even then, it has been identified that water and sample heterogeneity are major problems with FTIR instruments (Murray, 2012, Soriano-Disla et al., 2014). The presence of water in the sample may cause the strong MIR absorption of FTIR to swamp

other more useful information (Murray, 2012). Therefore, FTIR instruments may not be the best method for analysing natural heterogeneous products including soil (Murray, 2012). More information about the strengths and limitations of MIR systems are summarised in Table 2-2.

2.3.2.6 Visible and near-infrared spectroscopy

For the purpose of this study, only the quantitative cases regarding measurement of TPHs and PAHs in soils are reviewed. A comprehensive review of non-invasive NIR spectroscopy involving several other processes and matrixes can be found in Workman (1999) and Schwartz et al. (2011).

Toward the end of the 1980s, the spectral characteristics of hydrocarbons were reported by Cloutis (1989). Following the work by Cloutis (1989), by the mid-1990s a fibre-optic NIR reflectance sensor for the detection of organics such as benzene and toluene in soils has been developed and tested (Schneider et al., 1995). Shortly afterward, a small-scale study was initiated to investigate the applicability of reflectance spectroscopy (1,600–1,900 nm) on sandy loam artificially contaminated with motor oil (Stallard et al., 1996). Two years later, a relatively more comprehensive study was conducted involving three soil types artificially contaminated with diesel and gasoline with reported minimum detection limits of 0.1 and 0.5% by weight, respectively (Zwanziger and Foster, 1998). Away from the practice of using laboratory-constructed samples, Malley et al. (1999) used NIR reflectance spectroscopy (1,100–2,498 nm) and stepwise multiple linear regression to predict concentrations of TPH in diesel-contaminated soils collected from the field with low accuracy and high prediction

error for the first time. This low performance was attributed to the small number of sample sets used and the inconsistency in the reference laboratory results (among others). A decade later, using different calibration models, Chakraborty et al. (2010) reported a fair TPH validation r^2 for field-collected intact soils from oil spill sites with vis-NIR reflectance spectroscopy. They acknowledged that a small number of sample sets resulted in the failure to develop robust calibration models. The possibility of using vis-NIR spectroscopy (400–2,500 nm) in reflectance mode as a rapid measurement tool for TPHs in crude oil and diesel-spiked soil minerals (kaolinite, illite, smectite, carbonate, and quartz) has also been reported by Forrester et al. (2010) with relatively high calibration errors for some of the minerals. Bray et al. (2010) used an ordinal logistic regression technique for total PAHs and benzo[a]pyrene predictions using vis-NIR spectroscopy with good accuracy and moderate to high false-positive rate at low and high total PAH threshold, respectively. These results were attributed to a lack of samples. In 2012, two studies were reported utilizing laboratory-constructed hydrocarbon-contaminated soil samples. Using several multivariate techniques and a slightly higher number of sample sets, Chakraborty et al. (2012) predicted with significantly improved accuracy the amount of petroleum contamination in soil samples using a vis-NIR spectroscopic method. Schwartz et al. (2012) employed several petroleum hydrocarbons (PHCs) for the simulated contamination of a total of 750 soil samples and used a vis-NIR spectroscopic method to predict their TPH levels using PLS with unsatisfactory accuracy but good correlation results; a development they also attributed to the confirmed inter- and intra-laboratory inconsistencies in reference TPH results.

Within the first quarter of 2013, one study predicted PAHs in sets of artificially contaminated soils with vis-NIR diffuse reflectance spectroscopy with reasonably high prediction accuracy (Okparanma and Mouazen, 2013b). This was the second study on PAH to be reported in two years after Bray et al. (2010).

It must be pointed out, however, that the interaction of NIR radiation with a soil sample produces soil spectra with fewer absorption features due to weak vibrational modes of molecular functional groups and broad and overlapping bands, which make NIR spectra difficult to interpret (Stenberg, 2010). As discussed, this difficulty has largely been overcome by the use of advanced chemometrics and data-processing techniques (Pasquini, 2003). Again, the long pathlength of probe may decrease the resolution and accuracy of the vis-NIR spectroscopic method. Nevertheless, intensive research is being conducted to improve the accuracy of vis-NIR spectroscopy.

2.4 Integration, analysis, and discussion

When comparing technologies for the analysis of TPHs and/or PAHs in contaminated soils, several important attributes of general applicability can be used. Here, the attributes discussed include economic considerations, operational time, occupational health and safety, portability, accuracy, and precision.

2.4.1 Economic considerations

The current capital equipment costs of different TPH and/or PAH analytical devices are shown in Table 2-3, which shows that the laboratory-based GC systems appear to be the most expensive probably due to their size and sophistication. Although the capital equipment cost of the portable EM 640 system (now replaced with E^2M system – Bruker-Daltonics GmbH, Germany) is relatively on the high side, the capital equipment cost of the portable zNose 4200 or portable 3000 Micro GC 1-, 2-channel systems is comparatively low (Table 2-3). The capital cost of the NIR-based systems is much lower than that of laboratory-based GC systems. It is about the same as the portable MIR-based system, ExoScan 4100 (Agilent Technologies, CA, USA), but higher than those of immunoassay, fluorescence, portable Raman spectroscopic, and GC–MS systems (Table 2-3). In terms of analytical cost, a standard PAH or TPH analysis by GC-based method currently costs ~$157 per sample, whereas the cost by IR-based methods is ~$51 per sample in a commercial laboratory in Nigeria, which is equivalent to the cost for a similar analysis in most developed countries like in the UK in 2005 (Askari and Pollard, 2005). Currently, it is possible that this cost may have increased in the UK considering the present economic realities. Analysts have always attributed this to the constantly increasing running costs of GC-based systems.

Table 2-3 Capital cost of selected total petroleum and polycyclic aromatic hydrocarbon analytical devices (as of 2012)

Technique	Make	Model	Price ($) (approx.) [a]
Gas chromatography	Agilent Technologies	6890 N GC–FID	28,000 (used)
	Agilent Technologies	GC-MS 6890/5975	39,900 (used)
	Thermo Fisher	LTQ Orbitrap	362,486
	Agilent Technologies	3000 Micro GC 1-,2-channel Systems (portable)	24,000–57,000
	Electronic Sensor Technology	zNose 4200 Portable Ultra-Fast GC–SAW	28,115–43,221
	Bruker-Daltonics	EM 640 (portable, bench-top)	157,900–185,000
		E^{2}M (now replaces EM 640)	–
Infrared spectroscopy	Shimadzu	IR400	3,528
Mid-infrared spectroscopy	Thermo Fisher	370 DTGS	7,055
	Thermo Scientific	Nicolet 6700 FTIR	14,900 (used)
	Agilent Technologies	4100 ExoScan FTIR (diffuse reflectance head)	43,497
	Agilent Technologies	4100 ExoScan FTIR (universal system)	61,301
Immunoassay	Pierce	Easy Titer ELISA Systems	546
Visible and near-infrared spectroscopy	Analytical Spectral Devices	Quality-SpecPro	48,563
		LabSpec 2500	56,688
		LabSpec 5000	61,078
Fluorescence spectroscopy	Hitachi	F-4010	2,500
	Tecan	Spectra Fluor	8,200
	Perkin Elmer	LS5B	4,500
Raman spectroscopy	—	Raman-HR-TEC-IG	17,777
	—	Raman-HR-TEC	5,295
	—	Raman-HR	4,000
	—	Raman-SR	3,500

[a] LabX, ON, Canada; Analytik Ltd., CB, UK; O. I. Analytical Corp. TX, USA; StellarNet Inc., FL, USA; TechMondial Ltd., London, UK.

It is known that the analytical costs for most alternative methods, such as fluorescence spectroscopy-based systems, increase as the number of samples analysed decreases as a result of the spread of the initial capital equipment cost across the number of samples in contrast to the fixed cost per sample of

reference laboratory methods (Billets, 2001). Consequently, the increase in the analytical cost per sample for such methods is not a fair comparison to the reference laboratory methods (Billets, 2001). The analytical costs of other methods like Raman and vis-NIR spectroscopic methods are not well documented at the moment due to the absence of well-developed and standardized operating protocols and cannot be fairly compared.

2.4.2 Operational time

Cycle time is the time it takes the analytical system to go from one analysis to the next (Harris, 2003); analysis run time is the sum of the cycle time and the time spent preparing the sample for analysis. Sample preparation has both occupational health (see subsection 2.4.3) and economic implications. During microcosm studies, experience shows that costly delays can result from overhead expenses due to high analysis run times. As shown in Table 2-2, high analysis run times are associated with techniques requiring lengthy initial sample preparation, particularly laboratory-based GC techniques, suggesting that the methods are uneconomical in the assessment of large-scale contamination involving dense sampling, as has been previously reported (Peterson et al., 2012). Portable GC systems can compare favourably with non-invasive devices in terms of cycle time (Table 2-2) but still involve time-consuming sample extraction protocols. Although current data for gravimetric methods were not available for comparison, the analyses run times for laboratory-based IR, fluorescence spectroscopic, and sorption-based immunoassay methods, which also require sample preparation, are comparable

to those of non-invasive Raman and NIR spectroscopic techniques. Table 2-2 shows that the portable MIR, Raman, and NIR spectroscopic techniques tend to have shorter analysis run times, because little or no sample preparation is required, compared to those methods that depend on prior sample extraction. The shorter analysis run times of the portable MIR, Raman, and NIR spectroscopic techniques make them potentially better techniques for cost-effective assessment of large-scale contamination involving dense sampling, rapid decision making, and accurate contaminant mapping.

2.4.3 Occupational health and safety

As stated in subsection 2.4.2, there are occupational health and safety concerns associated with sample preparation. Preparing samples for TPH and/or PAH analysis using the gravimetric method (Abu and Atu, 2008; Adesodun and Mbagwu, 2008), immunoassay ELISA test kits (Chuang et al., 2003; Nording et al., 2006), laboratory-based IR spectroscopy (Becker et al., 2002), fluorescence spectroscopy (Greason, 2009), and GC-based methods (Risdon et al., 2008) obviously involves handling of hydrocarbon-contaminated soil samples and noxious chlorinated extraction solvents, which exposes the user to potentially biological and chemical hazards. For instance, the solvent tetrachloroethylene, often used as a substitute for Freon 113 for extracting TPH compounds for analysis with laboratory-based IR spectroscopy, is a potential carcinogen according to the classification of the International Agency for Research on Cancer and is also a central nervous system depressant, which finds its way into the human body through inhalation and skin contact

(Wikipedia, 2012). Similarly, dichloromethane, a liquid extractant for TPH and PAH analysis by GC–FID or GC–MS, and chloroform and toluene, liquid extractants for TPH analysis by the gravimetric method, all have serious health implications, as evident in the risk phrases contained in their respective safety data sheets. Other equally dangerous chemicals (liquid or powder), which come in sealed bottles or bags that the operator must be exposed to when the containers are opened up to prepare standard solution mixes for GC-based analysis, are PAH and TPH calibration standards, internal standards such as deuterated alkanes and PAHs, and surrogate spike standards such as squalane, o-terphenyl, 2-fluorobiphenyl, heptamethylnonane, and so forth.

On the other hand, portable MIR devices such as the Agilent 4100 ExoScan FTIR (Agilent Technologies, CA, USA) (Forrester et al., 2010), vis-NIR devices, fluorescence devices such as the UVF 3100 (Sitelab Corporation, MA, USA) (Greason, 2009), and Raman devices employ electromagnetic radiation for property elucidation. In Raman spectroscopy, the electromagnetic radiation is capable of penetrating glass and plastic containers, enabling in-vessel analysis, thereby eliminating or reducing possible exposure to hazardous chemicals (GE Security Inc., 2006), but operators may still face possible exposure to fugitive lasers necessitating a Class 3B laser warning on the devices to comply with regulations. Though vis-NIR and portable MIR devices employ particle physics in their applications, which eliminates the need for a liquid extractant and concomitant exposure to chemical hazards, there is still a need, according to their data sheets, for a close proximity between the sample and the detectors

for better instrument sensitivity. This implies that contact with the contaminated soil sample and exposure to both biological and unknown hazards cannot be completely eliminated. In the same vein, fluorescence spectroscopy employing a charge coupled detector camera may still pose a radiation risk to the operator in addition to exposure to different proprietary extraction and calibration solvents often used during sample preparation.

2.4.4 Portability

Considering recent innovations, it appears that there is no clear-cut distinction between field-and laboratory-based techniques. Some techniques such as those based on MIR, Raman, and fluorescence spectroscopies and GC-based techniques, which are laboratory based, are now also available in manageable sizes that can be deployed for field measurements and are now more user friendly (Table 2-2). For example, with the introduction of portable GC–MS devices and portable MIR systems such as the Agilent 4100 ExoScan FTIR (Forrester et al., 2010), difficult circumstances can now be accessed in situ with relative ease. To the best of our knowledge, with the exception of gravimetric methods, virtually all of the techniques now come in field-implementable devices essentially for convenience and to reduce the cost of moving the equipment to and from the site. However, it has been reported that miniaturization of GC–MS (Harris, 2003) and Raman (Brown, 2010) devices results in loss of instrument performance. However, we were not able to ascertain whether this problem also affects the MIR- and immunoassay-based techniques or fluorescence and vis-NIR spectroscopic techniques. Field-

portable devices in general, in addition to being cost-effective and easy to use, are meant to reduce the amount of time spent on conventional laboratory-based analysis, expediting the screening of contaminated sites to pave the way, if need be, for more detailed investigation (Whittaker et al., 1995).

2.4.5 Accuracy

Table 2-4 shows the accuracy of selected innovative measurement techniques for petroleum hydrocarbons in contaminated soils. More often than not, the performance of standard analytical techniques is used as a benchmark for the innovative techniques because they are assumed to be more accurate than the innovative methods. Consequently, innovative techniques are often seen to play a complementary role to standard analytical techniques, because their data eventually will have to be verified by the relevant standard analytical techniques such as GC, IR spectroscopy, and general gravimetry. The accuracy of an innovative analytical technique relative to a standard analytical method can be determined using different indicators (Naes et al., 2002). Although the root mean square error (RMSE) is commonly used to estimate the performance of NIR and MIR techniques, the RMSE could not be used here as a basis for comparison because this was not reported while estimating the accuracy of other innovative methods discussed. But the value of coefficient of determination, r^2, appears to be the common denominator among other indicators reported by previous researchers (see Table 2-4). Therefore, it was used to fairly compare how well or poorly a particular innovative technique has been able to predict TPH and/or PAH values relative to a chosen standard

analytical method. For a fair comparison, average values of r^2 were used, and for the vis-NIR technique, only validated values, r^2_p, were used (see **Error! Reference source not found.**). We assumed that an r^2 value between 0.9 and 1.0 provided an excellent correlation, an r^2 value between 0.8 and 0.89 a good correlation, an r^2 value between 0.7 and 0.79 a fair correlation, and an r^2 value between 0.5 and 0.69 a poor correlation. Average values of r^2 deduced show that the fluorescence spectroscopic method performed best with the highest average r^2 value of 0.95, which implies an excellent correlation and suggests that this technique is best to assess TPHs and/or PAHs compared to the other innovative techniques. This is followed successively by the FTIR method (average r^2 = 0.87), and immunoassay-based and the vis-NIR methods (average r^2 = 0.81, respectively), implying good performance. Raman spectroscopic methods could not be fairly compared because the standard analytical method used in the studies we reviewed was either not stated or was high-performance liquid chromatography, which is outside the scope of this review.

2.4.6 Precision

Precision of an analytical instrument is a measure of reproducibility of analytical data. This is usually achieved by repeat analysis of a sample aliquots and estimating the relative standard deviation or coefficient of variation expressed as a percentage (Patnaik, 1997). In routine environmental analysis by the standard analytical techniques such as GC–MS where several repeat analyses is often not possible, precision is estimated by the relative percent difference of

the duplicate analysis (Patnaik, 1997) (see Chapter 5:). However, it is assumed that the standard analytical techniques are precise since they are used here as reference methods. In the analysis of spectroscopic data, the standard error of prediction (SEP) is used as a measure of precision of a prediction (Naes et al., 2002). SEP is defined as the standard deviation of the predicted residuals or the mean difference between the predicted and measured reference values in the prediction set (Naes et al., 2002). In NIR analysis of PHCs in soils, SEP is rarely reported probably because of the absence of a standardised operating protocol. Moreover, studies on the application of NIR method for the detection of PHCs in soils are few in literature at the moment. Nonetheless, we managed to find three studies on vis-NIR spectroscopy that reported SEP (Stallard et al., 1996; Zwanziger and Foster, 1998; Malley et al., 1999) (Table 2-4). With the reported low SEP (Table 2-4), it shows that vis-NIR is capable of high precision determination. On the other hand, it can be seen in Table 2-4 that studies that employed IMA and fluorescence techniques to determine PHCs in soils used relative standard deviation (RSD) as a measure of precision. Therefore, a common indicator could not be found to fairly compare the precision of the innovative techniques for PHCs in soils. This is the reason precision was not included in the attribute comparison table in subsection 2.5 as this would result in incomplete score for all techniques just as accuracy did for Raman spectroscopy.

Table 2-4 Accuracy of selected innovative measurement techniques for petroleum hydrocarbons in contaminated soils

Innovative method	Reference method	Deployed	Measured target	Multivariate approach	No. of samples	Spectral pre-processing	Wavelength range (nm)	Statistical parameters	Reference
Immunoassay-based	GC–MS	Field	PAH	–	52	–	–	r^2 (0.61–0.68) RSD (0.3–55 %) *p*-value (<0.001)	Chuang et al. (2003)
	GC–MS	✓	PAH	–	11	–	–	r^2 (0.94–0.99)	Nording et al. (2006)
Fluorescence-based	GC–FID	Field	PAH	–	595	–	254	r^2 (0.90–0.97) (For GRO→EDRO)	Greason (2009)
	IR–S	✓	TPH	–	30	–	370–524	RSD (5–124 %)	Lohmannsroben and Roch (2000)
	GC–MS	✓	PAH	–	08	–	–	r^2 (0.92)	Sitelab Corp. (2010)
	GC–MS	✓	PAH	–	Not stated	–	266	r^2 (0.997)	Schultze and Lewitzka (2005)
FTIR	Not stated	✓	TPH	PLSR	172	Not reported	1280–22000	r^2_{cv} (0.81) RMSECV (4500–8000 mg kg^{-1})	Forrester et al. (2010)
FTIR	GC–FID	✓	TPH	PLSR	205	Mean value centring, Linear Detrending	1280–22000	RPD (3.7) r^2 (0.93) RMSECV (564 mg kg^{-1})	Forrester et al. (2013)
Vis–NIR spectroscopy	Not stated	Laboratory and Field	Oil and Grease	PLSR	17	–	1600–1900	SEP (0.13–0.26 %) (CV)	Stallard et al. (1996)

Continued on next page

Table 2-4 Accuracy of selected innovative measurement techniques for petroleum hydrocarbons in contaminated soils *(Continued)*

Innovative method	Reference method	Deployed	Measured target	Multivariate approach	No. of samples	Spectral pre-processing	Wavelength range (nm)	Statistical parameter	Reference
Vis-NIR spectroscopy	Not stated	Laboratory and Field	Oil and Grease	PLSR	>25	Kubelka-Munk transformation, Saunderson correction, Mean value centring, MSC	800–2700	r^2_{cv} (0.968–0.998) SEP (0.116–1.04) (CV) Bias (-0.001–0.586) RMSD (0.106–0.948) (CV) (Diesel data only)	Zwanziger and Foster (1998)
	GC–FID	✓	TPH	Stepwise MLR	26	Wavelength average, first derivative, and smoothing splines	1100–2498	r^2_{p} (0.68–0.72) SEP (0.84–1.00) RPD (1.76–1.82)	Malley et al. (1999)
	Gravimetric	Laboratory and Field	TPH	PLSR, BRT	46	Parabolic splice, wavelength average, first derivative, second derivative, smoothing splines	350–2500	r^2_{cv} (0.64–0.85) r^2_{p} (0.42–0.64) RMSEP (0.335–0.589) RMSECV (0.303–0.436) RPD (1.35–1.94) Bias (-0.07–-0.14)	Chakraborty et al. (2010)

Continued on next page

Table 2-4 Accuracy of selected innovative measurement techniques for petroleum hydrocarbons in contaminated soils *(Continued)*

Innovative method	Reference method	Deployed	Measured target	Multivariate approach	No. of samples	Spectral pre-processing	Wavelength range (nm)	Statistical parameter	Reference
Vis-NIR spectroscopy	AAS	✓	PAH	OLR	65	Not reported	350–2500	Accuracy (65.85–90.25 %) FPR (0.57–0.91) FNR (0.03–0.13)	Bray et al. (2010)
	Not stated	Laboratory and Field	TPH	PLSR, MLR, Penalised Spline	68	First derivative, discrete wavelet transform	350–2500	r^2_{cv} (0.84–0.98) RMSECV (3010-4791) RPD (2.50–3.97)	Chakraborty et al. (2012)
	IR–S	✓	TPH	PLSR, ANN	750	SNV, MSC, smoothing, first derivative, second derivative, continuum removal	Not stated	r^2 (0.79–0.99) AVG delta (325–6187) AVG dev. (47–68 %) Max. delta (747–11494) Max. dev. (74–143 %)	Schwartz et al. (2012)
	GC–MS	✓	PAH	PLSR	150	Noise cut, wavelength average, maximum normalisation, first derivative, smoothing, baseline correction	350–2500	r^2_{cv} (0.56–0.86) r^2_{p} (0.89) RMSEP (0.20) RMSECV (0.14–0.48) RPD (1.52–2.79) (CV), 2.75(P)	Okparanma and Mouazen (2013b)

AAS, Atomic Absorption Spectrometry, IR-S = infrared spectrometry; FS = fluorescence spectroscopy; MLR = multiple linear regression; BRT = boosted regression tree; OLR = ordinal logistic regression; FTIR = Fourier transform infrared spectroscopy; SNV = standard normal variate; MSC = multiplicative scatter correction; RSD = relative standard deviation; GRO = gasoline range organics; EDRO = extended diesel range organics; SEP = standard error of prediction; CV = cross-validation; RMSD = root mean square deviation; RPD = residual prediction deviation; RMSEP = root mean square error of prediction; RMSECV = root mean square error of cross-validation; FPR = false-positive rate; FNR = false-negative rate; P = prediction.

2.5 Overall performance and method-specific recommendations

Table 2-5 shows a comparison of attributes of general applicability for selected analytical methods for TPH and PAH in soil. Overall, the table shows that most of the analytical techniques compare very well, with fluorescence and FTIR spectroscopies topping the table, whereas the GC-based method is at the bottom. Fluorescence spectroscopy has excellent records for 60% of the attributes but a good operational time and a fair occupational health and safety records (Table 2-5), which suggests that improvements should be geared toward reducing or completely eliminating the need for prior sample preparation.

Immunoassay evidently is economical and portable and has good accuracy but has only a fair record for operational time and health and safety issues (Table 2-5) due to its dependence on solvent extraction. Like every other solvent extraction–dependent method, immunoassay would require (among others) an operating protocol that will be less dependent on, or independent of, chlorinated extraction solvents. Although the GC-based method has 55% of the total attributes, its high accuracy is well depicted in Table 2-5. This low overall score obviously correlates to the effects of the poor health and safety records and the fair economic, operational time, and portability factors. The introduction of the portable GC–MS systems obviously did not have much of a positive impact on the mobility of the method as the system is still the bench-top type but is light enough to be moved to site. Similarly, the portable GC–MS devices did not have much of a positive impact on the economic (based on capital equipment

cost), operational time, and health and safety concerns. This is because the cost of most portable GC–MS devices is still relatively high (Table 2-5), and the method is still reliant on time-consuming extraction protocols and the use of extraction solvents. Therefore, operating protocols that encourage the use of non-chlorinated solvents without compromising analyte recovery may well improve on its health and safety impacts, and an instrument design that eliminates the need for prior sample extraction will definitely enhance timeliness. The gravimetric method is cheap and has an assumed excellent accuracy record, because it is assumed that standard methods are reference frames against which innovative methods are evaluated. However, gravimetry is a laboratory-based method with poor health and safety records and a fair operational time (Table 2-5) due to the extraction step involved. Therefore, the gravimetric method would require the same improvement recommended for other extraction-dependent analytical methods.

The FTIR spectroscopy outperformed vis-NIR spectroscopy on the overall score sheet (Table 2-5) mainly because of higher accuracy records. The significantly improved records for FTIR spectroscopy may be attributed to the enhanced accuracy and portability of the method brought about by the introduction of portable FTIR devices. FTIR devices would also require improvements on their accuracy. On the other hand, vis-NIR spectroscopy has excellent records in terms of portability and operational time and good records from economic and health and safety standpoints (Table 2-5). However, vis-NIR spectroscopy has only a fair accuracy record, which appears to be one of the lowest among other

methods (Table 2-5). This is because NIR prediction is based on overtones and combinations of fundamental vibrations occurring in the MIR region. Therefore, significant improvement in the accuracy of vis-NIR spectroscopy is required. Possible solutions may include the use of nonlinear multivariate analytical methods such as artificial neural networks (ANNs) as an alternative to PLS regression analysis and understanding the effects of other variables (e.g., soil texture, moisture content, and oil concentration) that affect the soil diffuse reflectance spectra of contaminated soils. Soil properties have been predicted with higher accuracy using an ANN than with PLS regression analysis using vis-NIR spectroscopy (Mouazen et al., 2010). Mouazen et al. (2006a, 2010) have reported that soil texture and moisture content negatively affect the performance of NIR spectroscopy in the prediction of soil chemical properties. By sufficiently understanding the individual and combined effects of these three factors, moisture content, clay content, and oil concentration, it is expected that techniques to reduce or even eliminate the effects of these factors and thus lead to improved accuracy of calibration models developed to predict TPH and/or PAHs in petroleum hydrocarbon–contaminated soils will be established. Laboratory- and field-scale studies to address these key areas are being undertaken by us, the authors.

Table 2-5 Attributes of general applicability for selected analytical methods for total petroleum and polycyclic aromatic hydrocarbons in soil

Analytical technique for petroleum hydrocarbon	Attribute of general applicability					Overall score (%)
	Economic considerations [a]	Operational time	Occupational health and safety	Portability	Accuracy [b]	
Gas chromatography	××	××	×	××	××××	55
Infrared spectroscopy	×××	×××	××	×	××××	65
Gravimetric	××××	××	×	×	××××	60
Immunoassay	××××	××	××	××××	××	70
Fluorescence spectroscopy	××××	×××	××	××××	××××	85
Raman spectroscopy	××××	××××	×××	××××	—	Incomplete
FTIR spectroscopy	×××	××××	×××	××××	×××	85
Vis-NIR spectroscopy	×××	××××	×××	××××	××	80

× = Poor; ×× = Fair; ××× = Good; ×××× = Excellent.

[a] Based on capital equipment cost only for fair comparison

[b] Note: The accuracy of the standard methods is assumed excellent, because they are used as benchmarks for the innovative methods but in practice this may not be the case. The accuracy of Raman spectroscopy could not be fairly compared because the standard analytical method used in the studies we accessed was either not stated or was high-performance liquid chromatography, which is outside the scope of this review.

2.6 Conclusions

On a regular basis, analytical techniques for TPHs and/or PAHs in soils are being developed and diverse factors are coalescing to bring this about. At the top of this list of factors is the increasingly persistent demand for more profitable and simple environmental diagnostic tools capable of generating reliable data on a shorter timescale. From this study, it is obvious that there are no analytical techniques that are problem free, although some are less problematic than others. This is why efforts have been made in this study to extract as much relevant information from the open access literature as possible on the application of vis-NIR spectroscopy in particular, and analytical techniques in general, in the detection of petroleum contamination in soils. If timeliness and operator health and safety concerns are to be addressed, improved protocols aimed at completely eliminating the need for prior sample preparation and the use of chlorinated extraction solvents are of the essence. This suggested method change should necessarily target gravimetric, GC-based, immunoassay-based, and fluorescence spectroscopic techniques. This is because of their low performance in those two attributes due to their dependence on lengthy sample preparation steps involving the use of chlorinated extraction solvents and proprietary extraction kits. Although the vis-NIR technique compares well with other techniques in terms of cost-effectiveness, timeliness, operator health and safety, and portability, it was documented to be the least accurate. As discussed, this has to do with the fact that NIR prediction is based on overtones and combinations of fundamental

vibrations occurring in the MIR region. However, if the individual and interaction effects of soil texture, moisture content, and oil concentration on soil diffuse reflectance spectra and calibration models developed are well understood, and if nonlinear analytical techniques such as ANNs or support vector machine and others could be applied as alternatives to linear PLS regression analysis used so far in the literature, the accuracy of the method may be improved. These are pertinent research questions that need to be answered. To this end, there always will be room for some breakthroughs in trying to improve on existing systems by scientists and for some techniques to complement others. Therefore, there is a need for constant reviews of the progress being made to help scientists avoid reinventing the wheel.

Chapter 3: Effects of oil concentration, and moisture and clay contents on the prediction of polycyclic aromatic hydrocarbons in soils by visible and near-infrared spectroscopy

3.1 Chapter summary

Removal of petroleum hydrocarbon (PHC) contamination that is hazardous and often prevalent in soils would benefit from a rapid detection technique. Vis-NIR spectroscopy has a large potential as a rapid detection technique for PHC in soils. Nevertheless, the combined influence of oil concentration, moisture content and clay content on soil reflectance spectra and the accuracy of the technique have yet received little attention. The objective of this chapter was to investigate the combined influence of oil concentration and moisture and clay contents on the spectral characteristics of diesel-contaminated soils and the quality of calibration models developed for PAH in soils using vis-NIR spectroscopy. With PLS regression data from a systematic experimental design using 150 artificially contaminated soil samples, results showed that soil diffuse reflectance decreased with increasing oil concentration, clay and moisture contents. The trend was less defined in relation to moisture and clay due mainly to the interaction effects of the soil matrices as mediated by the oil. The PAH PLS cross-validation showed best performance with the lowest oil concentration and clay content at 20% moisture with r^2 of 0.89, RMSEP of 0.20 $mgkg^{-1}$ and RPD of 2.75. Analysis of variance showed that the interaction effects of oil concentration, moisture and/or clay content significantly ($p<0.05$) affected the quality of the PAH models.[2]

[2] Reprinted after the permission of Springer

3.2 Introduction

Vis-NIR spectroscopy is an emerging and portable technique that is gradually becoming recognized for the rapid screening of petroleum hydrocarbon-contaminated soils with reasonable levels of accuracy (Graham, 1998; Malle and Fowlie, 1998; Malley et al., 1999; Forrester et al., 2010, 2013; Bray et al., 2010; Chakraborty et al., 2010, 2012; Schwartz et al., 2012).

Clay is described as soil particles less than 0.002 mm in size and consists mostly of clay minerals. The amount of clay in the soil has a large influence on soil structure and texture as it promotes the formation of aggregates. Soil structure and texture in turn affect the movement of water in the soil. It is known that the higher the clay content is, the larger is the water retention capacity (Mouazen et al., 2005a; Stenberg, 2010; Stenberg et al., 2010) and the lower is the soil diffuse reflectance (Viscarra Rossel and McBratney, 1998). Clay also has a significant influence on sorbed TPH spectral signals (Forrester et al., 2010). The reduced apparent intensity of sorbed TPH (straight chain alkanes, C_{15}–C_{28}) spectral signal has been attributed to the shielding of the TPH within the soil matrix from the NIR beam due to the mineralogy of clay (Forrester et al., 2010). Although the clay–organic matter agglomerate is known to have a large influence on the sorption of PAHs in soil (Dexter et al., 2008), the combined effect of clay, organic matter, and water on sorbed PAH spectral signals has not been explored.

Water has a strong effect on soil vis-NIR spectra, represented in the bands of 950, 1,450 and 1,950 nm due to the resonance in the molecular vibration of the

illuminated water molecules (Whalley and Stafford, 1992). The size of these water absorption bands in the first (1,950 nm) and second (1,450 nm) overtone regions increases with increasing soil moisture content (Tekin et al., 2012). This may be because in vis-NIR spectroscopy, soil is known to become darker as the amount of water increases, thereby absorbing more light and reflecting less (Mouazen et al., 2005a). For a sandy-loam soil, changes in absorption–reflectance decrease when soil approaches saturation, e.g. field capacity, between 20 and 25% gravimetric moisture content (Mouazen et al., 2006a). This also applies to other soil texture types. Although the literature attempts to explain the effect of both moisture content and texture (in terms of clay content) on soil spectral characteristics, minimal research can be found about the combined effect of these two parameters with oil contamination effect.

This study was carried out with two objectives, namely to investigate the feasibility of using vis-NIR spectroscopy as a quantitative tool for PAHs in diesel-contaminated soils and to characterize the combined effect of moisture content, soil texture, in terms of clay content and oil concentration on soil diffuse reflectance spectra and quality of calibration models for PAHs in diesel-contaminated soils.

3.3 Materials and methods

3.3.1 Sample collection and treatment

Using unaligned systematic random sampling, soil samples from the top 15 cm of soil were collected, stored in Ziploc bags and refrigerated at 4 °C. The soil

samples were obtained from the Cranfield University's Research Farm at Silsoe, Bedford, and Duck End Farm, Wilstead, both in Central Bedfordshire, UK. A total of 150 samples of 25 g each, representing five different soil textural classes (Table 3-1), were prepared. A total of 25 samples from each texture were oven-dried at 105 °C for 24 h, pulverized and sieved to <2 mm with a 2-mm sieve, while another 25 samples, five from each texture, were used as field-moist intact soils. Of the 125 dried ground samples, 25 samples, five from each texture, were scanned first as dry non-contaminated set. The remaining 100 samples, 20 from each texture, were conditioned to 5, 10, 15 and 20% moisture contents, five samples for each moisture level, and allowed to incubate for 24 h. Prior to oil addition, 20 samples (four for each texture) were scanned first as wet non-contaminated set. Then with a precision pipette, in separate vials, 852, 1,136, 1,420, 1,705 and 1,989 µL of diesel fuel #2, corresponding to expected oil concentrations of 30,000; 60,000; 90,000; 120,000; and 150,000 mg kg^{-1}, respectively, were added to all the soil samples, five samples for each concentration, thoroughly mixed and allowed to incubate for 48 h before scanning.

Table 3-1 Pertinent physical properties of the soil samples

Soil ID	Texture fractions (%)			Soil texture classification [a]	Moisture content of field-moist samples (%)
	Sand	Silt	Clay		
A	74	17	9	Sandy-loam	.5.41
B	61	19	20	Clay-loam	9.04
C	32	33	35	Sandy-clay-loam	.5.05
D	52	22	26	Loamy-sand	.6.13
E	9	17	74	Clay	.1.91

[a] United States Department of Agriculture (USDA) soil textural classification

3.3.2 Laboratory analysis of soil properties

3.3.2.1 Particle Size Measurements

Soil particle size was measured by means of low angle laser light scattering, i.e. laser diffraction, with a Mastersizer2000® (Malvern Instruments, Worcestershire, UK) coupled to a HydroMu® dispersing unit (Malvern Instruments, Worcestershire, UK). Measurements were taken in the range between 0.02 and 2,000 µm based on Sieve ASTM E1161 standard operating procedures, under the following instrument conditions: sample refractive index of 1.52, sample absorption coefficient of 0.1, water refractive index of 1.33, sample and background measurement times of 15 s, sample and measurement integration time of 15,000 ms, maximum obscuration of 20% (because of the presence of sand >50 µm) and five measurement cycles with 0 s delay between measurements. Particle size distribution was characterized in terms of volume percent of particles. Soil textural classification was done based on percent clay, silt and sand (Table 3-1) with reference to the United State Department of Agriculture soil textural classification scheme.

3.3.2.2 Moisture content (mass basis) determination

Soil moisture content was determined by oven drying the soil samples at 105±5 °C for 24 h. The amount of moisture in the soil was deduced from the difference in mass of 30 g (to 0.001 g) of field-moist soil sample before and after drying (See Table 3-1).

3.3.2.3 Extraction and analysis of polycyclic aromatic hydrocarbons

The sequential ultrasonic solvent extraction method (Risdon et al., 2008) was used to extract PAH compounds from the diesel-spiked soil samples. Dichloromethane (DCM) was used in place of acetone, and except the field-moist intact soils, the dried graded soil samples were not subjected to chemical drying with anhydrous Na_2SO_4 since there was no residual moisture in them. Five grams of the spiked soil samples was extracted with 6 mL DCM and sonicated for 2 min at 20°C. A 10-mL volume of hexane and 4-mL volume of DCM (1:1) was added to the sample, sonicated for a further 10 min at 20°C, shaken manually and then centrifuged for 5 min at 750 rpm. Samples were decanted, filtered and the supernatant was then subjected to a 2-step 20°C sonication with 10 mL 1:1 DCM/hexane mix, manually shaken for re-suspension, 5-min centrifugation at 750 rpm, decantation and filtration. The extract clean-up was carried out in a ~0.6-g Florisil®, on glass wool, column pre-washed with DCM. The final extract was then made up to 40 mL with a 1:1 DCM/hexane mixture. Prior to analysis by gas GC–MS, 100 µL of mixture of deuterated PAHs (naphthalene-d^8, anthracene-d^{10}, chrysene-d^{12} and perylene-d^{12}) internal standards of concentration 0.5 µg mL^{-1} were added to 900 µL of the extract in an amber coloured, polytetrafluoroethylene-capped GC vial.

Analysis of PAHs was carried out by GC–MS with a 6890N Network Gas Chromatographic System (Agilent Technologies Inc., USA) coupled to a 5973 Network Mass Selective Detector (MSD) (Agilent Technologies Inc., USA) operated at 70 eV in positive ion mode. Before use, an external five-level

calibration was carried out with the EPA 525 Standard PAH Mix A and deuterated PAHs at 1 to 5 and 0.5 µg mL^{-1}, respectively. A quality control standard, 1 mL of 1 µg mL^{-1} PAH Mix standard, was analysed every 20 samples. Also, a solvent blank (DCM/hexane, 1:1) was run each day prior to sample analysis. The GC column was a fused silica capillary column (30 x 0.25 mm id) coated with RTX®-5MS (0.25 µm film thickness; Restek Corporation, UK) with helium as the carrier gas. Measurement was taken under the following instrument conditions: Oven temperature was programmed at 85°C for 2 min and then increased to 310°C at 20°C min^{-1} and held at this temperature for 38 min. Splitless injection with a sample volume of 1 µL was applied. The mass spectrometer was operated using the full scan mode (range m/z 50–650) for quantitative analysis of target PAHs. For each compound, quantification was performed by integrating the peak at specific m/z using MSD ChemStation® software. Major hydrocarbon compounds of the diesel fuel #2 were identified on the basis of their retention time and by comparing them to those of analytical standards.

3.3.2.4 Visible and near infrared optical measurement

Optical measurement of the soil samples in diffuse reflectance mode was carried out using a mobile fibre-optic vis-NIR spectrophotometer (350–2,500 nm) (LabSpec2500® Near Infrared Analyser, Analytical Spectral Devices Inc., USA), equipped with one Si array (350–1,000 nm) and two Peltier cooled InGaAs detectors (1,000–1,800 and 1,800–2,500 nm). The sampling interval of the instrument was 1 nm. The spectral resolution was 3 nm at 700 nm and 10

nm at 1,400 and 2,100 nm. A high intensity probe with a built-in light source was used. A quartz-halogen bulb of 3,000-K light source and a detection fibre are gathered in the high-intensity probe enclosing a 35° angle. Prior to scanning the soil samples, and periodically at intervals of 30 min, white disc reference measurements were taken. Soil samples were packed into the sampling cuvette, gently pressed and levelled with a spatula, resulting in a smooth soil surface to ensure maximum diffuse reflection and, thus, a good signal-to-noise ratio (Mouazen et al., 2005a). Soil samples were then placed in direct contact with the high intensity probe. Reflectance spectra were measured from each sample at three different positions (120° apart), averaged and used for spectral pre-processing and multivariate analysis.

3.3.3 Spectral pre-processing

In order to correct for the scattering effect of light, it may be necessary to transform the spectral data prior to calibration. The pre-processing of soil spectra was carried with the Unscrambler® 9.8 (CAMO Software, Woodbridge, NJ, USA). Firstly, noisy portions of the spectrum at 350 to 449 and 2,451 to 2,500 nm were removed due to low instrument sensitivity at these wavelengths. Noise removal was followed by smoothing by averaging successive 5-nm wavelengths. Then the selected working wavelength range (452–2,450 nm) showed distinct reflectance dips and peaks. In this study, various data pre-processing models were combined. Eventually, maximum normalization, Savitzky–Golay first derivative and smoothing were successively implemented for all subset of samples based on the knowledge of the effects that resulted in

the largest cross-validation r^2. From the literature, normalization helps to bring all data to approximately the same scale or to achieve a more even distribution of the variances and the average values. The first derivative removes an additive baseline, and smoothing reduces the impact of noise, both aimed at reducing spurious peaks that do not hold physical or chemical information.

3.3.4 Calibration and validation models: partial least squares (PLS) regression analysis

PLS regression was carried out using the Unscrambler®. This combines both the independent variables (reference values of PAH) and the dependent variables (wavelengths) using them as regression generators for the independent variables (Maleki et al., 2006). PLS regression analysis was performed in two steps: firstly, to understand the effects of the design parameters on the quality of models developed for PAH and to select the optimal number of components (latent factors). The entire dataset (25 samples) in each subset of sample was used for training with full cross-validation considering up to 12 latent factors. Then, the optimal number of latent factors was determined based on the number of factors with the first local minimum. Secondly, because of the limited number of samples in each subset, the entire 150 samples from the six subsets were randomly split into 90% calibration (training) and 10% validation (test) sets. Partial least squares regression with leave-one-out cross-validation was performed using the training dataset. The same variable selection and optimization scheme described above was also carried out. The validation dataset was then used to test the prediction accuracy

of the calibration model. While building the PAH calibration models, the occurrence of outliers and their effects on models prompted the need to remove influential X- and Y-outliers that differed from the reference by three times the standard deviation of the predicted residuals. The criteria used to evaluate the quality of the models were based on the root-mean-square error (RMSE) of cross-validation and prediction, the ratio of standard deviation of sample concentration to the root-mean-square error (RPD), and corresponding coefficient of determination, r^2, according to Equations 3-1, 3-2, 3-3, and 3-4 (Naes et al., 2002). RPD was originally defined as the ratio of the standard error of prediction to the standard deviation of the reference data in the validation set (Williams and Sobering, 1986).

$$\mathrm{RMSECV} = \sqrt{\frac{\sum_{i=1}^{N}(\hat{y}_{CV,i} - y_i)^2}{N}} \tag{3-1}$$

$$\mathrm{RMSEP} = \sqrt{\frac{\sum_{i=1}^{N_P}(\hat{y} - y_i)^2}{N_P}} \tag{3-2}$$

$$\mathrm{RMSE} = \sqrt{\mathrm{MSE}} = \sqrt{E(\hat{y} - y)^2} \tag{3-3}$$

$$\mathrm{RPD} = \frac{\mathrm{SD}}{\mathrm{RMSE}} \tag{3-4}$$

Where $\hat{y}_{CV.i} =$ Estimate for y_i based on the calibration equation with *i* deleted; $\hat{y}\ and\ y_i =$ Predicted and measured reference values respectively; $E(.) =$ Statistical expectation (average) over the population of future samples; N =

Number of samples in the set; SD = Standard deviation of the measured reference values.

Model prediction ability was categorised based on the following criteria: excellent if RPD > 2.0, almost good if 1.4 ≤ RPD < 2.0, and unreliable if RPD < 1.4 (Chang et al., 2001).

3.3.5 Principal component analysis (PCA)

To extract relevant information on the vis-NIR reflectance spectra of the various oil-impacted soil groups, the PCA was applied on the overall dataset. The PCA transforms the original independent variable (wavelengths) into mathematical terms called principal components (PCs), which express the total variability in the dataset in a few PCs in order of decreasing variance. The first PC covers most of the variability in the dataset; the second PC is orthogonal to the first and explains as much of the remaining variability as possible and so on (Viscarra Rossel et al., 2006a). A plot of the PCs can be used to view the relationships between different variables and discriminate samples according to the underlying design. In this study, PCA was carried out with the Unscrambler®.

3.3.6 Statistical analysis

To evaluate the significance of the effects of oil concentration, moisture and clay contents on prediction of PAH, one-way analysis of variance (ANOVA) was carried out using RPD and RMSE of cross-validation as accuracy indicators with the ANOVA data analysis toolbox in Microsoft® Excel 2010. Differences were considered significant at $p<0.05$.

3.4 Results and discussion

3.4.1 Combined effects of oil concentration, clay and moisture contents on soil diffuse reflectance spectra

The plots in Figure 3-1 were average reflectance spectra of the five diesel concentrations, five soil textures (in terms of clay content) and four moisture contents. For the uncontaminated soils, the majority of the spectra showed that soil diffuse reflectance decreased with increasing clay content (Figure 3-1a), which agrees with a previous report (Viscarra Rossel and McBratney, 1998). Due to the interaction effect of molecular absorption of clay minerals and the scatter effect of sand fractions, the odd behaviour of soil spectrum of 74% clay can be explained. The other reason might be the lighter colour of this soil as compared to the other soils, which masks the texture effect.

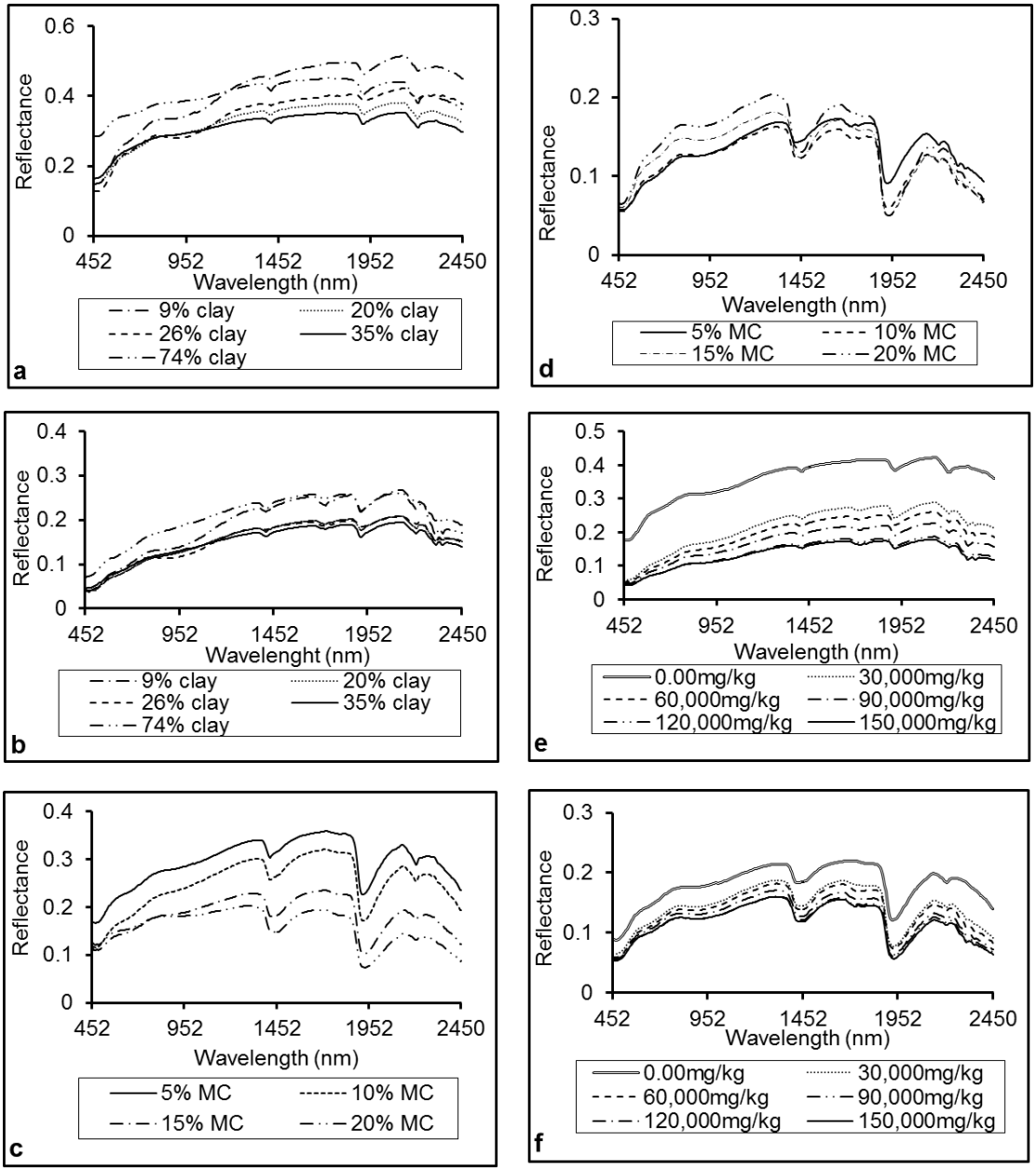


Figure 3-1 Variation of visible and near infrared red average soil diffuse reflectance spectra

(a) With clay content of dry soil samples before contamination, (b) With clay content of dry soil samples after contamination, (c) With moisture content of soil samples before contamination, (d) With moisture content of soil samples after contamination, (e) With diesel concentration of dry soil samples, and (f) With diesel concentration of field-moist intact soils.

Similarly, before contamination, soil diffuse reflectance decreased with increasing moisture content (Figure 3-1c), which also is in line with earlier reports (Mouazen et al., 2005a; Stenberg, 2010). Nevertheless, after oil contamination, soil diffuse reflectance up to about 1,200 nm decreased with moisture, whereas it decreased with increasing moisture, particularly at larger wavebands around 1,900 nm (Figure 3-1d). This is clearly a new feature regarding the effect of moisture in NIR spectroscopy. This can be explained by the interaction effect of oil concentration with moisture. The increased absorption or decreased reflectance with moisture in the visible range up to 1,200 nm might be due to the colour change of the oil–soil mix material. It is possible that the addition of water to diesel actually lightens the colour of diesel in soil. Furthermore, after oil contamination, soil diffuse reflectance decreased with increasing oil concentration both for the dry soil samples (Figure 3-1e) and the field-moist (representing mixed moisture) soil samples (Figure 3-1f). This is explained by the increase in soil darkness with oil concentration, which increases the overall absorption, in a similar way as that resulting from increase in moisture content of soil sample. Nevertheless, when soil is wet, the decrease in reflectance or increase in absorption is of much smaller increment, as compared to dry soil (Figure 3-1e).

Currently, no reports could be found to explain these spectral characteristics, which have resulted from the interaction of oil and water within the soil matrix. In the interim, the relationship between soil colour and shape of spectra reported in Mouazen et al. (2007b) was adopted in this study to account for these

spectral responses. As it turns out, examining the variation of the average soil diffuse reflectance spectra with respect to the soil matrices (clay content, moisture and oil concentration) before and after oil contamination (Figure 3-1) confirms the presence of colour associated variation. According to Mouazen et al. (2007b), the shape of spectra relates to soil colour such that the slope of the linear regression line increases and intercept decreases with soil darkness and vice versa. Table 3-2 shows the results of a linear regression analysis between wavelength and reflectance carried out in this study between the wavelength ranges of 452–1,702 nm. The table shows significant changes in the shape of spectra after oil contamination.

Table 3-2 Linear regression results of the variation of visible and near infrared red diffuse reflectance spectra with oil concentration of non-contaminated and contaminated dry and field-moist soil samples (average of five textures) between the wavelength ranges of 452 to 1,702 nm.

Oil concentration ($mgkg^{-1}$)	Dry soils			Field-moist soils		
	r^2	Slope (nm^{-1})	Intercept	r^2	Slope (nm^{-1})	Intercept
0.0 (no oil)	0.91	8.0×10^{-4}	23.1×10^{-2}	0.77	4.0×10^{-4}	12.9×10^{-2}
30,000	0.96	8.0×10^{-4}	8.5×10^{-2}	0.72	4.0×10^{-4}	10.0×10^{-2}
60,000	0.96	8.0×10^{-4}	7.4×10^{-2}	0.72	4.0×10^{-4}	9.6×10^{-2}
90,000	0.96	7.0×10^{-4}	6.7×10^{-2}	0.69	3.0×10^{-4}	8.8×10^{-2}
120,000	0.96	6.0×10^{-4}	5.8×10^{-2}	0.67	3.0×10^{-4}	9.3×10^{-2}
150,000	0.96	5.0×10^{-4}	5.7×10^{-2}	0.65	3.0×10^{-4}	8.7×10^{-2}

r^2, coefficient of determination

Both the slope and intercept of the linear regression line decrease with different degrees after spectral contamination with oil (Table 3-2). The oil contamination also led to the decrease in these two indicators from the uncontaminated spectra through the spectra of the lowest oil concentration to the highest (Table

3-2). Although the reduction in the intercept values with oil concentration and thus darkness was in line with the conclusion of Mouazen et al. (2007b), opposite variation was found for the slopes. The reason might be the different changes in spectral geometry associated with colour change in the visible range with oil contamination, as compared to the natural change of soil colour reported by Mouazen et al. (2007b). Although Mouazen et al. (2007b) did also state that the degree in linearity increases with increasing soil darkness; the r^2 values for the dry soils after contamination were the same for all concentrations (Table 3-2). Again, the field-moist soils contradicted the same assertion by Mouazen et al. (2007b) as degree of linearity decreased with increasing oil concentration and soil darkness (Table 3-2), which might be attributed to changes in spectral geometry caused by colour changes due to the interaction effect of moisture and oil concentration as distinct from that due to the natural darkness of soil samples. The dramatic drop of more than an average of 4-fold in the dry soils and 2-fold in the field-moist soils in the slope and intercept responds to the additional increase in the soil darkness imparted by the contaminating oil resulting in increased soil absorption. The magnitude of this additional increase in darkness was dependent on the magnitude of the difference between the values of the intercept of the regression lines before and after oil contamination (Table 3-2). This difference appeared to increase with oil concentration (Table 3-2). For instance, the darkest soil appears to be the one with the highest oil concentration and the largest difference in intercept while the lightest soil appears to be the one with the lowest oil concentration and the smallest difference in intercept. Table 3-3 shows the linear regression results of

the variation of vis-NIR spectra with oil concentration and moisture contents before and after diesel contamination between the wavelength ranges of 452 to 1,602 nm. By extension, the decreases in soil diffuse reflectance or increase in soil absorption after contamination with increasing clay content (Figure 3-1b) may be attributed to the difference in the intercept of the regression line before and after oil contamination as shown in Table 3-3.

Table 3-3 Linear regression results of the variation of visible and near infrared red average diffuse reflectance spectra with clay content (average of four moisture contents and five oil concentrations) and moisture contents before (for only 74% clay) [a] and after (average of five textures and five oil concentrations) diesel contamination between the wavelength ranges of 452 to 1702 nm.

Variable	Before contamination			After contamination		
	r^2	Slope (nm^{-1})	Intercept	r^2	Slope (nm^{-1})	Intercept
Clay (%)						
9	0.94	12.0×10^{-4}	2.2×10^{-1}	0.98	9.0×10^{-4}	5.3×10^{-2}
20	0.90	8.0×10^{-4}	2.0×10^{-1}	0.97	7.0×10^{-4}	5.5×10^{-2}
26	0.90	10.0×10^{-4}	1.9×10^{-1}	0.96	7.0×10^{-4}	5.5×10^{-2}
35	0.86	6.0×10^{-4}	2.2×10^{-1}	0.95	6.0×10^{-4}	6.5×10^{-2}
74	0.92	6.0×10^{-4}	3.2×10^{-1}	0.95	8.0×10^{-4}	9.9×10^{-2}
Moisture content (%)						
5	0.89	7.0×10^{-4}	2.1×10^{-1}	0.85	4.0×10^{-4}	8.4×10^{-2}
10	0.88	7.0×10^{-4}	1.6×10^{-1}	0.67	3.0×10^{-4}	9.1×10^{-2}
15	0.76	7.0×10^{-4}	1.4×10^{-1}	0.50	3.0×10^{-4}	11.0×10^{-2}
20	–	–	–	–	–	–

[a] For non-contaminated soils, diffuse reflectance would normally decrease with increasing moisture content across a wide range of soil textures from previous study (Mouazen et al., 2007b), so spectra from only one texture were considered reasonable to illustrate this view.

– Data with r^2 <0.5; r^2 coefficient of determination

Nevertheless, the irregularities observed in the diffuse reflectance spectra in Figure 3-1a, b, which were slightly less defining of the expected trend, may be attributed to the marginal differences in slopes and intercepts of their respective

regression lines (Table 3-3). After oil contamination, the increases in intercepts of the regression lines and hence decrease in soil darkness with increase in moisture (Table 3-3) suggest that the change in spectral geometry is dominated by colour associated changes in the visible range up to around 1,200 nm. This corroborates our earlier suggestion that the oil–soil mix material in the presence of water may well be lighter in colour.

Five reflectance shoulders in the soil spectra before oil contamination between the wavelengths of 452–2,201 nm were identified at 600, 780, 1,350, 1,830 and 2,100 nm (Figure 3-1a, c). Six reflectance shoulders after oil contamination between the same wavelengths were identified at 600, 780, 1,350, 1,630, 1,830 and 2,100 nm (Figure 3-1b, d–f). Reflectance shoulders appear between two absorption wavelengths of colour and/or water, the magnitude of which depends on soil colour and moisture content (Mouazen et al., 2007b). Of the six reflectance shoulders appearing after oil contamination, colour-associated variations were responsible for the appearance of the shoulders at 600 and 780 nm due to blue colour absorption band around the 450 nm, red colour absorption band at 680 nm and water absorption band at 960 nm (Mouazen et al., 2007b). Water absorption at the second overtone region of water at 1,450 nm is responsible for the shoulders at 1,350 and 1,630 nm (Mouazen et al., 2007b). Nevertheless, we observed in this study that the reflectance shoulder at 1,630 nm may also be attributed to hydrocarbon absorption band in the first overtone region around 1,712 nm (Osborne et al., 1993; Workman and Weyer, 2008). This is a very distinguishing hydrocarbon band, which can be clearly

observed by comparing non-contaminated (Figure 3-1a, c) with contaminated spectra (Figure 3-1b, d). The reflectance shoulder at 1,830 nm may be due to water absorption in the first overtone region at 1,950 nm (Stenberg, 2010) and hydrocarbon absorption at 1,712 nm. As seen in Figure 3-1, the two reflectance shoulders at 1,630 and 1,830 nm are present only when there is absorption in the first overtone region due to C–H stretching modes of ArCH linked to PAHs, and saturated CH_2 and terminal CH_3 groups linked to TPH all present in the contaminating hydrocarbon oil. The remaining shoulder at 2,100 nm may be attributed to absorptions in the combinations band due to long chain C–H+C–H and C–H+C–C around 2,212 nm linked to soil organic carbon, which is outside the scope of this study. It, therefore, may follow that the relationships between soils diffuse reflectance spectra and oil concentration also associate with soil colour absorption bands in the visible region of the vis-NIR spectrum.

3.4.2 Effects of moisture, clay and oil concentration on calibration precision of polycyclic aromatic hydrocarbon models

Three sample outliers were eliminated from the cross-validation dataset in each subset of samples. All six subset of sample generated a total of 16 cross-validation models (Table 3-4). When all calibration models were compared, the best model was the 30,000-$mgkg^{-1}$ oil concentration with the least RMSE of cross-validation and the highest RPD values (Table 3-4). As shown in Table 3-4, the PLS models used different numbers of latent factors in the cross-validation models. Ideally, the optimal number of factors might be expected to be three or less if Beer's law applies and the pre-processing removed unwanted

physical effects (Naes et al., 2002). Nevertheless, because of the systematic stratification of the samples and the interaction effect of the experimental design parameters, the variability in the number of latent factors can be explained. Although most authors suggest that a smaller number of these factors are used, many more have reported even higher factors with a related sample matrix, e.g. Aske et al. (2001). An equally important fact, nonetheless, is that the validation, when done on a separate sample set, should give stable and good prediction accuracy.

Table 3-4 Summary of calibration results of polycyclic aromatic hydrocarbons using partial lease squares cross-validation models at different oil concentrations, moisture and clay contents

Variable	Calibration quality indicators			Number of LV
	RMSE ($mgkg^{-1}$)	RPD	r^2	
Moisture content (%)				
0 (dry)	0.36	1.99	0.72	1
5	0.48	1.53	0.61	2
10	0.47	1.59	0.62	5
15	0.44	1.62	0.61	3
20	0.30	2.53	0.82	7
Field-moist (mixed MC)	0.39	1.93	0.72	4
Clay content (%)				
9	0.39	1.99	0.70	5
20	0.48	1.55	0.51	3
26	0.41	1.51	0.54	4
35	0.44	1.72	0.64	2
74	0.35	1.79	0.63	5
Oil concentration ($mgkg^{-1}$)				
30,000	0.12	3.20	0.89	3
60,000	0.26	2.21	0.79	2
90,000	0.40	2.05	0.72	2
120,000	0.35	2.04	0.75	1
150,000	0.44	1.67	0.59	1

RMSE, root-mean-square error; RPD, ratio of the standard error of prediction to the standard deviation of the reference data in the validation set; LV latent variables

3.4.2.1 Effect of Moisture Content

Although views expressed in the literature show that calibration models developed for dry samples always are better than those for wet samples because water is known to decrease the accuracy of vis-NIR-predicted soil properties (Tekin et al., 2012; Nocita et al., 2013), the lower model accuracy achieved in the current study for the dry samples as compared to the 20% moisture samples is in line with the findings of Charkraborty et al. (2010) who reported very large prediction error (0.55 $\log_{10}$ $mgkg^{-1}$) and even lower RPDs (1.34 and 1.06) for air-dried ground scans for reflectance and first-derivative TPH models, respectively. Moreover, Tekin et al. (2012) reported an excellent model performance with wet samples of up to 10% moisture for soil organic carbon. Nonetheless, the performance achieved in this study based on RPD values can be classified as good for the dried ground samples, whereas the performance for the 20% moisture and field-moist intact samples can be classified as excellent and almost good, respectively (Chang et al., 2001). Nevertheless, between the fairly performing 5% and 10% moisture models, the 5% moisture model showed worse performance with lower RPD and higher RMSE values as confirmed in Figure 3-2a. Tekin et al. (2012) and Nocita et al. (2013) also reported similar results with soil organic carbon for 5 and 15% moisture, respectively. The good model performance of the field-moist samples suggests the possibility of quantitative predictions without recourse to lengthy prior sample preparations.

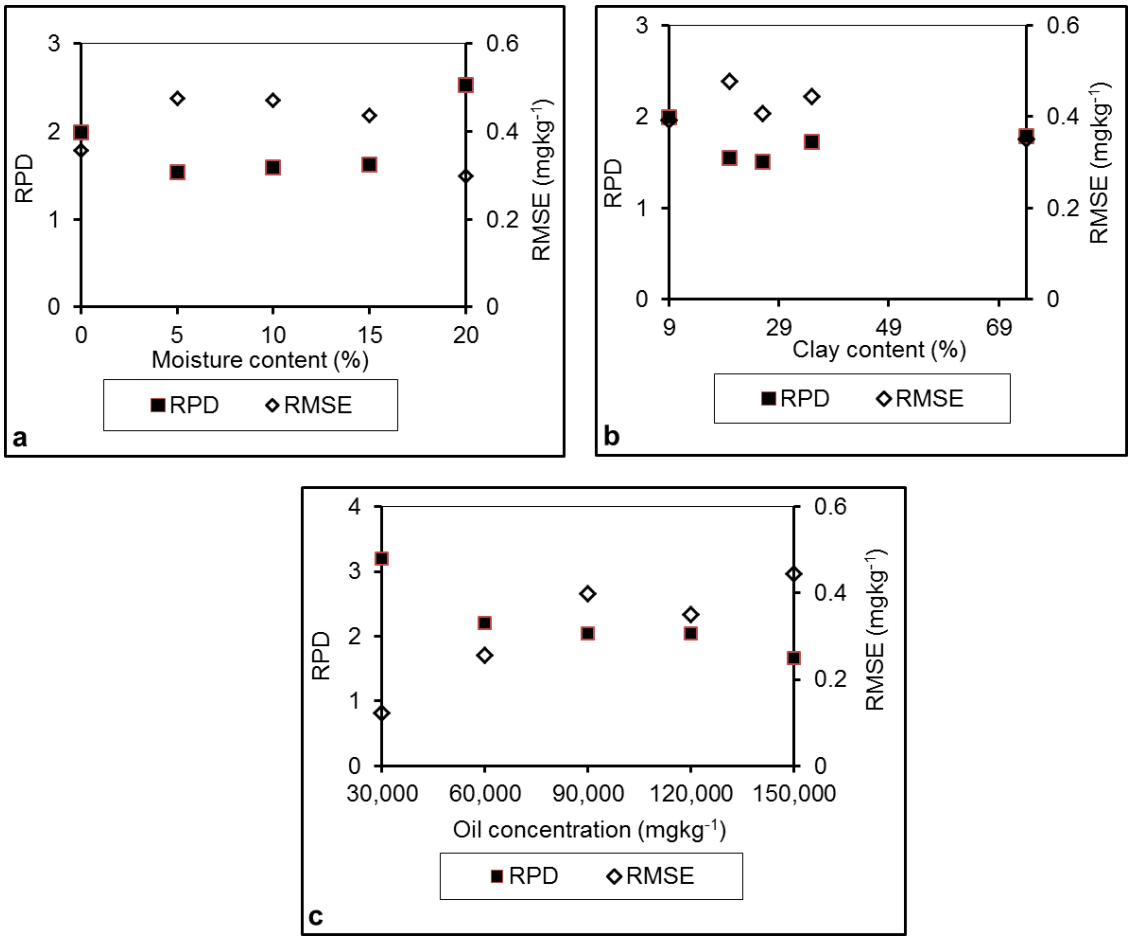


Figure 3-2 Variation of RPD and RMSE values of partial least-squares cross-validation models developed for polycyclic aromatic hydrocarbons (PAH) of diesel-contaminated soils

(a) With moisture content, (b) With clay content, and (c) With oil concentration

3.4.2.2 Effect of clay content

The good calibration performance observed with the least clay content model, 9% clay (Table 3-4), agrees with the findings of Forrester et al. (2010) that minor amounts of clay can seriously affect the apparent intensity of sorbed hydrocarbon (TPH) spectral signals and hence the calibration performance with vis-NIR spectroscopy. As shown in Figure 3-2b, the poorest calibration performance occurred at 26% clay content. This result is in agreement with the

report of Forrester et al. (2010) that a dramatic signal reduction by up to a factor of 100 for sand containing up to 25% clay was observed for TPH calibration. Nevertheless, the unexpected relatively high calibration performance of the 35 and 74% clay models may not be unconnected with the interaction effect of water and oil.

3.4.2.3 Effect of Oil Concentration

Table 3-4 shows that calibration performance, in terms of the RPD, of PAH models with vis-NIR spectroscopy decreased with increasing oil concentration. As shown in Figure 3-2c, the RPD value decreased progressively with increasing oil concentration and attaining saturation at about 90,000 $mgkg^{-1}$ oil con-centration. It thus appears that for measurement of PAHs with vis-NIR spectroscopy, best models performance can be expected at very low clay content and oil concentrations at moisture of no more than 20% (Figure 3-2).

The ANOVA results in Table 3-5 showed that the combined effect of oil concentration, clay and/or moisture resulted in significant ($p<0.05$) differences in the quality of the PAH calibration models. This implies that the interaction of the experimental design parameters significantly ($p<0.05$) affected the precision, in terms of the RMSE and RPD values, of measurement of PAH with vis-NIR spectroscopy. Nevertheless, Table 3-5 showed that the accuracy of vis-NIR spectroscopy was mostly affected by the combined influence of the three design parameters, namely oil concentration, moisture and clay contents as shown by the p values.

Table 3-5 One-way ANOVA on the analysis of the significance of the combined effect of oil concentration, moisture and clay contents on the accuracy of prediction of polycyclic aromatic hydrocarbons using results of partial least squares cross-validation.

Option	Source of variation	S.S.	d.f.	M.S.	F	p value	F crit.
A	Between groups	14.5348	3	4.8449	36.61*	2.1×10^{-7}	3.24
	Within groups	2.1174	16	0.1323	–	–	–
	Totals	16.6522	19	–	–	–	–
B	Between groups	13.6446	3	4.5482	47.00*	3.7×10^{-8}	3.24
	Within groups	1.5484	16	0.0968	–	–	–
	Totals	15.1930	19	–	–	–	–
C	Between groups	9.4512	3	3.1504	56.72*	9.6×10^{-9}	3.24
	Within groups	0.8887	16	0.0555	–	–	–
	Totals	10.3399	19	–	–	–	–
D	Among groups	18.8735	5	3.7747	39.78*	7.5×10^{-11}	2.62
	Within groups	2.2773	24	0.0949	–	–	–
	Totals	21.1508	29	–	–	–	–

A combination of moisture content and oil concentration, **B** combination of clay content and oil concentration, **C** combination of moisture content and clay content, **D** combination of oil concentration, moisture and clay contents; S.S., sum of squares, M.S., mean squares, d.f., degree of freedom;

*$p<0.05$ (significant)

3.4.3 Accuracy of prediction of polycyclic aromatic hydrocarbons

No outlier was removed from the validation set while three sample outliers were eliminated from the cross-validation dataset. Table 3-6 summarises the sample statistics and results of PLS model for the prediction of PAH in cross-validation and validation data sets. Although samples of the cross-validation and validation sets were randomly selected, Table 3-6 shows that calibration set mean (2.22) and standard deviation (0.75) of PAH are similar to the validation set mean (1.92) and standard deviation (0.55). This similarity among calibration and validation data sets implied that the prediction model was not skewed.

Table 3-6 Sample statistics and results of partial least squares model for the prediction of polycyclic aromatic hydrocarbons in cross-validation and validation data sets for diesel-contaminated soil samples using visible and near infrared spectroscopy

Number of samples	Minimum ($mgkg^{-1}$)	Maximum ($mgkg^{-1}$)	Mean ($mgkg^{-1}$)	SD	RMSE ($mgkg^{-1}$)	RPD
Cross-validation set						
135	1.08	3.94	2.22	0.75	0.28	2.65
Validation set						
15	1.08	2.86	1.92	0.55	0.20	2.75

SD, standard deviation; RMSE, root-mean-square error; RPD, ratio of the standard error of prediction to the standard deviation of the reference data in the validation set

Figure 3-3 shows the scatter plots of measured versus predicted PAH of the PLS models in cross-validation and validation sets for the entire 150 randomly split samples. The two parameters, validation r^2 of 0.89 (Figure 3-3) and RPD of 2.75 (Table 3-6), obtained for PAH prediction in this study indicate the potential of vis-NIR spectroscopy to predict PAH in diesel-contaminated soils. Also, the RMSECV of approximately 7% achieved in this study is smaller than the approximately 10% reported in Forrester et al. (2010) for TPH prediction. An RPD value of 2.75 means excellent model predictions (Chang et al., 2001).

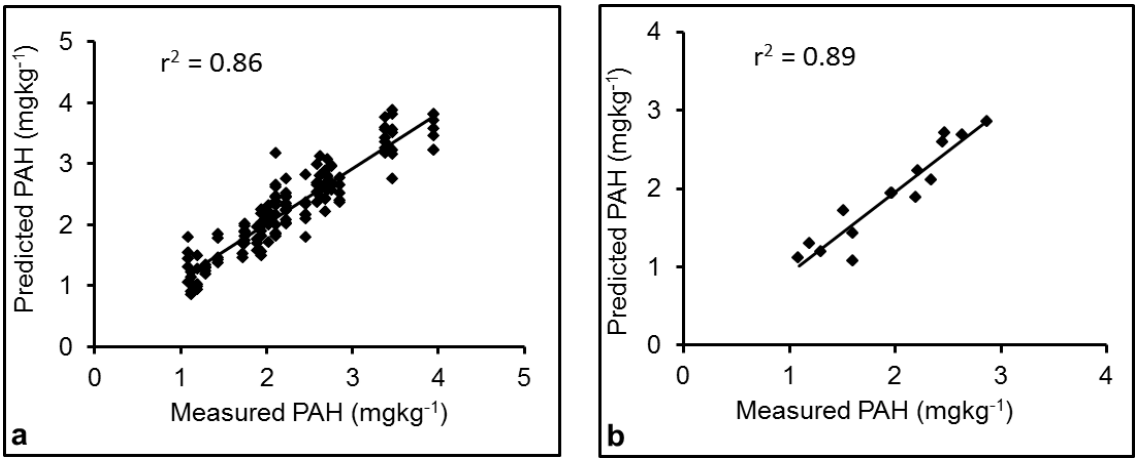


Figure 3-3 Scatter plot of chemical versus visible and near infrared spectroscopy predicted values of polycyclic aromatic hydrocarbons (PAH) using partial least squares models developed with 150 soil samples and validated (a) with cross-validation (135 samples), and (b) with validation (15 samples) sets.

3.4.4 Regression coefficients

Regression coefficients approximate observed model parameters resulting from the linear combination of the predictors in PLS regression. The regression coefficients plot is used to identify important wavelengths for the prediction of relevant soil property. Figure 3-4 was derived from the PLS analysis using vis-NIR diffuse reflectance spectra of diesel-contaminated dry graded (Figure 3-4a), field-moist (Figure 3-4b) and wet graded (Figure 3-4c) soil samples. The height of each bar at each wavelength is equivalent to the intensity of the regression coefficient. This plot over the modelling wavelength range of 452–2,450 nm for the three differently prepared contaminated soil samples confirmed the presence of peaks similar to the ones in the contaminated spectra shown earlier in Figure 3-1. These peaks and their corresponding wavelengths are shown in Table 3-7. The entire wavelength range (452–2,450 nm) appear to be

dominated by peaks associated with overtones of C–H stretching modes of ArCH, saturated CH_2 and terminal CH_3 groups, as well as O–H stretching modes of H_2O (Osborne et al. 1993; Workman and Weyer, 2008).

Nevertheless, the peaks linked to ArCH, probably PAHs usually found in diesel, appear around 1,647 nm in the first overtone region (Figure 3-4 and Table 3-7) while the peaks linked to saturated CH_2 and terminal CH_3 group, peculiar to TPH content of diesel fuel, appear around 1,712 nm in the first overtone region (Figure 3-4 and Table 3-7). These peaks are characteristic of the contaminated spectra observed previously in Figure 3-1. These also are in agreement with the widely acknowledged view that the spectra of hydrocarbon derivatives originate mainly from combinations or overtones of aromatic C–H functional groups or C–H stretching modes of saturated CH_2 and CH_3 (Aske et al., 2001; Forrester et al., 2010). This distinction in spectral signatures may be very useful in separating vis-NIR signals for ArCH associated with PAHs from saturated CH_2 and short-chain CH_3 groups associated with TPH in diesel. Equally important peaks associated with colour absorption band in the blue region of the visible spectrum were conspicuous (even in the dry graded samples plot [Figure 3-4a]) around the 497 nm in the fifth overtone region. This result agrees with our earlier observation that colour associated changes in the visible region might also be responsible for the characteristic decrease in soil diffuse reflectance with oil concentration. Other peaks observed, although not within the scope of this study, include RNH_2 at 1,002 nm in the second overtone, cellulose at 1,820

nm in the first overtone and long chain C–H+C–H and C–H+C–C at 2,212 nm in the combinations band of water and mineralogy of clay.

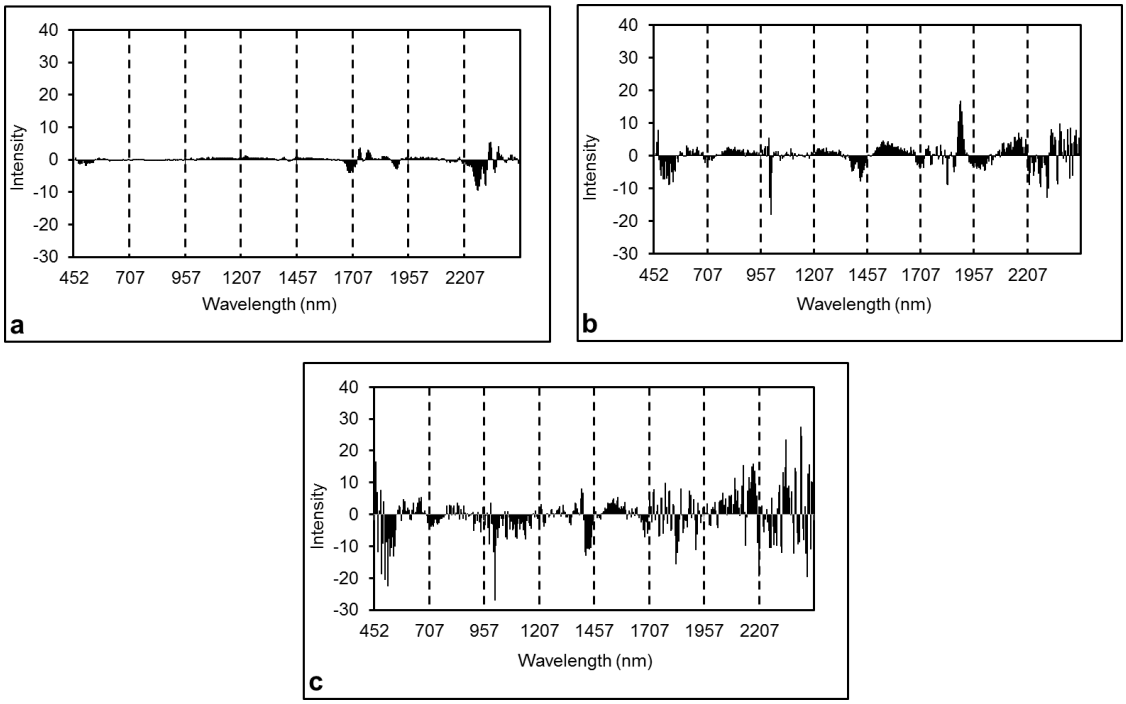


Figure 3-4 Regression coefficient plots derived from the partial least squares analysis using visible and near infrared diffuse reflectance spectra of diesel-contaminated (a) dry graded, (b) field-moist, and (c) wet graded soil samples.

Table 3-7 Some observed visible and near infrared absorption bands and their corresponding wavelengths in the range 452–2,452 nm

Absorption band	Vibration mode	Chemical structure	Wavelength (nm)
Blue	–	–	497
4th overtone	C–H stretch	ArCH and CH_3	712
2nd overtone	N–H stretch	RNH_2	1,002
2nd overtone	C–H stretch	ArCH	1,097
2nd overtone	C–H stretch	CH_3	1,137
2nd overtone	C–H stretch	CH_2	1,162
2nd overtone	C–H stretch	CH	1,227
Combinations	C–H stretch and C–H deformation	ArCH	1,417
1st overtone	C–H stretch	ArCH	1,647
1st overtone	C–H stretch	CH_3, CH_2, and CH	1,712
2nd overtone	O–H stretch and C–O stretch	Cellulose	1,820
Combinations	C-H+C-H Stretch and C-H+C-C stretch	CH_3, CH_2, and CH	2,212

3.4.5 Unsupervised classification by principal component analysis

The PCA score plots of the vis-NIR diffuse reflectance spectra of the entire dataset are shown in Figure 3-5. The PCA did provide qualitative information about the main patterns latent in the entire dataset. For the contaminated (wet and dry) samples, the first two PCs accounted for 98.4% of the total variance with PC1 accounting for 78.1% of the total variance (Figure 3-5a). One fact was easily discernible, namely that there was no clear separation of the soil samples into the different levels of oil contamination, which might be due to the interaction effect among oil, water and soil texture. For the contaminated (wet and dry) and non-contaminated dry samples, two PCs explained 98.6% of the initial variance, with PC1 accounting for 94.3% of the initial variance (Figure 3-5b). In this case, a clear separation of the contaminated and non-contaminated samples was observed along PC1. This agrees with the result of Chakraborty et al. (2010) that PCA could reasonably separate contaminated and non-contaminated samples. Similarly, along PC2, the samples were clearly separated according to their water content, with the dry samples partitioning towards the positive (north) end of PC2 and the wet samples towards the negative (south) end of PC2. How-ever, in Fig. 5c (PC1, 94%; PC2, 5%), there was an incomplete separation of the samples into contaminated and non-contaminated sets due to some degree of overlap of the wet non-contaminated samples with their contaminated counterparts. Considering these latter two cases, it therefore appears that PC1 characterized the contaminated nature of the samples while PC2 accounted for their wetness and dryness. Hence, the lower the PC1 scores, the more contaminated the samples were and the lower

the PC2 scores, the more wet the samples were and vice versa (Figure 3-5b, c). These results suggest that oil contamination accounts for most of the variability, hence most of the information, in the spectral dataset.

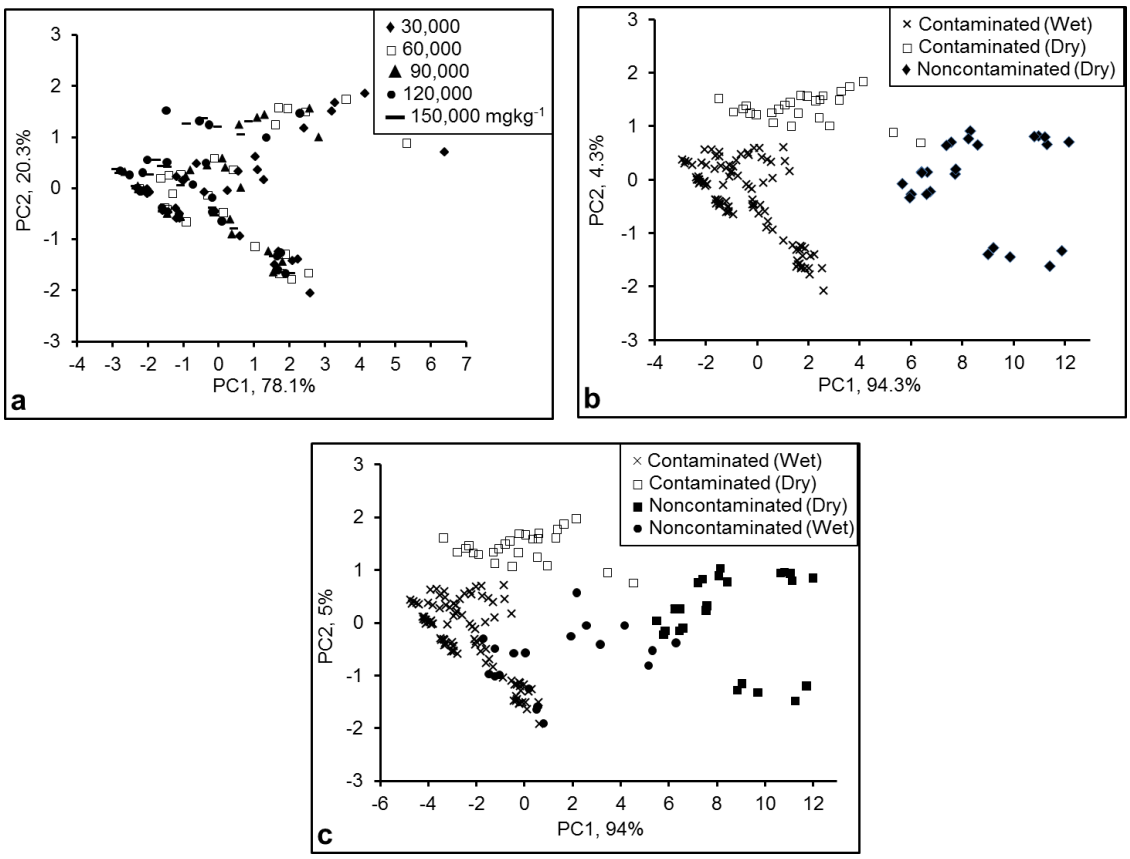


Figure 3-5 Principal component analysis scores plots of visible and near infrared diffuse reflectance spectra of overall dataset evaluated for diesel-contaminated and non-contaminated (wet and dry) soils

3.5 Conclusions

In the current study, results confirmed the following conclusions:

1. Soil diffuse reflectance decreased with increasing oil concentration, moisture content and clay content due to colour-associated variations in

the visible region, as well as water, mineralogy and hydrocarbon absorption in the NIR region.

2. The interaction of oil concentration, clay content and soil moisture content had significant ($p<0.05$) effects on the accuracy of PAH determination with vis-NIR spectroscopy but mostly with the interaction of the three design parameters combined.

3. Investigations with field-moist intact soils (mixed moisture) without a need for prior sample preparations may be possible with vis-NIR spectroscopy as good model performance was achieved (RMSE = 0.39 mg kg^{-1} and RPD = 1.93). Nevertheless, best model performance was achieved with the lowest oil concentration of 30,000 mg kg^{-1} (RMSE = 0.12 mg kg^{-1} and RPD = 3.20), least clay content of 9% (RMSE = 0.39 mg kg^{-1} and RPD = 1.99) and moisture of 20% (RMSE = 0.30 mg kg^{-1} and RPD = 2.53).

4. Prediction accuracy for PAH with vis-NIR spectroscopy for mixed moisture soils of validation r^2 of 0.89 and RPD of 2.75 implies an excellent model performance enabling quantitative applications. Nevertheless, we recommend that for potentially improved models, non-linear multivariate calibration techniques like artificial neural net-work may be applied.

5. Vis-NIR diffuse reflectance spectra may be used to predict and distinguish between PAH and TPH since distinct peaks identified

respectively around the 1,647 and 1,712 nm in the first overtone region of the NIR spectrum were peculiar to C–H stretching modes of ArCH linked to PAHs and saturated CH_2 and terminal CH_3 groups linked to TPH.

6. PCA may only be used to differentiate the samples into contaminated and non-contaminated as well as dry and wet groups as it could not be used to account for the different levels of oil contamination in the samples.

Chapter 4: Prediction of selected individual polycyclic aromatic hydrocarbons in soils by visible and near-infrared spectroscopy

4.1 Chapter summary

A rapid and cost-effective measurement method for PAH in soils will assist in mitigating the environmental impact of the soil contaminant by informing risk assessment and remediation. Vis-NIR spectroscopy and multivariate statistical analysis have shown a valuable synergy for rapid measurement of petroleum hydrocarbons in contaminated soils. This chapter investigated the potential application of vis-NIR spectroscopy (350–2500 nm) for the prediction of selected PAHs. A total of 150 artificially diesel-contaminated soil samples of five textural classes at various moisture contents (MC) were used. PLS regression analysis with full cross-validation was used to develop models to predict acenaphthylene (ACE), fluorene (FLU), phenanthrene (PHE), anthracene (ANT), and pyrene (PYR). Results showed that only PHE could be predicted with reasonable accuracy (ratio of prediction deviation, RPD = 2.0–2.32; root-mean-square error of prediction, RMSEP = 0.21–0.25 mg kg^{-1}; and coefficient of determination, r^2 = 0.65–0.83). The failure to develop robust models for ACE, FLU, ANT and PYR was because of low concentration ranges of the compounds in the soil samples (0.02–1.09 mg kg^{-1}). The mechanism of prediction was attributed to co-variation of PAH with other soil properties that have direct spectral responses in the NIR spectral range (e.g., clay and organic carbon). Overall, the results demonstrated that the vis-NIR spectroscopic method coupled with PLS regression analysis may be suitable for PAH concentrations higher than 1.09 mg kg^{-1} in soils.

4.2 Introduction

The concentration and compositional distribution of PAHs in an environmental sample is widely used to identify their origin or source (Hites and Gschwend, 1982; Yunker and MacDonald, 1995; Okoro and Ikolo, 2007). This is vital in tackling their proliferation in the environment by informing risk assessment and remediation. Conventional methods of quantifying PAHs in contaminated soils such as GC–MS are highly sensitive and specific (Brassington et al., 2010). But, as discussed, they are relatively expensive, involve time-consuming sample preparation protocols, and rely on the use of noxious extraction solvents that tend to pose a health risk to operators (Okparanma and Mouazen, 2013a). This has prompted increasing demand for alternative methods capable of overcoming most of those challenges, without necessarily having to trade-off instrument performance, to complement the conventional methods.

One of these emerging alternatives is the vis-NIR spectroscopy analysis. The TPH content in hydrocarbon-contaminated soils has been predicted by vis-NIR spectroscopy using various multivariate techniques (Graham, 1998; Malle and Fowlie, 1998; Malley et al., 1999; Forrester et al., 2010, 2013; Chakraborty et al., 2010, 2012; Schwartz et al., 2012). PAHs in hydrocarbon-contaminated soils have also been predicted by vis-NIR spectroscopic method using multivariate data analysis (Bray et al., 2010; Okparanma and Mouazen, 2013b). A recent review shows that the quality of TPH and PAH models have improved recently, suggesting a greater likelihood of using vis-NIR spectroscopy as a

screening tool for hydrocarbon-contaminated soils (Okparanma and Mouazen, 2013a).

It must be pointed out that because of the high variability in sample or site matrix, more research is still necessary to achieve the required level of accuracy prescribed by the industry and/or regulatory agencies for an on-line manufacturing system. This is because the models reported so far appear to be local models developed for specific sampling sites or individual soil types. As a result, such models lack robustness as their extrapolation is limited to the sampling sites or individual soil types for which the calibrations were developed. Therefore, to develop robust and stable models, global calibrations are essential. Global calibrations are achieved when all sources of variability in the calibration are taken into account. Nonetheless, issues concerning the change in stability of calibration models over time still exist. This may cause erroneous predictions due to several reasons including (among others) changes in the instrument response due to variations in temperature or pressure, non-standardisation of equipment and optical accessories, and changes in physical conditions of the samples (Macho and Larrechi, 2002). Interestingly, several correction strategies have recently been suggested by researchers, according to Macho and Larrechi (2002), to overcome the problem without a complete recalibration procedure. Some of the strategies proposed are transformation, robust models, and model updating (Macho and Larrechi, 2002).

At the moment, only one study could be found in the literature on the application of vis-NIR spectroscopy for the prediction of an individual PAH (e.g.,

benzo[a]pyrene) in contaminated urban soils. The study demonstrated good accuracy (average prediction accuracy = 78.9%) but moderate to high false positive rates (Bray et al., 2010). It is important to point out that benzo[a]pyrene is merely one of the sixteen PAHs classified as priority pollutants by the USEPA that require accelerated detection and removal from the environment because of their hazardous nature. The high false positive rates reported for benzo[a]pyrene by Bray et al. (2010) and the absence of literature for the remaining PAHs accentuate the need for further research on the application of vis-NIR spectroscopic method for the determination of PAHs in soils.

The objective of this chapter was to evaluate the prediction accuracy of five PAH compounds including ACE, ANT, PHE, FLU, and PYR in diesel-contaminated soils by vis-NIR spectroscopy and PLS regression analysis. This was evaluated for five clay contents, MC, and oil concentrations with artificially diesel-contaminated soil samples.

4.3 Materials and methods

4.3.1 Sample collection and treatment

The soil samples used in this study were obtained from the experimental farm of Cranfield University at Silsoe (52^{o}00'N, 0^{o}26'W) and Duck End Farm at Wilstead (52^{o}05'N, 0^{o}27'W), all in Central Bedfordshire, United Kingdom. The soil samples consist of five textural classes belonging to the Arenosols, Cambisols and Luvisols (Table 4-1). A total of 150 samples were given various treatments including drying, sieving, wetting, and contamination with diesel. The samples

were separated into five groups of 25 samples each according to their textural classes first, and then oven-dried at 105±5^{o}C for 24 h and sieved to <2 mm. Then each group was separated further into five subgroups consisting of 5 samples each according to moisture treatment levels (0, 5, 10, 15, and 20%, w/w). Finally, in each subgroup, each sample was contaminated with 852, 1136, 1420, 1705, and 1989 mm^{3} of diesel, respectively. The expected concentrations of the added diesel were 30,000; 60,000; 90,000; 120,000; and 150,000 mg kg^{-1} of soil, respectively. A sixth group consisting of 25 samples, 5 from each textural class, was also prepared and used as field-moist intact samples without drying, sieving, or wetting apart from diesel contamination. The weight of each sample was approximately 25 g.

Table 4-1 Details of the sampling sites and selected physicochemical properties of the soil samples

Soil ID	Farm name	Farm Location [a]	WRB Order [b]	Soil texture fractions (%)			USDA Soil Textural Classification	Soil organic carbon (%)	Moisture content of field-moist samples (%)	*n* [c]
				Sand	Silt	Clay				
A	College Farm, Silsoe	52°00'30"N, 0°26'54"W	Arenosols	74	17	9	Sandy-loam	0.76	15.41	11.8
B	College Farm, Silsoe	52°00'32"N, 0°26'49"W	Cambisols	61	19	20	Clay-loam	1.89	9.04	10.6
C	College Farm, Silsoe	52°00'01"N, 0°26'36"W	Cambisols	32	33	35	Sandy-clay-loam	2.04	15.05	17.2
D	College Farm, Silsoe	52°00'34"N, 0°25'60"W	Cambisols	52	22	26	Loamy-sand	1.15	16.13	22.6
E	Duck End Farm, Wilstead	52°05'08"N, 0°27'10"W	Luvisols	9	17	74	Clay	1.63	11.91	45.4

[a] Google Earth

[b] World Reference Base (WRB) classification

[c] *n*, Dexter index = clay content/soil organic carbon (Dexter et al., 2008)

4.3.2 Reference laboratory analysis of soil physicochemical properties

Soil particle size distribution was determined by laser diffraction with a Mastersizer2000® (Malvern Instruments, Worcestershire, UK) coupled to a HydroMu® dispersing unit (Malvern Instruments, Worcestershire, UK). We used the United States Department of Agriculture (USDA) soil textural classification scheme to determine the soil textural classes on the basis of percent clay, silt, and sand (Table 4-1). Soil moisture content (w/w), on dry basis, was determined by the oven-drying method at 105±5^{o}C for 24 h (Table 4-1). Soil OC (Table 4-1) was determined by the standard operating procedures of the Cranfield University based on British Standard 7755 Section 3.8:1995 (British Standard Institute, 1995) with a Vario EL III Analyzer (Elementar Analysensysteme, Hanau, Germany).

We extracted PAH compounds from the diesel-spiked soil samples by the sequential ultrasonic solvent extraction method (Risdon et al., 2008) with a mixture of dichloromethane (DCM) and hexane (1:1). Clean-up of the PAH extract was carried out with a ~0.6 g Florisil (Fisher Scientific Ltd., Loughborough, UK), on glass wool, micro-scale column prewashed with DCM. PAH analysis was carried out with a 6890N Network GC System (Agilent Technologies Inc., USA) coupled to a 5973 Network MSD (Agilent Technologies Inc., USA) operated at 70 eV in positive ion mode. Each PAH compound was quantified by the internal standard method. The instrument was calibrated beforehand with a 5-level calibration solution mix. The calibration solution mix

was made up with EPA 525 PAH Mix-A standard solution (Sigma-Aldrich Co. Ltd., Dorset, UK) and deuterated PAH internal standard solutions – Naphthalene-d_8, Anthracene-d_{10}, Chrysene-d_{12}, and Perylene-d_{12} (Sigma-Aldrich Co. Ltd., Dorset, UK). Quantification of each PAH was performed by integrating the peak at specific m/z (mass-to-charge ratio) by means of MSD ChemStation. Major hydrocarbon compounds were identified on the basis of their retention time, and by comparing them to those of analytical standards. Matrix spikes, duplicates, solvent and method blanks were also analysed as quality control samples.

4.3.3 Optical scanning of soil samples

Diffuse reflectance spectra were taken from the soil samples with a mobile fibre-optic LabSpec2500® vis-NIR spectrophotometer (350–2500 nm) (Analytical Spectral Devices Inc., USA) coupled to a high-intensity probe (Analytical Spectral Devices Inc., USA). The spectrophotometer has one Si array (350–1000 nm) and two Peltier-cooled InGaAs detectors (1000–1800 nm and 1800–2500 nm). Spectral sampling interval of the instrument was 1 nm across the entire spectral range. However, the spectral resolution was 3 nm at 700 nm and 10 nm at 1400 and 2100 nm. The high-intensity probe has a built-in light source made of a quartz-halogen bulb of 3000K. The light source and detection fibres are assembled in the high-intensity probe enclosing a 35-degree angle.

Before wetting, five dry-graded non-contaminated samples from each of the five textural classes (25 samples) were scanned and labelled as dry non-contaminated set. Before diesel addition, four wet-graded samples from each of

the five textural classes (20 samples) of 5, 10, 15, and 20% MC were scanned and labelled as wet non-contaminated set. Then, the remaining 105 contaminated wet-graded and field-moist intact samples were scanned. The 25 dry non-contaminated samples were then contaminated, scanned and labelled as dry contaminated set. Finally, the 20 wet non-contaminated samples were also contaminated and scanned. Before the soil samples were scanned and at intervals of 30 min, white-referencing with a Spectralon disc of approximately 100% reflectance was carried out to optimize the instrument. Measurements were taken from the soil sample, tightly packed and levelled in a cuvette, at three equidistant positions, 120^{o} apart. Each sample was scanned nine times, three times per spot, and averaged for spectral preprocessing and multivariate analysis.

4.3.4 Spectral preprocessing

Spectral preprocessing aims to reduce spurious peaks that do not contain physical or chemical information and to correct physical scatter effects. Perceived noise at the extremes of the spectrum (i.e., at 350–449 nm and 2451–2500 nm) were removed because of low instrument sensitivity at these wavelengths. Spectral truncation was followed by smoothing by averaging adjacent 5-nm wavelengths to reduce the impact of noise. Thus, the final wavelength range used for modelling was 452–2450 nm consisting of 401 wavelengths in the range. Afterward, the spectra were first derivative transformed by Savitzky-Golay method to remove additive baseline shift. This

was implemented for all sample subset with the Unscrambler 9.8 (CAMO Software, Woodbridge, NJ).

4.3.5 PLS regression analysis

Before calibration, spectral reflectance (R) was transformed to the logarithm of the relative intensity (1/R) or absorption (Naes et al., 2002). The PLS regression analysis is a bilinear modeling method where information in the original x-data is projected onto a small number of underlying ("latent") variables called PLS components. The y-data are actively used in estimating the "latent" variables to ensure that the first components are those that are most relevant for predicting the y-variables. Interpretation of the relationship between x- and y-data is then simplified as this relationship is concentrated on the smallest possible number of components (latent variables, LV). More detailed information about the PLS can be found in Martens and Naes (1989). We used PLS regression analysis with full cross-validation to relate the variation in a single-component variable (e.g., PAH) to the variation in a multi-component variable (e.g., wavelength) by Unscrambler 9.8 (CAMO Software, Woodbridge, NJ). In the current study, two categories of PLS models were developed. In the first category, models were developed for the five selected PAHs, namely, ACE, FLU, PHE, ANT, and PYR for each level of MC, clay content (9, 20, 26, 35, and 74%), and oil concentration. To do this, the vis-NIR spectra were separated into subgroups of 25 spectra according to treatment levels. Then the 25 spectra and chemical variables for each subgroup were used together to develop a PLS regression model for each PAH at each level of MC, clay content, and oil concentration. Up

to twelve LVs were considered, and the optimal number of LVs for future predictions was determined on the basis of the number of factors at the first local minimum (Naes et al., 2002). In the second category, general models were developed to predict each PAH compound with the entire 150 samples. In this category, the entire dataset was randomly separated into calibration (76%) and prediction (24%) sets. The ratio of calibration/prediction samples was chosen to ensure that each sample subset (group) was equally represented in the prediction set by randomly choosing six samples from each sample subset. PLS regression analysis with full cross-validation was carried out with the calibration set. The prediction set was then used to test the prediction accuracy of the calibration models. During model calibration in the second model category, potential outliers were identified on the basis of their influence on the X–Y relationship. Spectra that differed from the reference by three times the standard deviation of the predicted residuals were eliminated from the calibration dataset (Tekin et al., 2012). Model quality was statistically analysed by the root-mean-square error (RMSE) of cross-validation and prediction, ratio of prediction deviation (RPD) (i.e., the ratio of standard deviation of laboratory-measured sample concentration to the RMSE), and corresponding coefficient of determination (r^2) (Naes et al., 2002). Model prediction ability was categorised based on the following criteria: excellent if RPD > 2.0, almost good if 1.4 ≤ RPD < 2.0, and unreliable if RPD < 1.4 (Chang et al., 2001).

4.4 Results and discussion

4.4.1 Calibration models of PAH compounds

For the pre-treated (i.e., dried, sieved, and wetted) soil samples, most PLS models for ACE, FLU, ANT, and PYR have RPDs ranging from <1.0 to <1 (Tables not shown) demonstrating very poor to poor quality (Chang et al., 2001; Viscarra Rossel et al., 2006a). Table 4-2 summarizes the statistics of the PLS models for PHE. The quality of more than thirty percent of the models are classified as very good and excellent (RPD >2.0) signifying the possibility for quantitative applications (Chang et al., 2001; Viscarra Rossel et al., 2006a). For PHE, PLS model developed for 90,000 $mgkg^{-1}$ oil treatment level by first derivative transformed spectra has the best quality (RPD = 3.88, RMSE = 0.1379 $mgkg^{-1}$, and r^2 = 0.93) as compared to others (Table 4-2). For the field-moist intact soil samples, model statistics for all five PAH are summarized in Table 4-3. Apparently, only PLS models for PHE have excellent quality (RPD = 2.85–3.18). Two major factors were responsible for the failure to develop robust models in this model category, particularly for ACE, FLU, ANT, and PYR. These are the small number of sample set used for PLS analysis and the low initial concentration range of the PAHs in the samples. In this study, reference PAH analysis shows relatively low concentrations of ACE, FLU, ANT, and PYR in the soil samples (0.02–1.09 $mgkg^{-1}$) (Table 4-4). But, relatively high levels of PHE (0.58–2.49 $mgkg^{-1}$) are detected in the samples (Table 4-4). This is perhaps the reason for the smaller PLS correlations observed for ACE, FLU, ANT, and PYR compared to PHE. Besides, sample concentrations are not uniformly distributed

over the working concentration range of 0.02–2.80 $mgkg^{-1}$ (i.e., range of the calibration curve). This result is corroborated in part by reports that models for the prediction of total nitrogen and OC developed by vis-NIR spectroscopy with larger dataset resulted in smaller RMSE values compared to models developed with smaller dataset (Kuang and Mouazen, 2012). These results demonstrate that the vis-NIR spectroscopic method may be suitable for evaluating PAH concentrations higher than some 1.09 $mgkg^{-1}$ in soils.

Nonetheless, the encouraging result obtained for PHE raises two important research questions. The first question is what type of physicochemical characteristics influences fate and transport and absorption–reflection of PAHs in contaminated agricultural soils? The second question is how can vis-NIR spectroscopy be used successfully in broader environmental applications? In the following subsection, we answer the first question within the scope of the current study. The answer to the second question underlines the need for further research on the field application of the vis-NIR method. The field application of the vis-NIR method is the focus of the remaining two chapters.

Table 4-2 Summary of calibration results for phenanthrene obtained by partial least-squares (PLS) cross-validation analysis carried out with 25 samples for various concentrations of diesel, and moisture and clay contents.

Treatment level	Reflectance spectra					First derivative spectra				
	r^2	RMSE ($mgkg^{-1}$)	SD	RPD	LV	r^2	RMSE ($mgkg^{-1}$)	SD	RPD	LV
Diesel Conc. ($mgkg^{-1}$)										
30,000	0.86	0.1067	0.2958	2.77	2	0.84	0.1169	0.2958	2.53	2
60,000	0.75	0.1810	0.3726	2.06	2	0.74	0.1891	0.3726	1.97	2
90,000	0.89	0.1722	0.5354	3.11	4	0.93	0.1379	0.5354	3.88	4
120,000	0.50	0.3631	0.5144	1.42	2	0.46	0.3790	0.5144	1.36	2
150,000	0.81	0.1967	0.4590	2.33	2	0.77	0.2179	0.4590	2.11	2
Moisture Content (%)										
0	0.79	0.2190	0.5137	2.35	5	0.72	0.2464	0.4758	1.93	3
5	0.42	0.4470	0.4948	1.11	6	0.71	0.2632	0.4948	1.88	4
10	0.50	0.3493	0.4948	1.42	5	0.35	0.3999	0.4948	1.24	3
15	0.53	0.3420	0.4948	1.45	5	0.67	0.2768	0.4948	1.79	3
20	0.80	0.2345	0.4948	2.11	10	0.69	0.2712	0.4948	1.82	4
Clay Content (%)										
9	0.78	0.2290	0.4948	2.16	7	0.70	0.2668	0.4948	1.92	5
20	0.65	0.2878	0.4948	1.72	2	0.72	0.2576	0.4948	1.75	3
26	0.70	0.2685	0.4948	1.84	3	0.69	0.2684	0.4948	1.84	1
35	0.56	0.3247	0.4948	1.52	3	0.56	0.3234	0.4948	1.53	2
74	0.76	0.2296	0.4948	1.85	4	0.67	0.2350	0.4948	1.81	2

LV, Latent variable; RMSE, Root-mean-square error; RPD, Ratio of prediction to deviation; SD, Standard deviation

Table 4-3 Summary of calibration results of partial least-squares (PLS) cross-validation analysis carried out with 25 samples for polycyclic aromatic hydrocarbons (PAHs) in field-moist intact soil samples (moisture content = 9.04–16.13%; clay content = 9–74%; diesel concentration = 30,000–150,000 $mgkg^{-1}$).

PAH	Raw spectra					First derivative of spectra				
	r^2	RMSE ($mgkg^{-1}$)	SD	RPD	LV	r^2	RMSE ($mgkg^{-1}$)	SD	RPD	LV
ACE	0.20	0.08	0.09	1.21	1	<0.2	0.09	0.09	1.08	1
FLU	0.59	0.14	0.24	1.69	5	0.73	0.12	0.24	1.94	4
PHE	0.90	0.16	0.52	3.18	6	0.86	0.18	0.52	2.85	6
ANT	0.26	0.05	0.06	1.23	2	0.70	0.03	0.06	1.73	6
PYR	0.63	0.01	0.01	1.72	4	0.62	0.01	0.01	1.75	2

ACE, Acenaphthylene; ANT, Anthracene; FLU, Fluorene; LV, Latent variable; PHE, Phenanthrene; PYR, Pyrene; RMSE, Root-mean-square error; RPD, Ratio of prediction deviation; SD, Standard deviation

Table 4-4 Statistics of the chemical analysis result for polycyclic aromatic hydrocarbons in the diesel contaminated soil samples

PAHs	Variable statistics					
	Number of samples	Min. ($mgkg^{-1}$)	Max. ($mgkg^{-1}$)	Range ($mgkg^{-1}$)	Mean ($mgkg^{-1}$)	SD
ACE	150	0.02	0.39	0.37	0.16	0.09
FLU	150	0.36	1.09	0.73	0.60	0.23
PHE	150	0.58	2.49	1.91	1.18	0.50
ANT	150	0.04	0.43	0.39	0.12	0.84
PYR	150	0.02	0.05	0.03	0.03	0.01

ACE, Acenaphthylene; ANT, Anthracene; FLU, Fluorene; LV, Latent variable; PHE, Phenanthrene; PYR, Pyrene; SD, Standard deviation

4.4.2 Accuracy of prediction of PAH compounds

In this model category, the entire 150 vis-NIR spectra were combined including both the treatments and levels in the overall model to predict the five PAH compounds. The entire dataset used in this model category will enable exploring the influence of limited number of samples on PLS models and their prediction capability. Three outliers were eliminated from the calibration dataset. Table 4-5 summarizes the statistical results of PLS models in cross-validation and prediction sets for the prediction of PAHs. While other PAH–PLS regression models use between one and three LVs as recommended (CAMO Software, 2012), the PHE models use six LVs for the first derivative and ten LVs for the reflectance models (Table 4-5). Nonetheless, these are comparable to the range of 6–8 LVs reported for the prediction of saturates, aromatics, resins, and asphaltenes (SARA) fractions in crude oil (Aske et al., 2001). The means and standard deviations for the calibration and prediction dataset are similar, suggesting that the prediction models are not skewed. Of the five PAHs, only

PHE could be predicted with reasonable accuracy (RPD = 2.00–2.32, RMSE = 0.21–0.25 mgkg^{-1}, and r^2 = 0.65–0.83) (Table 4-5 and Figure 4-1). This does not suggest that PHE could be used to indicate the presence of other PAH compounds. But the results confirm that one of the reasons for the failure to develop robust models with vis-NIR spectroscopic method in this study was because of low initial PAH concentrations in the soil samples. Low concentration ranges of OC and total nitrogen in soils have also been reported to affect their prediction accuracy by vis-NIR spectroscopy (Kuang and Mouazen, 2011). This large concentration range may explain why PHE was predicted reasonably well as compared to other PAH compounds.

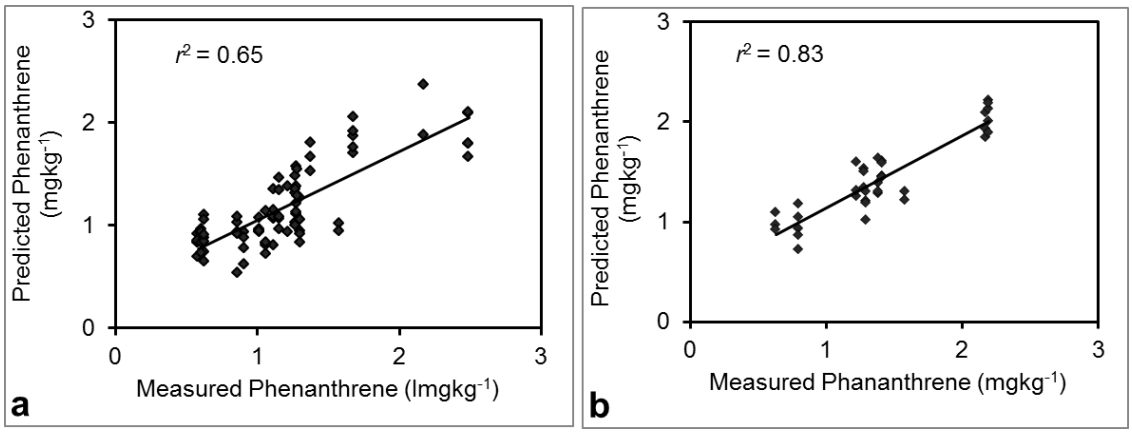


Figure 4-1 Scatter plots of gas chromatography–mass spectroscopy (GC–MS)-measured vs. visible and near-infrared (vis-NIR)-predicted values of phenanthrene developed by partial least-squares (PLS) regression analysis with raw spectra and validated

(a) with cross-validation (114 samples) and (b) with prediction (36 samples) sets. The soil samples contain various amounts of moisture (0–20%), clay (9–74%), and diesel (30,000–150,000 mg kg^{-1}).

Table 4-5 Sample statistics and results of partial least-squares (PLS) models for the prediction of selected polycyclic aromatic hydrocarbons (PAHs) in cross-validation and prediction data sets for diesel-contaminated soil samples by visible and near infrared (vis-NIR) spectroscopy.

PAHs	Variable statistics						Model quality							
							Reflectance spectra				First derivative spectra			
	No. of samples	Min. ($mgkg^{-1}$)	Max. ($mgkg^{-1}$)	Mean ($mgkg^{-1}$)	SD	No. of outliers removed	r^2	RMSE ($mg\ kg^{-1}$)	LV	RPD	r^2	RMSE ($mgkg^{-1}$)	LV	RPD
Cross-validation set (76%)														
ACE	114	0.02	0.38	0.16	0.09	3	< 0.30	0.09	1	< 1.40	< 0.30	0.10	1	< 1.40
FLU	114	0.36	1.09	0.60	0.24	3	< 0.30	0.23	3	< 1.40	< 0.30	0.24	1	< 1.40
PHE	114	0.58	2.49	1.18	0.48	3	0.65	0.28	10	1.71	0.62	0.30	6	1.63
ANT	114	0.04	0.43	0.12	0.09	3	0.32	0.07	2	< 1.40	0.38	0.07	2	< 1.40
PYR	114	0.02	0.05	0.03	0.01	3	< 0.30	0.01	1	< 1.40	< 0.30	0.01	1	< 1.40
Prediction set (24%)														
ACE	36	0.04	0.32	0.18	0.09	N/A	< 0.30	0.08	1	< 1.40	< 0.30	0.09	1	< 1.40
FLU	36	0.38	0.99	0.73	0.20	N/A	< 0.30	0.18	3	< 1.40	< 0.30	0.20	1	< 1.40
PHE	36	0.63	2.20	1.40	0.50	N/A	0.83	0.21	10	2.32	0.65	0.25	6	2.00
ANT	36	0.06	0.24	0.15	0.06	N/A	0.64	0.05	2	< 1.40	0.64	0.05	2	1.40
PYR	36	0.03	0.05	0.04	0.01	N/A	< 0.30	0.01	1	< 1.40	< 0.30	0.01	1	< 1.40

ACE, Acenaphthylene; ANT, Anthracene; FLU, Fluorene; LV, Latent variable; N/A, Not applicable; PHE, Phenanthrene; PYR, Pyrene; RMSE, Root-mean-square error; RPD, Ratio of prediction deviation; SD, Standard deviation

4.4.3 Regression coefficients

In PLS regression analysis, regression coefficients are approximations of model parameters resulting from the linear combination of the predictors. The regression coefficients plot is used to identify important wavelengths for the prediction of relevant soil properties. In this study, we derived bar plots of regression coefficients versus wavelength after PLS regression analysis using first derivative spectra with 114 calibration samples (Figure 4-2). In the bar plots, the absolute value of each bar at each wavelength is equivalent to the intensity of the regression coefficient. The intensity of the regression coefficients shows the relative importance of the X-variables in the model. Variables with large coefficient play an important role in the model; a positive coefficient shows a positive link to the response, and a negative coefficient shows a negative link (CAMO Software, 2012). The plots of regression coefficients versus wavelength over the modelling wavelength range of 452–2450 nm show that the intensities of the regression coefficients vary considerably in magnitude for each PAH. But the positions of significant wavelengths are largely similar (Figure 4-2). Larger regression coefficients were observed for PHE (38.12–119.92) compared to other PAHs (0.03–1.45).

In the bar plots, absorption bands around 1647 and 1712 nm indicate the presence of petroleum-based hydrocarbons in the diesel-contaminated soils (Osborne et al., 1993; Workman and Weyer, 2008). These absorptions bands are linked to overtones of C–H stretching modes of ArCH, saturated CH_2 and terminal CH_3 groups (Osborne et al., 1993; Workman and Weyer, 2008).

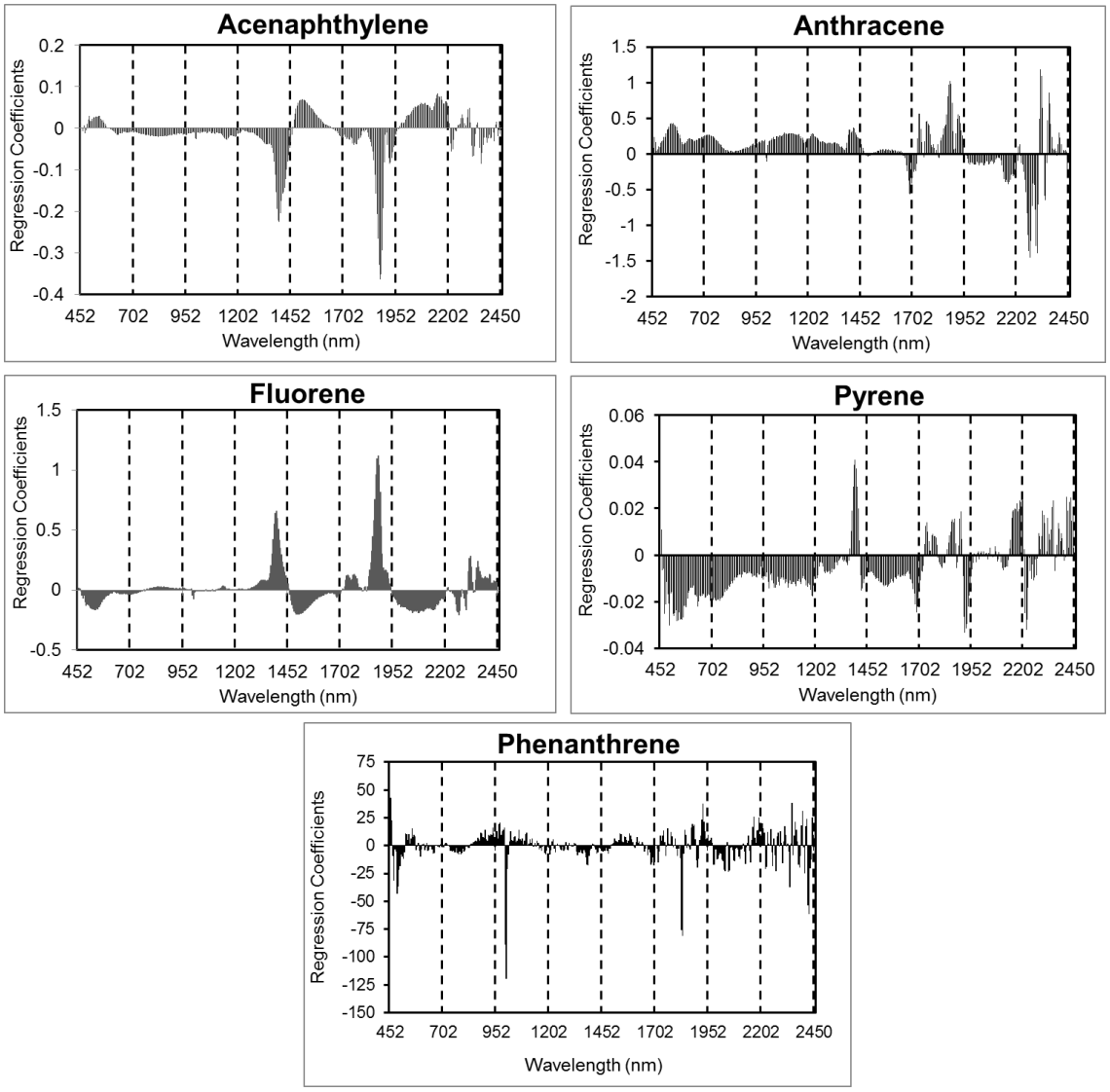


Figure 4-2 Plots of regression coefficients vs. wavelength derived from partial least squares (PLS) regression analysis with first derivative of 114 sample spectra of diesel-contaminated soils.

The spectral signature of ArCH around 1647 nm in the first overtone region suggests the presence of PAHs, which are usually found in diesel. The signatures of saturated CH_2 and terminal CH_3 groups around 1712 nm are unique to the TPH content (Forrester et al. 2010). These are in agreement with the well-known fact that the spectra of hydrocarbon derivatives originate mainly from combinations or overtones of ArCH functional groups or C–H stretching

modes of saturated CH_2 and terminal CH_3 groups (Aske et al., 2001; Forrester et al., 2010). This distinction in spectral signatures may be useful in separating vis-NIR signals for PAHs from TPH signals in diesel-contaminated soils.

The relatively large regression coefficients linked to moisture, clay minerals, and OC absorption bands shown in the regression coefficients plots for studied PAHs (Figure 4-2) is indicative of the influence of these soil variables on the spectral absorption of PAHs in the NIR band. This may also suggest indirect measurement capability through co-variation with these soil properties in the NIR range. These soil properties are also known to have direct spectral responses in the NIR range (Stenberg et al., 2010; Kuang et al., 2012). The absorption around 497 nm (Figure 4-2) is linked to blue colour absorption band in the visible range of the spectrum. Reports have shown that changes in soil colour are also linked to changes in the amounts of water and/or diesel in the soil (Mouazen et al., 2005a; Okparanma and Mouazen, 2013b). Soil becomes darker with increase in water content and diesel concentration resulting in overall increase in absorption or decrease in reflection (Mouazen et al., 2005a; Okparanma and Mouazen, 2013b). The blue absorption band also deepens (Mouazen et al., 2005a). Absorption bands linked to water were also observed around 950, 1450, and 1950 nm in the NIR spectrum (Figure 4-2). In the NIR range, O–H stretching modes of water are responsible for the absorptions around 950, 1450, and 1950 nm in the O–H second and first overtones, and combinations band respectively (Whalley and Stafford, 1992). Absorption bands were also observed in the range of 2150–2450 nm (Figure 4-2). Absorption

features linked to metal-OH bend plus O–H stretch combinations around 2200 nm and 2300 nm are characteristic of clay minerals (Clark et al., 1990; Viscarra Rossel et al., 2006b). Absorption features attributed to long-chain C–H+C–H and C–H+C–C stretch combinations around 2150 nm and 2212 nm are unique to soil organic matter (Forrester et al., 2010).

As stated, PAHs exhibit strong hydrophobicity and largely sorb to the organic matter or OC in the soil (Huang et al., 2003). Dexter et al. (2008) suggested that it is not the total amount of OC that controls soil physical behaviour but the amount of complexed and non-complexed organic carbon (COC and NCOC, respectively). Dexter et al. (2008) further stated that the NCOC is present in soil only if the Dexter index, n = clay/OC, is less than 10. It is also reported that the NCOC has a higher sorption affinity for PAH than the COC (Soares et al., 2013). In this study, the Dexter n for each soil texture was deduced before contamination using their OC and clay contents shown in Table 4-1. Table 4-1 shows that only COC was present (n >10) in all soil texture. As a result, a significant proportion of PAH is not sorbed to soil because of the absence of NCOC and low PAH sorption affinity of COC. The intensity of regression coefficients shown in the bar plots, hence NIR spectral signal, reflect the amount of sorbed PAH as dictated by COC. On the other hand, this implies that the presence of NCOC in soil, when n <10, immobilizes PAH within the soil matrix, which might cause increased soil absorption and reduced soil reflection. Therefore, the inter-relationship between clay and OC may be useful for evaluating the presence of PAH in soils. At the moment, the tendency for clay–

OC interactions to partly dictate the behaviour of PAH sorption to soil is not well-known and requires further investigation by conventional laboratory methods (Soares et al., 2013).

4.5 Conclusions

In the current study, results confirmed the following conclusions: (1) of the five PAHs investigated, only phenanthrene could be predicted with reasonable accuracy (RPD = 2.0–2.32, RMSE = 0.21–0.25 mg kg^{-1}, and r^2 = 0.65–0.83); (2) the failure to develop robust models for acenaphthylene, fluorene, anthracene, and pyrene was because of their low concentration ranges (0.02–1.09 mg kg^{-1}) in the samples. Another reason was that sample concentrations were not uniformly distributed over the working concentration range of 0.02–2.80 $mgkg^{-1}$; (3) the vis-NIR spectroscopic method may be suitable for evaluating PAH concentrations above 1.09 $mgkg^{-1}$; and (4) the inter-relationship between clay and OC influenced the intensity of NIR spectral signal of sorbed PAH in soil. Therefore, the clay–OC interactions may be useful for evaluating the presence of PAH in soils by vis-NIR spectroscopy. However, the unique spectral signal of sorbed PHE observed in the current study requires thorough investigation in relation to decision-making for risk assessment and remediation of contaminated sites.

The bulk of the remaining two chapters of this thesis will be concerned with attempts at expanding the implementation of the findings of the current work for broader environmental applications; the methodology was applied on real oil-contaminated soils collected from contaminated agricultural fields in the Niger

Delta province of Nigeria. Succinctly, the study in Nigeria aims to predict and map PAH contamination in Nigerian soils to inform risk assessment and remediation. If successful, this would enable rapid collection of large dataset at relatively low cost to explore the spatial variation in contaminated soils. This can also reduce the likely impacts of future pollution incidents through early identification of pollutants and/or by incorporation into verification testing of remediation treatment.

Chapter 5: Analysis of petroleum-contaminated soils by diffuse reflectance spectroscopy and sequential ultrasonic solvent extraction–gas chromatography

5.1 Chapter summary

In this study, we demonstrate that partial least-squares regression analysis with full cross-validation of spectral reflectance data estimates the amount of PAH in petroleum-contaminated tropical rainforest soils. We applied the approach to 137 field-moist intact soil samples collected from three oil spill sites in Ogoniland in the Niger Delta province (5.317°N, 6.467°E) of Nigeria. We used sequential ultrasonic solvent extraction–gas chromatography as the reference chemical method. We took soil diffuse reflectance spectra with a mobile fibre-optic visible and near-infrared spectrophotometer (350–2500 nm). Independent validation of combined data from studied sites showed reasonable prediction precision (root-mean-square error of prediction = 1.16–1.95 mg kg^{-1}, ratio of prediction to deviation = 1.86–3.12, and validation r^2 = 0.77–0.89). This suggests that the methodology may be useful for rapid assessment of the spatial variability of polycyclic aromatic hydrocarbons in petroleum-contaminated soils in the Niger Delta to inform risk assessment and remediation.[3]

5.2 Introduction

In the analysis of petroleum-contaminated soils for PAH, gas chromatography–mass spectrometry (GC–MS) is mostly preferred because of their relative selectivity and sensitivity (Brassington et al., 2010; Wang and Fingas, 1995).

[3] Reprinted after the permission of Elsevier Ltd

But as discussed, GC–MS method is relatively expensive, involves time-consuming sample preparation protocols, and relies on the use of noxious extraction solvents. On the other hand, rapid measurement techniques based on vibrational spectroscopy, particularly those employing diffuse reflectance sample interfaces are simpler, safer, faster, and cost-effective (Okparanma and Mouazen, 2013a; Forrester et al., 2013).

Techniques employing diffuse reflectance sample interfaces include FTIR and vis-NIR spectroscopies. Diffuse reflectance in the IR spectrum came to limelight in 1976, and became known as diffuse reflectance infrared Fourier transform (DRIFT) spectroscopy (Willey, 1976). Before its advent in the IR spectrum, diffuse reflectance has been commonly available in the visible and NIR spectrum (Willey, 1976). Early applications of DRIFT have been widely reported (e.g., Fuller and Griffiths, 1978). Both DRIFT and vis-NIR are very much similar in many ramifications, but maybe more importantly are the differences between the two of them. Both techniques are now available in portable devices that can be deployed for in-field measurements without sample preparation. In terms of capital equipment cost, the prices of their latest portable models are comparable. The commercially available DRIFT (portable) system developed by Forrester et al. (2011), the 4100 ExoScan FTIR (Agilent Technologies, CA, USA), currently costs $61 301 while the latest portable model of VIS–NIR analyser, the LabSpec 5000 (Analytical Spectral Devices, Inc., CO, USA), costs $61 078 (Okparanma and Mouazen, 2013a). The former system predicted TPH with root-mean-square error of 903 mg kg^{-1} in cross-validation for TPH ranging

from 0 to 11000 mg kg^{-1} (Forrester, 2010). On the other hand, vis-NIR spectroscopy has been successfully used for on-line (tractor-mounted) measurement of soil properties (e.g., Mouazen et al., 2007a; Kuang and Mouazen, 2013), which opens the possibility for on-line measurement of PAHs using the technology. So far, no reports about the use of DRIFT for on-line measurement can be found in the literature. Although a tractor-mounted vis-NIR technology may not be realistic for application in Nigeria because of the terrain in which these oil spills occur, a hand-held or 'knapsack' variant of the technology is feasible.

As stated, vis-NIR spectroscopy has been used in conjunction with various multivariate analytical techniques to predict TPH in contaminated soils (Chakraborty et al., 2010, 2012; Forrester et al., 2010; Graham 1998; Malley et al., 1999; Schwartz et al., 2012). However, studies on the application of the methodology to predict PAH in contaminated soils are few in literature (e.g., Bray et al., 2010; Okparanma and Mouazen, 2013b). High false positive rates were reported for PAH prediction by Bray et al. (2010), which underscore the need for further research on the application of the approach. On the other hand, models reported by Okparanma and Mouazen (2013b) for PAH prediction appear to be local models. This limits extrapolation of the models to studied soil types. Moreover, none of these two studies considered contaminated tropical rainforest soils in their applications. Therefore, to set up a vis-NIR-based methodology to model vis-NIR spectra for broader environmental application, it

is essential to expand the approach to cover a wider range of soil types and environmental conditions.

The objective of this chapter was to evaluate the performance of vis-NIR-based methodology in the prediction of PAH in contaminated tropical rainforest soils. To do this we used sequential ultrasonic solvent extraction–gas chromatography (SUSE–GC) as the benchmark method. Sites investigated in this study are in the tropical rainforests of the oil-rich Niger Delta province in Nigeria. To the best of our knowledge this study is the first attempt to adopt vis-NIR spectroscopy for the determination of hydrocarbons in contaminated arable lands in Nigeria.

5.3 Materials and methods

5.3.1 Brief description of the study area

Available geological data for the study area show that Ogoniland is located within the Niger Delta basin with soils that are broadly classified as tropical rainforest soils, which occur in the southern part of Nigeria (SPDC, 2006). Niger Delta province covers a total land area of 70000 km^2 (Niger Delta Environmental Survey, 1995). According to the United States Department of Agriculture (USDA) soil taxonomic order, soils in the Niger Delta belong to the Oxisols. In this study, soils were collected from within the shallow geology (top-soils) of Ogoniland, which consists of sandy clay (UNEP, 2011). Typical ranges of soil nutrients concentrations at all soil depths reported for similar ecosystems in the Niger Delta show that nitrate-nitrogen range from 0.01–1.96 $mgkg^{-1}$,

phosphorous range from 0.21–6.92 mg kg^{-1}, sulphate range from 0.20–10.91 mg kg^{-1} and soil pH range from 5.2–6.4 (SPDC, 2006). Total organic carbon range from 3.63–4.11% (Tanee and Albert, 2011). In this study, we collected soil samples from three oil spill sites located at Baraboo (4.652°N, 7.249°E), Bomu 1 (4.662°N, 7.277°E), and Bomu 2 (4.662°N, 7.249°E) all in K-Dere, Ogoniland in the Niger Delta province of Nigeria (Figure 5-1).

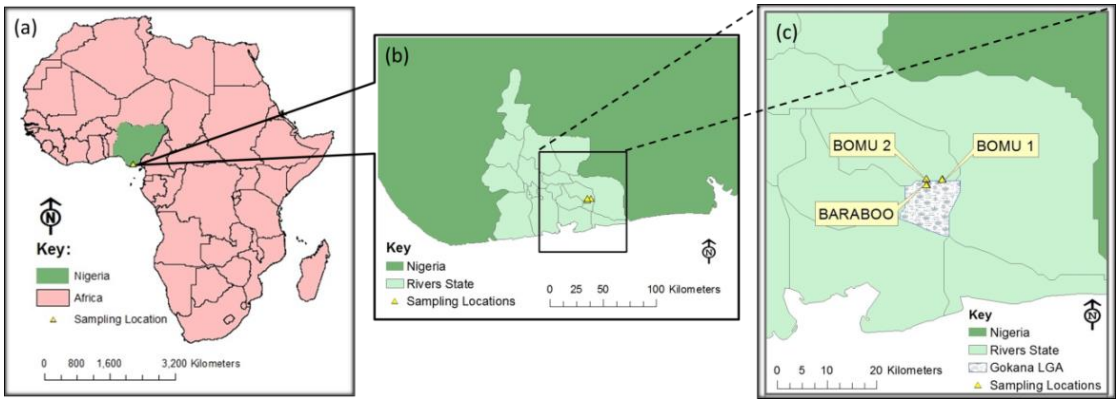


Figure 5-1 Sampling locations

(a) Nigeria, (b) Rivers State in the Niger Delta province, and (c) Gokana Local Government Authority in Ogoniland (Datum and Projection: GCS WGS 1984. Shape files source: ESRI®, CA, USA).

5.3.2 Sample collection

By targeted sampling method, we collected soil samples from the top 10–15-cm depth in plastic containers and preserved them in a cooler containing ice blocks until shipment (Figure 5-2).

Figure 5-2 Soil sampling sites and methods

(a) Bomu 1 spill site has no vegetation because it was recently remediated (b) Baraboo spill site, (c) Bomu 2 spill site, (d) Holes, 10–15cm deep, were dug with hand trowel, (e) soils from different depth were homogenised with a spatula and composite sample collected in a 28 mL vial, (f) the vial was properly labelled, sealed, collected in black polyethylene bags and stored in a cooler containing ice-blocks until shipment to the UK.

We adopted targeted sampling method because of the need to cover as much of the potentially contaminated spots in the sites as possible. Sample management was done according to the standards of the Nigerian Government's Department of Petroleum Resources (2002). We carried out both reference SUSE–GC analysis and vis-NIR optical scanning of soil samples in the Environmental Analytical Facility of Cranfield University, United Kingdom.

5.3.3 Reference chemical analysis of PAHs

We used SUSE–GC method to determine the concentrations of PAH compounds in the soil samples as described by Risdon et al. (2008). But, in place of acetone we used dichloromethane (Rathburn Chemicals Ltd., Walderburn, UK). Analysis of PAHs was carried with a 6890N Network GC system coupled to a 5973 Network MSD (Agilent Technologies Inc., USA) operated at 70 eV in positive ion mode. Each PAH compound was quantified by the internal standard method. In this study, we adopted the same limit of quantitation (LOQ) of 0.02 mg kg^{-1} routinely used by laboratories in Nigeria for PAH quantitation in Nigerian soils because our soil samples were collected from Nigeria. LOQ is the lowest concentration at which an analyte can be reliably detected (Mitra, 2003). Consequently, concentrations less than the LOQ were removed from the total PAH computation as they were considered unreliable. A priori 5-level calibration was carried out with a calibration solution mix. The calibration solution mix was made up with EPA 525 PAH Mix-A standard solution (Sigma-Aldrich Co. Ltd., Dorset, UK), surrogate standard solution mix, and solution mix of deuterated PAHs as internal standards. Deuterated PAHs

used were Naphthalene-d_8, Anthracene-d_{10}, Chrysene-d_{12}, and Perylene-d_{12} while surrogates used were 2-Fluorobiphenyl and *o*-Terphenyl (Sigma-Aldrich Co. Ltd., Dorset, UK). Matrix spikes, duplicates, solvent and method blanks were also analysed as quality control samples.

5.3.4 Accuracy, precision, and experimental uncertainty of reference PAH analytical method

Accuracy of the reference SUSE–GC chemical method was evaluated from the percent recovery of the surrogate spiked into the sample prior to extraction according to equation 5-1.

$$\%\ \text{Recovery} = 100 \times \frac{\text{Measured concentration}}{\text{Theoretical concentration}} \qquad (5\text{-}1)$$

Where, theoretical concentration is equal to the concentration of the surrogate standard spiked into the sample. PAH final concentrations were bias-corrected since percent spike recoveries were between 0 and 100 as recommended in literature (Patnaik, 1997). Biases in measured concentrations of 2- and 3-ring PAHs were corrected with the percent recovery of 2-Fluorobiphenyl while the percent recovery of *o*-Terphenyl was used to correct for biases in measured concentrations of 4-, 5-, and 6-ring PAHs. Individual PAH concentrations were then summed up to get total PAH. Since several replicate analyses of samples were not possible in this study, reproducibility (precision) of the analytical method was determined by estimating the relative percent difference of duplicate analyses of one sample in each batch of samples as recommended in

literature (Patnaik, 1997). Relative percent difference was deduced by means of equation 5-2.

$$\text{Relative percent difference}(\%) = \frac{(a_1 - a_2)}{\frac{(a_1 + a_2)}{2}} \times 100 \quad (5\text{-}2)$$

Where a_1 and a_2 are the PAH concentrations in the first and second duplicate sample, respectively.

For each batch of samples, experimental uncertainty was estimated using confidence interval at 95% level of confidence for duplicate analyses of one randomly chosen sample. The confidence interval was deduced by the well-known expression in equation 5-3.

$$\text{Confidence interval} = \bar{x} \pm \frac{ts}{\sqrt{n}} \quad (5\text{-}3)$$

Where, $\bar{x}$ = the sample mean, t = Student's *t* for a desired level of confidence, s = the sample standard deviation, and n = the number of measurements.

Then, percent error (absolute) in the current measurement was checked against PAH results obtained from a commercial laboratory in Nigeria. To do this, we ran duplicate analyses of the same sample from Baraboo site that was previously analysed in the commercial laboratory. Percent error was then computed by means of equation 5-4.

$$\%\ \text{Error} = \left| \frac{T - E}{T} \right| \times 100 \quad (5\text{-}4)$$

Where, T = “known” PAH value from commercial laboratory, and E = measured mean PAH value from current study.

5.3.5 Optical measurement of soil samples

We took diffuse reflectance spectra from the soil samples with a mobile fibre-optic LabSpec2500® vis-NIR spectrophotometer (350–2500 nm) coupled to a high-intensity probe (Analytical Spectral Devices Inc., CO, USA). The spectrophotometer has one Si array (350–1000 nm) and two Peltier-cooled InGaAs detectors (1000–1800 nm and 1800–2500 nm). Spectral sampling interval of the instrument was 1 nm across the entire spectral range. However, the spectral resolution was 3 nm at 700 nm and 10 nm at 1400 and 2100 nm. The high-intensity probe has a built-in light source made of a quartz-halogen bulb of 2727^{o}C. The light source and detection fibres are assembled in the high-intensity probe enclosing a 35-degree angle. Before use, and after every 30 minutes, the instrument was optimised by white-referencing with a white Spectralon disc of almost 100% reflectance. Reflectance spectra were taken from each soil sample, tightly packed and levelled out in a cuvette, at three different positions, 120° apart. Each sample was scanned three times at each position and averaged before spectral pre-processing and multivariate analysis.

5.3.6 Spectral pre-processing

The pre-processing of soil spectra was carried out with the Unscrambler® X (CAMO Software AS, Oslo, Norway). Noisy portions at the extremes of the spectrum (i.e., from 350–449 nm and 2451–2500 nm) were removed due to low

instrument sensitivity at these wavelengths. Spectral truncation was followed by smoothing (by averaging successive 5-nm wavelengths), leaving a total of 401 wavelengths in the modelling range of 452–2450 nm. Then spectral transformation was carried out by successive combination of maximum normalization and Savitzky-Golay first derivative of polynomial order of two and two smoothing points. Normalization helps to bring all data to approximately the same scale or to get a more even distribution of the variances and the average values. First derivative removes additive baseline shifts in the data and smoothing reduces the impact of noise. These measures were aimed at reducing spurious peaks that do not hold physical or chemical information (Aske et al., 2001; Naes et al., 2002).

5.3.7 Partial least-squares (PLS) regression analysis

Before calibration, spectral reflectance (R) was transformed to the logarithm of the relative intensity (1/R), or absorption (Naes et al., 2002). The PLS regression analysis combines both the independent variables (reference values of PAH) and the dependent variables (wavelengths) using them as regression generators for the independent variables (Maleki et al., 2007). Detailed information about the PLS can be found in Martens and Naes (1989). We used PLS regression analysis with full cross-validation to relate the variation in a single-component variable (e.g., PAH) to the variation in a multi-component variable (e.g., wavelength) by means of Unscrambler® X. The optimal number of latent factors for future predictions was determined on the basis of the number of factors with the smallest total residual validation Y-variance or highest total

explained validation Y-variance (CAMO Software, 2012). Site-specific calibration models were developed with vis-NIR spectral data and reference SUSE–GC chemical data for each site. Also, a generalised model was developed for all three sites with the entire dataset. To develop the generalised model, 78% of the samples were used for cross-validation (calibration) while the remaining 22% were used for validation (prediction). The ratio of calibration/validation samples was chosen to ensure an equal representation of samples in the validation set by randomly choosing ten samples from each site.

5.3.8 Statistical evaluation of model performance

Criteria used to evaluate the quality of the models were based on the root-mean-square error (RMSE) of cross-validation and prediction (equation 3-1 and 3-2, respectively), ratio of prediction deviation (RPD) (equation 3-4), and corresponding coefficient of determination (r^2) (Williams and Sobering, 1986). Model prediction ability was categorised based on the following criteria: excellent if RPD > 2.0, almost good if 1.4 ≤ RPD < 2.0, and unreliable if RPD < 1.4 (Chang et al., 2001).

5.3.9 Outlier detection techniques

During model calibration, a priori outlier detection techniques were adopted to remove influential X- and Y-outliers. The influence of sample outliers was ascertained using the influence plot (Figure 5-3).

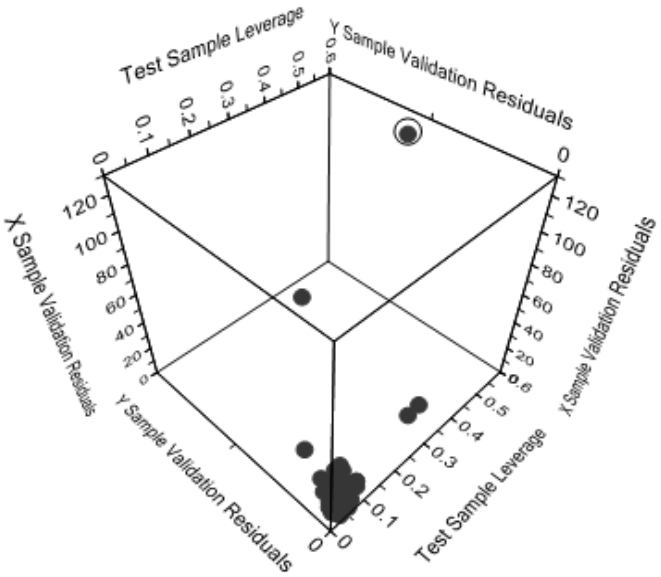


Figure 5-3 Detection of outliers after partial least-squares regression analysis. A potential sample outlier, marked in circle, detected among samples from Baraboo site in Ogoniland, Niger Delta province of Nigeria is shown as an example.

This was done by plotting the residual X- and Y-variances against leverages. Samples with a high leverage and high residual X- or Y-variance were considered as potential outliers (CAMO Software, 2012). Then the projected influence plots were used to confirm those samples with high residuals. Nonetheless, before such samples were treated as outliers, they were studied in more details by marking them one-by-one and plotting the X-Y relation outliers for several model factors to observe their influence on the shape of the X-Y relationship, as recommended in literature (CAMO Software, 2012).

5.4 Results and discussion

5.4.1 Accuracy, precision and level of uncertainty in reference SUSE–GC analysis of PAH

Table 5-1 shows the statistical results of the chemical SUSE–GC analysis showing the sum of individual PAHs quantified for each site. The high spatial variability of the compounds within and among sites is clearly demonstrated in the table. It is important that the sum of a PAH fraction in the entire sample set (Table 5-1) is differentiated from the sum of the entire PAH fraction in a sample (Appendix E to Appendix G). The latter is the total PAH in a sample, which is used for model development. In this study, total PAH concentration is the sum of the 13 individual PAH concentrations in a sample. The 13 PAHs in Table 5-1 were the relevant compounds contained in the PAH standard solution used for prior instrument calibration. The concentrations of the lower boiling 2- to 4-ring PAHs (i.e., acenaphthylene to chrysene) were higher than the higher boiling 5- to 6-ring PAHs. The difficulty in quantitation increased with boiling point. The lower boiling PAHs eluted the column before the higher boing PAHs because of their shorter retention times. Overall, Bomu 2 appeared to be the most contaminated site with the highest mean total PAH (5.39 $mgkg^{-1}$) compared to either Baraboo site (3.38 mg kg^{-1}) or Bomu 1 (4.47 mg kg^{-1}) as shown in Appendix E to Appendix G.

Table 5-1 Statistics of the chemical analysis result showing the sum of individual polycyclic aromatic hydrocarbons (PAHs) quantified for each site by reference sequential ultrasonic solvent extraction–gas chromatography (SUSE–GC).

PAH	LOQ (mg kg^{-1})	Baraboo					Bomu 1					Bomu 2				
		N	Min. (mg kg^{-1})	Max. (mg kg^{-1})	Mean (mg kg^{-1})	Range (mg kg^{-1})	*N*	Min. (mg kg^{-1})	Max. (mg kg^{-1})	Mean (mg kg^{-1})	Range (mg kg^{-1})	*N*	Min. (mg kg^{-1})	Max. (mg kg^{-1})	Mean (mg kg^{-1})	Range (mg kg^{-1})
Acenaphthylene	0.02	43	0.07	3.06	0.83	2.98	58	0.02	2.40	0.46	2.38	36	<0.02	8.53	1.41	8.53
Fluorene	0.02	43	<0.02	4.80	0.73	4.80	58	<0.02	2.70	0.45	2.70	36	<0.02	6.25	0.88	6.25
Phenanthrene	0.02	43	<0.02	3.64	0.47	3.64	58	0.02	3.15	0.81	3.13	36	<0.02	6.86	0.76	6.86
Anthracene	0.02	43	<0.02	4.22	0.41	4.22	58	<0.02	2.57	0.36	2.57	36	<0.02	6.86	0.75	6.86
Pyrene	0.02	43	<0.02	1.03	0.30	1.03	58	<0.02	2.83	0.56	2.83	36	<0.02	1.73	0.52	1.73
Benz[a]anthracene	0.02	43	<0.02	0.88	0.18	0.88	58	<0.02	2.61	0.39	2.61	36	<0.02	1.01	0.15	1.01
Chrysene	0.02	43	<0.02	1.17	0.18	1.17	58	<0.02	1.07	0.22	1.07	36	<0.02	0.55	0.09	0.55
Benz[b]fluoranthene	0.02	43	<0.02	0.15	0.03	0.15	58	<0.02	1.33	0.23	1.33	36	<0.02	0.18	0.04	0.18
Benzo[k]fluoranthene	0.02	43	<0.02	0.15	0.03	0.15	58	<0.02	1.23	0.19	1.23	36	<0.02	0.18	0.04	0.18
Benzo[a]pyrene	0.02	43	<0.02	0.73	0.17	0.73	58	<0.02	1.80	0.14	1.80	36	<0.02	9.36	0.68	9.36
Indeno[1,2,3-cd]pyrene	0.02	43	<0.02	0.22	0.02	0.22	58	<0.02	1.48	0.23	1.48	36	<0.02	0.82	0.06	0.82
Dibenzo[a,h]anthracene	0.02	43	<0.02	0.15	0.03	0.15	58	<0.02	1.82	0.22	1.82	36	<0.02	<0.02	-	-
Benzo[g,h,i]perylene	0.02	43	<0.02	0.22	0.02	0.22	58	<0.02	1.48	0.22	1.48	36	<0.02	0.33	0.02	0.33

LOQ, Limit of quantitation. It is the lowest concentration at which an analyte can be reliably detected (Mitra, 2003).

N, Number of samples

PAH, Polycyclic aromatic hydrocarbons

Percent recoveries of spiked surrogates for the three sites, as shown in Table 5-2, are within the acceptable range of 40–120% (EPA, 1999a), which is typical of a reasonably accurate PAH extraction procedure.

Table 5-2 Accuracy of the reference sequential ultrasonic solvent extraction–gas chromatography (SUSE–GC) method used in the chemical analysis of polycyclic aromatic hydrocarbons (PAH) in topsoils from petroleum-contaminated sites in Ogoniland, Niger Delta province of Nigeria.

Sampling site	No. of samples	Spiked ($mgkg^{-1}$)	Measured ($mgkg^{-1}$) [a]		% Recovery of spiked surrogates
			2-Fluorobiphenyl	o-Terphenyl	
Baraboo	43	2.50	1.36	1.33	53–54
Bomu 1	58	2.50	1.94	1.63	65–78
Bomu 2	36	2.50	1.19	1.40	48–56

[a] Mean concentrations of the spiked surrogates.

Table 5-3 shows the reproducibility and experimental uncertainty for 95% confidence interval (n = 2) of the reference PAH measurement procedure. The relative percent difference for duplicate samples for all sites is less than 20%, suggesting that the reference SUSE–GC method was within precision standards (Mayer, A. S. [2008] – Lecture Note, GE3850: Geohydrology, Department of Civil and Environmental Engineering, Michigan Technology University, MI. – Unpublished result). The rather large confidence intervals in Table 5-3 are attributed to the fewer number of measurements (n = 2). It is known that the confidence interval reduces with number of measurements as the sample mean approaches the true population mean (Harris, 2010). Nevertheless, the estimated confidence interval suggests with 95% confidence that the true mean PAH value from the duplicate measurements would lie within the estimated range.

Table 5-3 Experimental uncertainty and precision of reference sequential ultrasonic solvent extraction–gas chromatography (SUSE–GC) method used for analysis of polycyclic aromatic hydrocarbon (PAH) in topsoils from petroleum-contaminated sites in Ogoniland, Niger Delta of Nigeria. Test sample was randomly chosen for each site.

Sampling site	First duplicate (mg kg^{-1})	Second duplicate (mg kg^{-1})	Relative percent difference (%)	Confidence interval
				Mean ± uncertainty [a] (mg kg^{-1})
Baraboo	3.67	3.22	13	3.45 ± 2.86
Bomu 1	4.49	3.88	14	4.18 ± 3.84
Bomu 2	2.89	2.66	8	2.78 ± 1.42

[a] Experimental uncertainty for 95 % confidence interval for duplicate measurements.

Inter-laboratory differences in reported PAH values are shown in Table 5-4. For the test soil sample, Table 5-4 shows that mean PAH value obtained in this study is comparable to PAH result from the commercial laboratory in Nigeria. This implies that the current PAH measurement by Cranfield University's operating procedures described above is in agreement with those of the commercial laboratory. The margin of error may be attributed to differences in operating procedures used in both laboratories (Table 5-4). Moreover, inter-laboratory differences in reported values of an analyte in chemical analysis are a common occurrence (Risdon et al., 2008).

Table 5-4 Inter-laboratory differences in reported PAH values and selected rubrics in operating procedures for test soil sample from petroleum-contaminated site in Baraboo in Ogoniland, Niger Delta of Nigeria.

Analyte and selected rubrics	This study	Commercial Laboratory
PAH ($mgkg^{-1}$)	3.45 ± 2.86	2.73 [a]
Percent error (%) [b]	26	-
Quantitation method	Internal standard	External standard
Extraction method	Sequential ultrasonic solvent extraction	Sonication water bath (5-hour sonication)
Extracting solvent(s)	Dichloromethane and hexane (1:1)	*n*-Pentane
Surrogate standard(s)	2-Fluorobiphenyl and o-Terphenyl	1-Chlorooctadecane

[a] Uncertainty was not reported.

[b] Commercial laboratory PAH result was taken as the "known" value.

5.4.2 The PAH partial least-squares regression models

Table 5-5 summarises the statistical results of the PLS models. The cross-validation models for pre-processed spectra used smaller number of latent variables than raw spectra. Similarly, cross-validation models developed by pre-processed spectra had better quality than raw reflectance models. Overall, the quality of the cross-validation models is classified as ranging from very good to excellent on the basis of RPD values. Figure 5-4 shows the linear relationship between reference SUSE–GC-measured and vis-NIR-predicted PAH data for pre-processed spectra of entire dataset. Only three sample outliers were removed from the cross-validation datasets while none was removed from the independent validation set. On the basis of RPD values, the model quality is classified as excellent prediction. The quality of the model demonstrates the possibility of using vis-NIR spectroscopy for quantitative

determinations (Chang et al., 2001; Viscarra Rossel et al., 2006a). Moreover, as shown in Table 5-5 the similarity between cross-validation and validation standard deviations indicates that the prediction models are not skewed. These results are comparable with RPD values of 1.7 and 2.5 reported by previous researchers for TPH prediction in field-moist soils by PLS regression analysis (Chakraborty et al., 2010, 2012).

Table 5-5 Statistical results of partial least-squares (PLS) site-specific calibration and general prediction models for polycyclic aromatic hydrocarbons (PAH) in petroleum-contaminated tropical rainforest soils from Ogoniland in the Niger Delta province of Nigeria developed by visible and near-infrared (vis-NIR) spectroscopy.

Sampling site	No. of samples	Reflectance spectra					Combined pre-processing spectra [a]				
		r^2	RMSE ($mgkg^{-1}$)	SD	RPD	LV	r^2	RMSE ($mgkg^{-1}$)	SD	RPD	LV
Cross-validation set											
Baraboo	43	0.78	0.82	1.81	2.20	8	0.84	0.64	1.81	2.81	6
Bomu 1	58	0.76	1.41	2.95	2.09	11	0.83	1.22	2.95	2.41	3
Bomu 2	36	0.77	1.98	4.08	2.07	5	0.83	1.42	4.08	2.87	7
General	107	0.69	1.61	3.05	1.81	11	0.82	1.30	3.05	2.34	8
Validation set											
General	30	0.77	1.95	3.63	1.86	11	0.89	1.16	3.63	3.12	8

[a] Combination of maximum normalisation, and first derivative and smoothing by Savitzky-Golay. LV, Latent variable; RMSE, Root-mean-square error; RPD, Ratio of prediction deviation; SD, Standard deviation

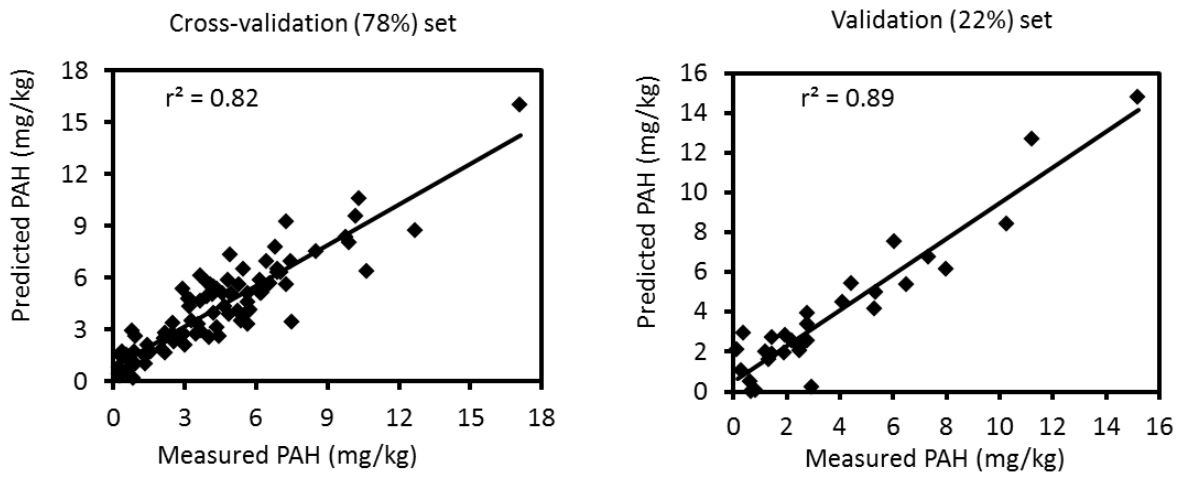


Figure 5-4 Scatter plots of chemically measured vs. predicted values of polycyclic aromatic hydrocarbons (PAH).

These were developed by partial least-squares (PLS) regression analysis with pre-processed spectra of 137 soil samples from petroleum-contaminated tropical rainforest soils collected from Ogoniland, Niger Delta, Nigeria.

5.4.3 Spectral reflectance of petroleum-contaminated tropical rainforest soils

Figure 5-5 shows mean vis-NIR spectral reflectance curves of field-moist tropical rainforest soils consisting of sandy clay from three oil spill sites in Ogoniland, Niger Delta, Nigeria. Generally, soil spectral reflectance decreased with increasing mean PAH concentration particularly in the NIR region (700–2500 nm). There was however, a slight shift in this trend for Baraboo and Bomu 1 reflectance curves in the visible range (452–700 nm). This is attributed to soil colour-associated changes in the visible range brought about when water and/or oil (e.g., diesel) is added to soil (Mouazen et al., 2005; Okparanma and Mouazen, 2013b). In the reflectance curves for the contaminated sites, spectral absorption minima of hydrocarbon-based oil were observed around 1647, 1712, and 1752 nm in the first overtone region of the NIR band (Figure 5-5).

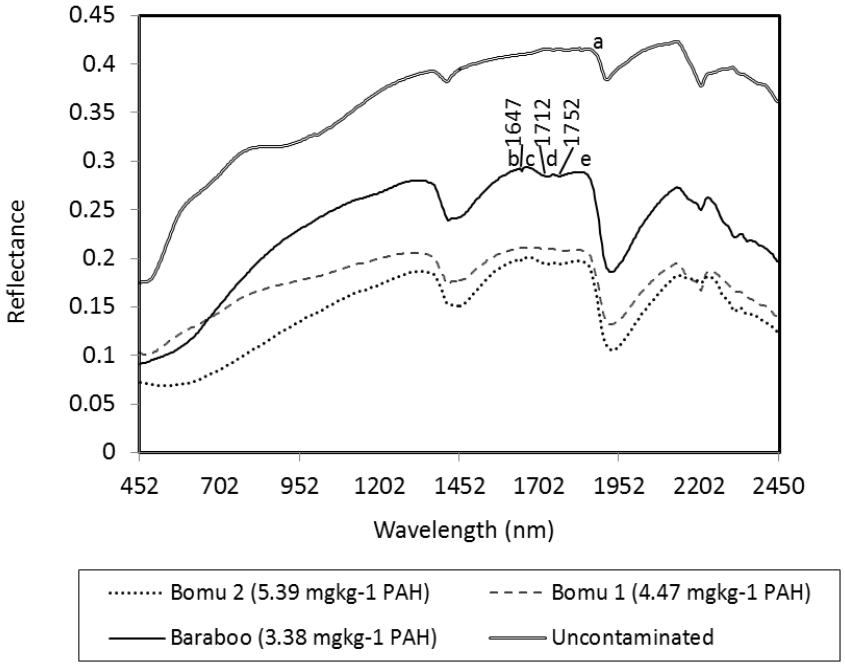


Figure 5-5 Mean vis-NIR spectral reflectance curves of petroleum-contaminated tropical rainforest soils from three oil spill sites in Ogoniland in the Niger Delta province of Nigeria. Values in the legend are average PAH concentrations.
Point a, b, c, d, and e are reflectance shoulders.

The absorption around 1647 nm is attributed to C–H stretching modes of ArCH linked to PAH. Absorptions around 1712 and 1752 nm are attributed to C–H stretching modes of terminal CH_3 and saturated CH_2 groups linked to TPH, both present in the contaminating hydrocarbon-based oil (Forrester et al., 2010; Osborne et al. 1993; Workman and Weyer, 2008). These features are practically absent in the uncontaminated reflectance curve as shown by the number of reflectance shoulders that appeared in the reflectance curves between 1452 and 1952 nm (Figure 5-5).

One reflectance shoulder appeared around 1830 nm (point a) in the uncontaminated reflectance curve while four appeared around 1630, 1675, 1737, and 1830 nm (point b, c, d, and e, respectively) in the contaminated reflectance curves (Figure 5-5). A reflectance shoulder appears between two absorption wavelengths of colour, water

and/or hydrocarbon (Mouazen et al., 2007b; Okparanma and Mouazen, 2013b). The reflectance shoulder at 1830 nm is attributed to water absorption in the first overtone band around 1950 nm (Stenberg, 2010) and hydrocarbon absorption in the first overtone band around 1712 nm (Osborne et al. 1993; Workman and Weyer, 2008). The shoulder at 1630nm is attributed to water absorption in the second overtone region around 1450 nm (Mouazen et al., 2007b) and hydrocarbon absorption in the first overtone band around 1647 nm (Workman and Weyer, 2008). The shoulder at 1675 nm is attributed to hydrocarbon absorptions in the first overtone band around 1647 and 1712 nm (Workman and Weyer, 2008). The shoulder at 1737 nm is attributed to hydrocarbon absorptions in the first overtone band around 1712 and 1752 nm (Workman and Weyer, 2008). Therefore, the hydrocarbon absorption bands around 1647 and 1712 nm differentiate the uncontaminated from contaminated reflectance curves (Figure 5-5).

5.4.4 Regression coefficients

Regression coefficients in PLS regression analysis provide a summary of all predictors and a given response. The plot of regression coefficients is used to identify important wavelengths for the prediction of relevant soil properties. Figure 5-6 shows bar plots of regression coefficients versus wavelength derived after PLS regression analysis using raw vis-NIR spectral data and reference SUSE–GC chemical data for each site. In the bar plots, the absolute value of the regression coefficients indicate the relative importance of the wavelength on the basis of explained X-variance in the model.

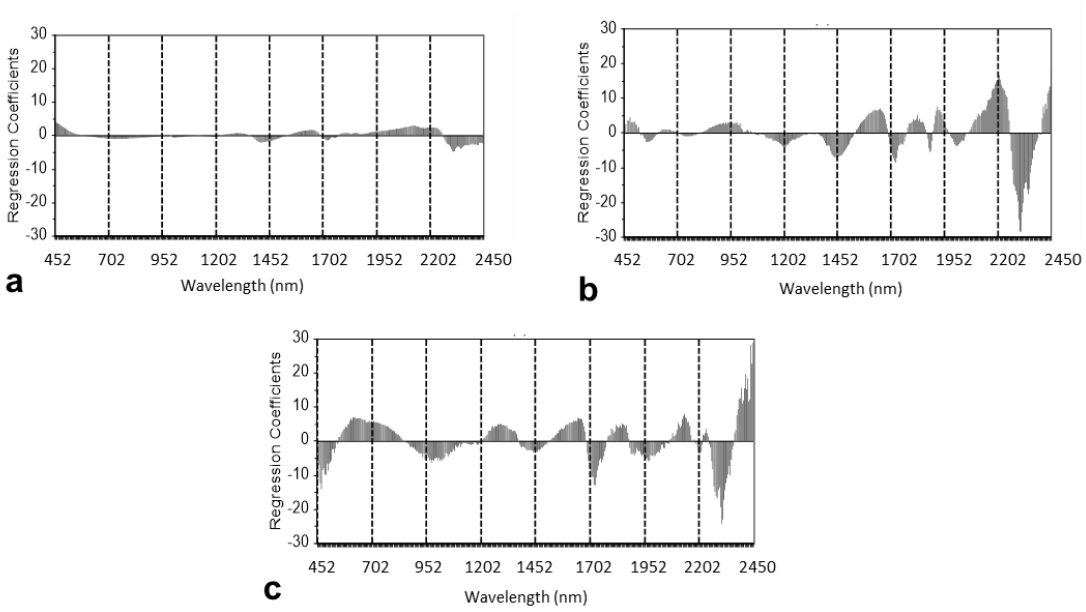


Figure 5-6 Plots of regression coefficients vs. wavelength after partial least squares (PLS) regression analysis by visible and near infrared (vis-NIR) diffuse reflectance spectroscopy for petroleum-contaminated tropical rainforest soils:
collected from (a) Baraboo, (b) Bomu 1, and (c) Bomu 2 in Ogoniland, Niger Delta, Nigeria.

In Figure 5-6, negative coefficients around 1712 and 1752 nm show a positive link to absorptions due to vibrational C–H stretching modes of terminal CH_3 and saturated CH_2 functional chemical groups linked to TPH. But positive coefficients around 1647 nm are consistent with absorptions due to vibrational C–H stretching modes of ArCH functional groups linked to PAH. These results agree with absorption bands observed in the reflectance curves (see Figure 5-5). Over the modelling wavelength range of 452–2450 nm, the bar plots show that intensities of regression coefficients vary in magnitude for each site. Nonetheless, the positions of important wavelengths are largely similar (Figure 5-6) even though the samples are from different sites. This shows a similarity in soil property among the three sites. We observed larger regression coefficients in Figure 5-6b and 5c than in Figure 5-6a. The intensity of

regression coefficients for Baraboo site was the least (Fig. 5a), and samples from Baraboo site showed the least level of contamination among the three sites as well (see Figure 5-5). This may be explained by the fact that the mechanism of PAH prediction is attributed to co-variation of PAH with other soil properties that have direct spectral responses in the NIR spectral range, particularly water, clay minerals and organic carbon (Okparanma and Mouazen, 2013c). This observation was also typified in Figure 5-6 by the sizes of coefficients around 950, 1450, 1950, 2200, 2212, and 2300 nm. Absorptions around 950, 1450 and 1950 nm in the NIR band are due to O–H stretching modes of water in the O–H second and first overtones, and combinations band respectively (Whalley and Stafford, 1992). Absorption features linked to metal-OH bend plus O–H stretch combinations around 2200 and 2300 nm are characteristic of clay minerals (Clark et al., 1990; Viscarra Rossel et al., 2006b). Absorption features attributed to long-chain C–H+C–H and C–H+C–C stretch combinations around 2150 nm and 2212 nm are unique to soil organic matter (Forrester et al., 2010).

5.5 Conclusions

In this study, results show that soil diffuse reflectance decreased with increasing PAH concentration. Positive regression coefficients around 1647 nm show a link to PAH. Additionally, the quality of generalized PLS models for PAH predictions ranged from good to excellent (validation r^2 = 0.77–0.89, RMSEP = 1.16–1.95 $mgkg^{-1}$, and RPD = 1.86–3.12). This demonstrates the possibility of using vis-NIR spectroscopy and PLS regression analysis for rapid quantitative determination of PAH in petroleum-contaminated tropical rainforest soils in the Niger Delta province of

Nigeria. However, it should be pointed out that the extrapolation of the model is limited to the three sites investigated in this study namely Baraboo, Bomu 1, and Bomu 2 in Ogoniland, Niger Delta region of Nigeria. To set up a model based on vis-NIR spectroscopy for general application in the Niger Delta province of Nigeria, we recommend the use of larger dataset covering both the concentration range and all the other sources of variability in oil spill sites in the Niger Delta region.

Chapter 6: Mapping PAH and total toxicity equivalent soil concentrations by vis-NIR spectroscopy for hazard assessment of petroleum release sites

6.1 Chapter Summary

In Chapter 5:, we have shown that spectroscopic models based on vis-NIR spectroscopy (350–2500 nm) predict PAH in tropical rainforest soils with reasonable accuracy. In this chapter, we used these models to develop soil maps of PAHs and total toxicity equivalent soil concentrations (TTEC) of the PAH mixture. We used the soil maps of TTEC for hazard assessment of the three petroleum release sites in Ogoniland in the Niger Delta province of Nigeria from where the soils were collected. Inverse distance weighting soil maps show a near-perfect match of high and low PAH and TTEC zones between measured and predicted maps. There were non-significant (p<0.05) differences between measured and predicted soil maps as the ANOVA revealed. There was a fair to good agreement between soil maps of PAH developed by reference GC-MS data and vis-NIR data (Kappa coefficient = 0.19–0.56). Hazard assessment based on Generic Assessment Criteria (GAC) established for the studied sites in this study show that the degree of action for site-specific risk assessment and/or remediation differs among the different zones. These results are indications that vis-NIR spectroscopy has the potential for mapping PAH and TTEC of the PAH mixture for hazard assessment of petroleum release sites in the Niger Delta province of Nigeria.

6.2 Introduction

Polycyclic aromatic hydrocarbons (PAHs) are widely distributed organic pollutants. The Water Framework Directive (WFD, 2000/60/EC) identified them as priority hazardous substances (PHS), and listed five key indicator compounds including benzo[a]pyrene, benzo[b]fluoranthene, benzo[k]fluoranthene, benzo[g,h,i]perylene, and indeno[1,2,3-cd]pyrene. Anthracene and naphthalene have been recently added to the list of PHS. In the EU, the cost of PAH decontamination is estimated to be approximately €17.3 billion per year (CEC, 2006).

Risk assessment is an established requirement for effective management of contaminated land, and now a widely-used support tool for environmental management decisions. It is employed as a means of assessing and managing potential impacts to human- and ecosystem health (Vegter, 2002). Assessing the risk of PAH at contaminated sites is complicated because of the profiles of compounds present. Specific indicator compounds (genotoxic carcinogens and non-carcinogens) should be assessed because these are often the key risk drivers at petroleum contaminated sites (Brassington et al., 2007). Genotoxic carcinogens are assumed not to have a threshold concentration as even very small concentrations (or doses) are assumed to pose some (albeit small) risk of cancer. There are cases in which carcinogenicity can be assumed to occur only after some dose or threshold concentration is reached, depending on the mode of action by which the contaminant is thought to cause cancer. To carry out such assessment, the United States Environmental Protection Agency (US EPA) developed the toxicity equivalency factor (TEF) methodology for a mixture of structurally related chemical compounds

with a common mechanism of action such as PAHs (WSDE, 2007). A TEF is an estimate of the relative toxicity of an individual PAH to benzo[a]pyrene, which is usually chosen as the reference chemical compound since its toxicity is well characterized (WSDE, 2007). The TEF approach aims to give a single concentration number (e.g., total toxicity equivalent concentration – TTEC), to environmental matrixes like soil with highly complex nature. To determine compliance for a particular soil sample, the deduced TTEC for the PAH mixture is compared with the applicable target value for benzo[a]pyrene (WSDE, 2007).

In petroleum release sites, mapping the spatial and temporal variation of PAH and the toxicity levels are required for a broad range of environmental applications. Mapping the toxicity of PAHs, which equates with hazard assessment or tier 1 Risk-Based Corrective Action (RBCA) for petroleum release sites (ASTM, 1995), helps in the delineation of potential management zones within the site. This also provides vital information for quantitative risk assessment and/or remediation (if action is required). Traditional methods of hazard assessment in petroleum release sites involve prior soil sampling, extraction of PAH compounds from the soil sample using various extraction solvents, and analysis of the liquid extract by gas chromatography–mass spectrometry (GC–MS) (EPA, 1999a). Soil sampling and solvent extraction of PAHs are tedious, labour-intensive, time consuming and require expert operators. Therefore, PAH analysis by GC–MS is hazardous and uneconomical; particularly when large-scale contamination is involved and dense sampling is required for high-resolution soil contaminant mapping (Peterson et al., 2002; Okparanma and Mouazen, 2013a). This has prompted increasing demand for

rapid and cost-effective methods capable of high-sampling resolution essential for spatial mapping to complement the conventional methods.

In response to this demand, scientists have recognised the potential of visible and near-infrared (vis-NIR) diffuse reflectance spectroscopy as a cost-effective tool for rapid identification of hydrocarbon contamination in soils (e.g., Malley et al., 1999). Over the years, several attempts have been made to use reflectance spectroscopy as an approach for mapping hydrocarbons in soils. Two studies employed airborne remote sensing with Landsat Thematic Mapper (TM) and Daedalus scanner to map hydrocarbons in soil (Bannert et al., 1994; Kuhn and Horig, 1995). Both studies yielded limited results because of the overly poor spectral resolution of the multispectral Landsat TM and Daedalus scanner (Horig et al., 2001). A later study employed airborne hyperspectral remote sensing using HyMap scanner in conjunction with a field GER Mark V IRIS® infrared spectroradiometer (385–2548 nm) to identify hydrocarbon-contaminated soils (Horig et al., 2001). However, this approach was amenable only to high hydrocarbon concentrations (2.5% w/w), and typical hydrocarbon absorption features in the pixel spectra were not very well pronounced (Horig et al., 2001)**.** In a similar study, Kuhn et al. (2004**)** developed the Hydrocarbon Index approach for mapping hydrocarbon bearing materials with the HyMap scanner. However, the approach was limited to sensors with very high signal-to-noise ratio like the HyMap scanner. Moreover, airborne hyperspectral remote sensing relies on bare earth imagery, which is affected by land cover and vegetation (Schwartz et al., 2011). Even then, the application of the methodology for soil investigation is still at the rudimentary stages. As a result, studies adopting airborne hyperspectral remote sensing are rarely found in literature since only few of

the sensors are available worldwide (Schwartz et al., 2011). Consequently, investigation into the possibility of using point reflectance spectroscopy for mapping hydrocarbons and the level of toxicity in petroleum release sites is needed. To the best of our knowledge, none of the studies reported in the literature on mapping hydrocarbon contamination in soils adopted vis-NIR spectroscopy – although the methodology has been proven to map soil properties for precision agriculture applications (e.g., Mouazen et al., 2007; Kodaira and Shibusawa, 2013; González et al., 2013 ; Quraishi and Mouazen, 2013).

Recently, the application of vis-NIR spectroscopy for the prediction of total PAH in petroleum contaminated tropical rainforest Oxisols of the Niger Delta province in Nigeria was reported by Okparanma et al. (2014a). They used a total of 137 field-moist intact soil samples for their study, and reported extraction efficiency ranging from 48 to 78 % for the chemical PAH extraction protocol used in the study, which were said to be within recommended standards (Okparanma et al., 2014a). They also reported that a total of 13 United States Environmental Protection Agency (US EPA) priority PAHs were identified and quantified in the three studied sites as shown in Table 5-1 (Okparanma et al., 2014a). Using partial least squares (PLS) regression analysis with full cross-validation, they developed vis-NIR-based models to predict total PAH in the soil samples. The site-specific calibration and generalised prediction models developed demonstrated reasonable accuracy with coefficient of determination (r^2) ranging from 0.69 to 0.89, root mean square error (RMSE) ranging from 1.16 to 1.98 mg kg^{-1}, and ratio of prediction deviation (RPD) ranging from 1.81 to 3.12 (Okparanma et al., 2014a). This result and the successful use of vis-NIR spectroscopy to map soil properties as reported by previous researchers prompted

further studies leading to the development of soil maps of PAH and TTEC of the PAH mixture for hazard assessment of the three petroleum release sites in Nigeria. Mapping PAH and TTEC of the PAH mixture using vis-NIR method should reduce the amount of time, energy and money usually expended in hazard assessment of petroleum release sites by the conventional methods.

The objectives of the current study were: (1) to develop maps of PAH and TTEC of the PAH mixture using data from vis-NIR-based models developed in Chapter 5: (Okparanma et al., 2014a), and (2) to use the soil maps of TTEC of the PAH mixture for hazard assessment of the studied sites based on Generic Assessment Criteria (GAC) established in this study for the three petroleum release sites in Ogoniland in the Niger Delta province of Nigeria.

6.3 Materials and Methods

6.3.1 The study area

The Niger Delta province (5.317°N, 6.467°E) covers a total land area of 70000 km^2, and is home to Nigeria's oil and gas industries with most of the country's oil and gas fields located in the area (Niger Delta Environmental Survey, 1995). Ogoniland, covering some 1000 km^2, is located within the Niger Delta province in southern Nigeria. Topsoils of Ogoniland are mainly sandy clay (UNEP, 2011), which is consistent with Oxisols that characterise tropical rainforest soils of southern Nigeria (USDA, 2005). In Ogoniland, soil total organic carbon range from 3.63 to 4.11 % (Tanee and Albert, 2011), and typical soil pH at all soil depths range from 5.2 to 6.4, according to studies reported for similar ecosystems in the Niger Delta (SPDC, 2006). The increased oil and gas activities in recent years make the Niger Delta

province most vulnerable to environmental pollution. For example, in Ogoniland alone, there are over 69 contaminated land sites (UNEP, 2011). In this study, we selected three oil spill sites in Gokana Local Government Authority (LGA) in Ogoniland – site A is located at Baraboo (4.652°N, 7.249°E); site B at Bomu 1 (4.662°N, 7.277°E); and site C is at Bomu 2 (4.662°N, 7.249°E) (Figure 6-1).

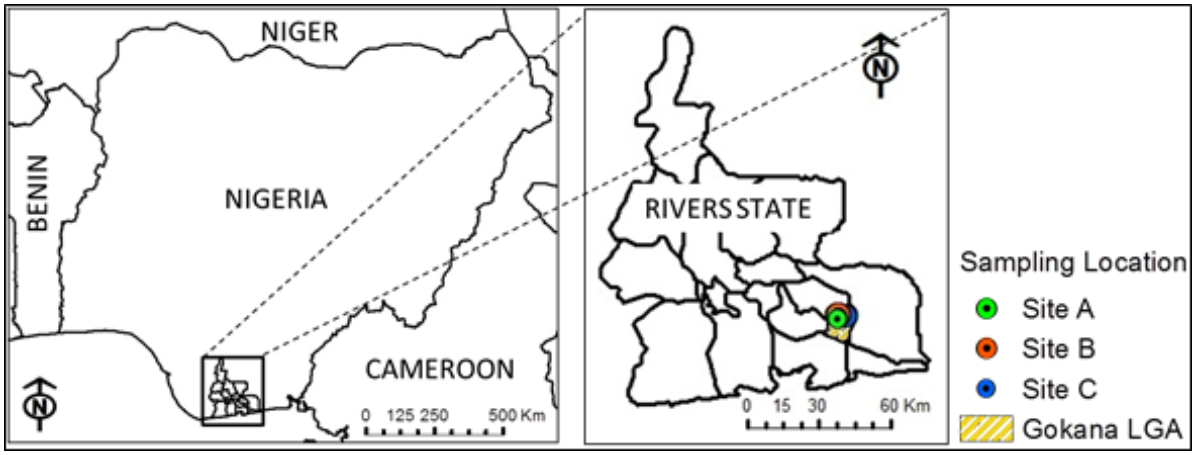


Figure 6-1 Sampling locations in Gokana Local Government Authority (LGA) in Ogoniland, Rivers State in the Niger Delta province of Nigeria.
(Projection: Geographic Coordinate System – World Geographic System 1984 in ArcGIS® 10.1; ESRI™, CA, USA)

6.3.2 Soil sampling

From three oil spill sites in Ogoniland, we collected a total of 137 soil samples from the top 15-cm depth in plastic containers and preserved them in a cooler containing ice blocks until shipment to the UK. The number of samples collected from site A, B, and C were 43, 58, and 36 respectively. We adopted targeted sampling method to cover as much of the potentially contaminated spots as possible. Sample management was strictly in line with the standards of the Nigerian Government's Department of Petroleum Resources (2002).

6.3.3 Soil chemical PAH analysis

Soil chemical analysis for PAHs was carried out by sequential ultrasonic solvent extraction–gas chromatography. Sequential ultrasonic solvent extraction was used to extract the PAH compounds from the soil samples as described by Risdon et al. (2008) (see subsection 5.3.3). Identification and quantification of the PAH compounds were carried by gas chromatography–mass spectrometry (GC–MS) as described by Okparanma et al. (2014a) (see subsection 5.3.3).

6.3.4 Soil optical measurement and development of PLS calibration models

Optical scanning of soil samples in diffuse reflectance mode was carried out with a mobile fibre-optic LabSpec2500® vis-NIR spectrophotometer (350–2500 nm) (Analytical Spectral Devices Inc., CO, USA). Soil scanning protocols, pre-processing of soil spectra, and establishment of PLS calibration models were carried as described by Okparanma et al. (2014a) (see subsection 5.3.5). Spectral pre-processing and PLS regression analysis were carried out with the Unscrambler® X version 10.2 (CAMO Software AS, Oslo, Norway).

6.3.5 Establishment of Generic Assessment Criteria (GAC) and delineation of potential management zones

The decision on the likely need for site-specific assessment and/or site-specific remediation of a petroleum release site following hazard assessment is done using GAC (Petts et al., 1997). This has the advantage of saving time and money that would otherwise be expended in developing site-specific assessment criteria (Ferguson et al., 1998; Cheng and Nathanail, 2009).

In this study, the threshold–action concept, which is the underlying concept for establishing GAC (Petts et al., 1997), was used to delineate potential management zones within the studied sites. To do this, the first step was to transform individual PAH concentrations to toxicity equivalent concentrations (TEC in mg kg^{-1}) using Equation 6-1, and compute the TTEC (mg kg^{-1}) for the PAH mixture using Equation 6-2 as follows (WSDE, 2007):

$$TEC = C \times TEF \tag{6-1}$$

$$TTEC = \sum C_n \times TEF_n \tag{6-2}$$

Where C = PAH concentration (in mg kg^{-1}), n = the individual PAH in the mixture, and TEF = toxicity equivalency factor (dimensionless). TEF values for PAHs are freely available on the internet. Table 6-1 shows TEF values of US EPA priority PAHs relevant to this study.

Table 6-1 Toxicity Equivalency Factor for selected United States Environmental Protection Agency (US EPA) priority polycyclic aromatic hydrocarbons (PAHs)

PAHs	Toxicity Equivalency Factor [a]
Acenaphthylene	0.01
Fluorene	0.001
Phenanthrene	0.001
Anthracene	0.01
Pyrene	0.001
Benzo[a]anthracene	0.1
Chrysene	0.01
Benzo[b]fluoranthene	0.1
Benzo[k]fluoranthene	0.1
Benzo[a]pyrene	1
Indeno[1,2,3-cd]pyrene	0.1
Dibenzo[a,h]anthracene	0.1
Benzo[g,h,i]perylene	0.01

[a] Data from Nisbet and LaGoy (1992), De Meulenaer (2006), WSDE (2007).

The second step was to compare the TTEC for the PAH mixture with applicable target value (screening level) for benzo[a]pyrene. In Nigeria, the screening level for TPH is 50 mg kg^{-1} while for monocyclic aromatic hydrocarbons including benzene, toluene, ethylbenzene, and xylene (BTEX) the screening level is 0.05 mg kg^{-1} (DPR, 2002). But, there are currently no screening levels for PAHs in Nigeria. As a result, in this study, we considered the screening levels for benzo[a]pyrene used among 9 European countries (Table 6-2) and used the average value as a guide.

Table 6-2 Soil screening levels for benzo[a]pyrene for selected European countries

European country	Screening value for agricultural/special/ natural/sensitive land use (mg kg^{-1}) [a]
Belgium	0.5
Czech Republic	0.1
Denmark	0.1
Finland	0.2
Germany	1
Lithuania	0.1
Poland	0.03
Slovakia	0.1
Spain	0.02
Average	*0.24*

[a] Data from Carlon (2007).

Before adopting the European screening level for benzo[a]pyrene, we ensured as much as we possibly could that the end-use of the sites and exposure scenarios in Nigeria were comparable to those considered when developing the applicable screening level for benzo[a]pyrene in the selected European countries. From our investigation, the studied sites in Nigeria are largely used for agricultural purposes for which several soil outdoor exposure pathways can be considered. Of these exposure pathways, majority of the European countries within the assessment considered oral ingestion of soil (Carlon, 2007). Since this exposure scenario is

applicable to the system of agriculture in Nigeria, we assumed that soil outdoor exposure in the studied sites was through oral ingestion by an adult human. By this assumption, we were able to apply the TEF approach directly since the assumption is the same as the one used in developing the TEF method by the US EPA (WSDE, 2007).

Therefore, the average screening level for benzo[a]pyrene of 0.24 mg kg^{-1} (Table 6-2) was adopted as the screening threshold of benzo[a]pyrene for the three sites in Nigeria. In Deutschland, when screening data against guidelines or standards, the acceptable risk level for excess lifetime cancer risk (ECLR) is 1 x 10^{-4} for which a site-specific risk assessment is mandatory (Petts et al., 1997). But, it has been recently proposed that the acceptable risk and unacceptable risk levels for ECLR are in the ratio of 1:9 (Kibblewhite, M. [2013], National Soil Resources Institute, Cranfield University, England – personal communication). Accordingly, the unacceptable risk level for ECLR at a contaminated site should be 9 x 10^{-4} for which site-specific remediation would be required. Therefore, we adopted an action value (i.e., threshold of acceptability) of 2.16 mg kg^{-1} for benzo[a]pyrene in the three studied sites in Nigeria. Using the threshold–action concept (Petts et al., 1997), three management zones were delineated as follows. The actual number of management zones created for each site was based on the range of their TTEC.

1) Zone 1 (TTEC < 0.24 mg kg^{-1}). In this zone, no action is required because there is no unacceptable risk to specified target;
2) Zone 2 (0.24 ≤ TTEC ≤ 2.16 mg kg^{-1}). In this zone, site-specific assessment may be required because there may be additional risk to some targets; and

3) Zone 3 (TTEC > 2.16 mg kg^{-1}). Here, action is required because there is unacceptable risk to specified target.

6.3.6 Development of full-data point soil maps

The GC–MS-measured and vis-NIR-predicted soil maps for PAH, TTEC of the PAH mixture, and management zones were prepared using ArcGIS® 10.1 (ESRI Inc., USA). Data points were interpolated by the inverse distance weighting (IDW) method using the default Power function of 2 and standard neighbourhood type. IDW is a quick deterministic interpolator that is exact and can be a good way to take a first look at an interpolated surface, and there are no assumptions required of the data (ESRI, 2012). For both PAH and TTEC soil maps, seven classes were used. The classification method used for the PAH and TTEC maps was geometrical interval because it was designed to accommodate continuous data (ESRI, 2012). These classes were represented with colour ramps. For the management zones, between two and three classes were used depending on the TTEC range of the site. The area in m^2 of each zone was determined by converting the geographic coordinates of the data points in decimal degrees from geographic coordinate system (GCS) to projected coordinate system (PCS) of Nigeria Mina mid-belt national grid using ArcGIS®. Prior to kappa statistics analysis, the geodata were converted to raster dataset. We used a 251 by 366 raster grid of cell size 3.27×10^{-6} m by 3.27×10^{-6} m for site A, for site C we used a 393 by 251 raster grid of cell size 3.15×10^{-6} m by 3.15×10^{-6} m, and for site C we used a 322 by 251 raster grid of cell size 3.09×10^{-6} m by 3.09×10^{-6} m. In all three cases, the number of bands is 1 and the pixel depth is 32 Bit. The raster dataset for each measurement method was then reclassified using the same classification and interval. Both layers (maps) were then combined.

6.3.7 Statistical data analysis

We used one-way analysis of variance (ANOVA) available in Microsoft® Excel 2010 to evaluate the significance of differences between soil maps of PAH and TTEC of the PAH mixture developed by reference GC–MS and vis-NIR spectroscopic methods. Differences were considered significant at $p < 0.05$. Histograms were used to analyse the distribution of the error between the measured and predicted maps. Kappa statistics (Cohen, 1960) were used to compare the soil maps developed by the two different measurement methods. We used the guidelines proposed by Landis and Koch (1977) to interpret the goodness of the agreement between the maps on the basis of the kappa coefficient.

6.4 Results and Discussion

6.4.1 Full-data point soil maps

Figure 6-2 shows the full-data point IDW soil maps developed for the three petroleum release sites using measured and predicted PAH and TTEC data. For fair comparison between measured and predicted maps, equal number of classes (7 classes) was used for the three maps in each site with identical class intervals (Mouazen et al., 2007a). As shown in Figure 6-2, there is a large spatial similarity between the measured and predicted maps for both PAH and TTEC with a near-perfect match of high and low concentrations zones. However, some differences are observed around some interpolated surfaces of the predicted maps as shown in the error maps (Figure 6-2).

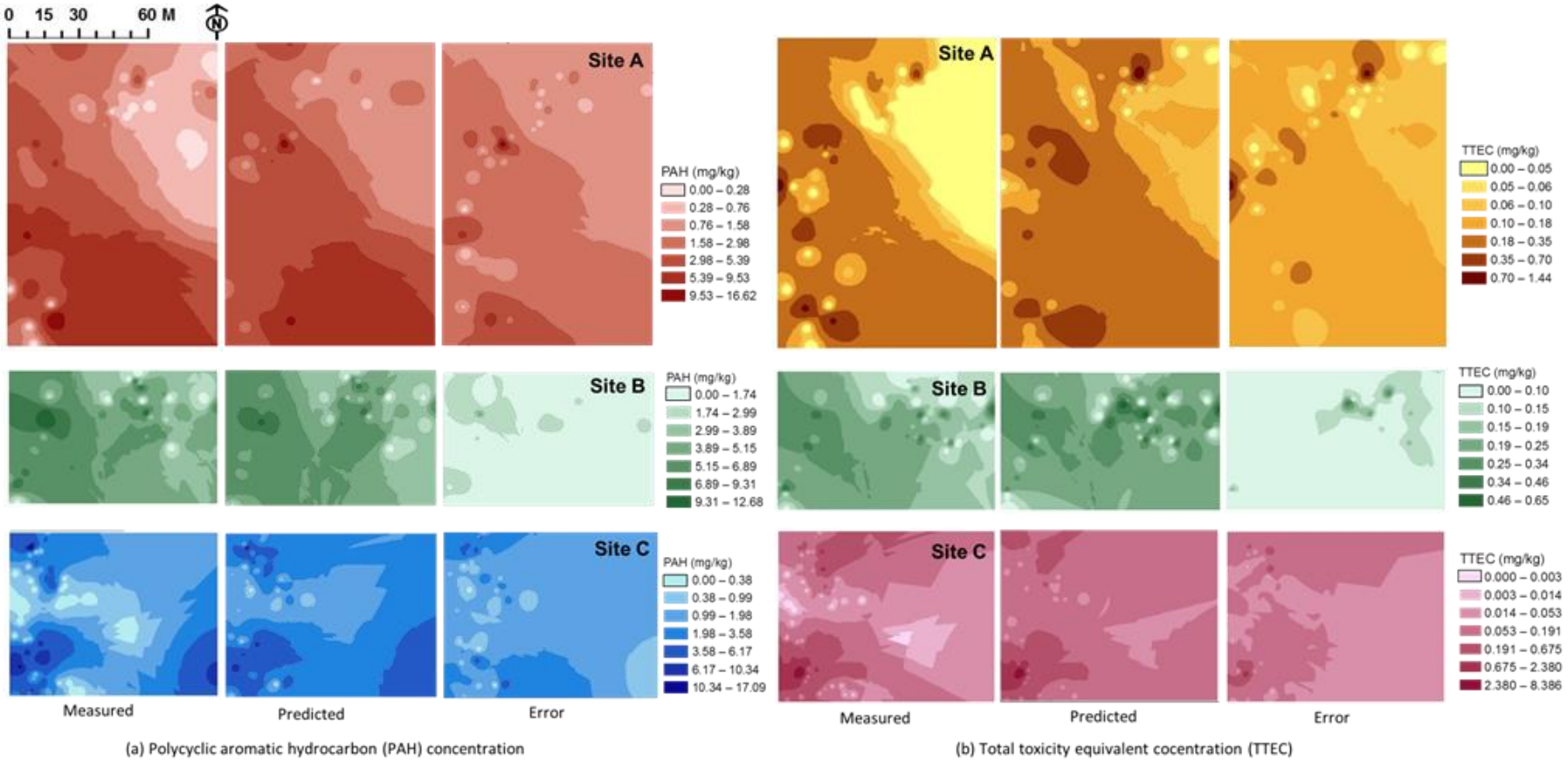


Figure 6-2 Comparison inverse distance weighting (IDW) soil maps for three petroleum release sites in Ogoniland in the Niger Delta province of Nigeria.

The histogram plot of the error between the two measurement methods for both PAH and TTEC is shown in Figure 6-3. For PAH measurement, the histogram shows that about 52 % of the error for site B with a PAH range of 12.68 mg kg^{-1} (see Figure 6-2) is less than 1 mg kg^{-1} in absolute values. Similarly, 49 % of the error for site A with a PAH range of 16.62 mg kg^{-1} (see Figure 2) and 42 % for site C with a PAH range of 17.09 mg kg^{-1} are less than 1 mg kg^{-1} in absolute values. The frequency of samples with low error is higher for TTEC measurement as compared to PAH measurement (Figure 6-3). Here, about 63 and 69% of the error are less than 0.2 mg kg^{-1} (absolute) for site A and C respectively while 62% are less than 0.1 mg kg^{-1} (absolute) for site B. Overall, the frequency of samples with low error is relatively larger than the frequency with high error (Figure 6-3). The histogram also shows that the error distribution is skewed for both PAH and TTEC measurements particularly in site A and C (Figure 6-3). The skewness in the negative range is largely attributed to the low concentration of PAH measured at some parts of the sites (Appendix E to Appendix G). As a result, PAH concentrations at these parts of the sites have been slightly overestimated by the vis-NIR method. There are also differences due to some cases of underestimation of the PAH as shown by the error distribution in the positive range (Figure 6-3). Nonetheless, these differences are non-significant ($p<0.05$) as the analysis of variance (ANOVA) reveals in Table 6-3.

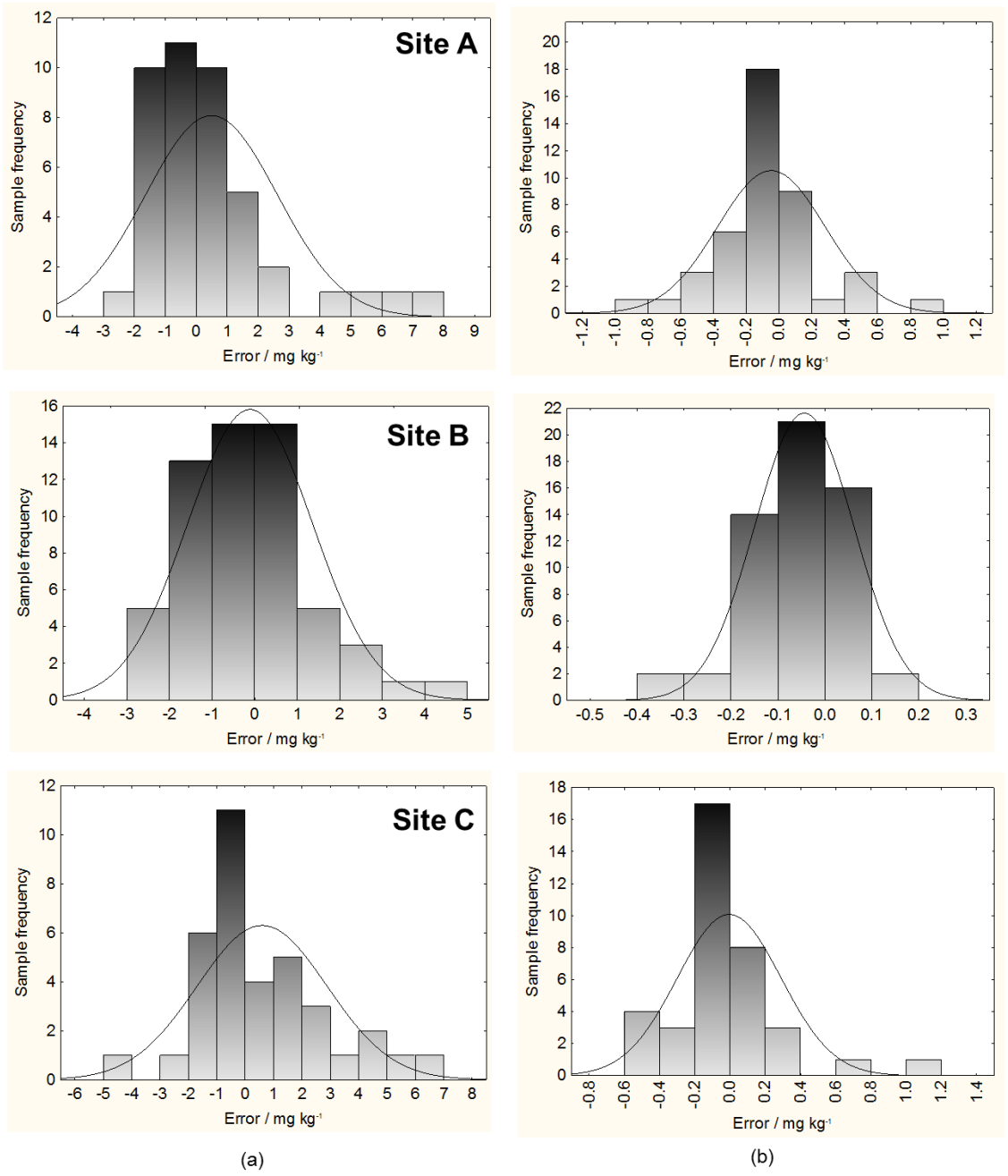


Figure 6-3 Histogram showing the distribution of error between measured and predicted soil maps of:

(a) polycyclic aromatic hydrocarbons (PAHs), and (b) total toxicity equivalent concentration (TTEC) of the PAH mixture in three contaminated sites at Ogoniland, Niger Delta province of Nigeria.

Table 6-3 One-way ANOVA on the analysis of the significance of the difference between measured and predicted soil maps of polycyclic aromatic hydrocarbon (PAH) and total toxicity equivalent concentration (TTEC) in three petroleum release sites in Ogoniland, Niger Delta province of Nigeria.

Site	Source of Variation	S.S.	*d.f.*	M.S.	*F*	*p* value	*F crit.*
Polycyclic aromatic hydrocarbon (PAH)							
Site A	Between Groups	0.05	1	0.05	0.01**	0.94	3.93
	Within Groups	1060.37	118	8.99	-	-	-
	Total	1060.42	119	-	-	-	-
Site B	Between Groups	0.07	1	0.07	0.01**	0.92	3.93
	Within Groups	770.78	104	7.41	-	-	-
	Total	770.85	105	-	-	-	-
Site C	Between Groups	2.12	1	2.12	0.23**	0.63	3.93
	Within Groups	1044.76	112	9.33	-	-	-
	Total	1046.88	113	-	-	-	-
Total toxicity equivalent concentration (TTEC)							
	Between Groups	0.08	1	0.08	1.38**	0.24	3.92
Site A							
	Within Groups	6.58	118	0.06			
	Total	6.66	119				
	Between Groups	0.06	1	0.06	3.15**	0.08	3.92
Site B							
	Within Groups	2.41	118	0.02			
	Total	2.47	119				
	Between Groups	0.00	1	0.001	0.001**	0.97	3.92
Site C							
	Within Groups	123.02	118	1.04			
	Total	123.02	119				

S.S., sum of squares; d.f., degrees of freedom; M.S., mean squares

**not significant ($p<0.05$)

The kappa statistic is used to compare two different classifications (reference GC-MS data and vis-NIR data) and to provide a value of how much they agree as shown in Table 6-4 to 6-6. This is often manifested in the Producer and User accuracies (not shown in Table 6-4 to 6-6). The Producer accuracy (usually in the vertical column) is how close one comes to the original GC-MS data, whereas the user accuracy (usually in the horizontal column) is an estimation of the accuracies when using the vis-NIR data. Here, we used two indicators to measure how much the two different data from GC-MS and vis-NIR agree – cross-tabulation and kappa coefficient. Although the value of the former may be higher, the latter is believed to be more accurate than the former because it also takes into account predictions made by chance (Cohen, 1960). As can be seen in Table 6-4 to 6-6, a non-zero kappa coefficient was achieved for each set of data in each site, which shows in the first instance an agreement between the dataset. However, the goodness of the agreement varied among the three sites. Since the soil maps are intended to be used for hazard assessment (.i.e., tier 1 risk assessment) and not for a quantitative risk assessment (tier 2), the guidelines proposed by Landis and Koch (1977) were considered sufficient to interpret the goodness of the agreement based on the kappa coefficient. The kappa coefficient of 0.56 achieved for site B (Table 6-5) shows a good agreement between GC-MS and vis-NIR data suggesting there is a 90% chance of correct predictions (Landis and Koch, 1977). For site A, the agreement is moderate with kappa coefficient of 0.42 (Table 6-4) indicating an 80% chance of correct predictions, while for site C the kappa coefficient is 0.19 (Table 6-6) showing a fair agreement where there is 70% chance of correct

predictions. The moderate and fair agreement achieved for site A and C respectively is mainly due to some overestimations by vis-NIR at some points in the sites. This was depicted in the error distribution discussed earlier. Nonetheless, this result shows promise that vis-NIR can be a good screening tool for PHCs in petroleum release sites.

Table 6-4 Kappa statistics for the comparison between soil maps of PAH developed by reference chemical GC-MS data and vis-NIR spectral data. The soils were collected from a petroleum contaminated site at Baraboo (Site A) in Ogoniland, Niger Delta province of Nigeria.

		Vis-NIR classification							Total
		1	2	3	4	5	6	7	
GC-MS classification	1	7	140	2146	75	54			2422
	2			13045	903	81			14029
	3			4738	4674	349			9761
	4	19	77	1631	8010	8165	244	102	18248
	5				728	19352	1498		21578
	6				80	6794	16972	125	23971
	7					111	380		491
Total		26	217	21560	14470	34906	19094	227	90500
Cross-tabulation									0.54
Kappa Coefficient									0.42

Table 6-5 Kappa statistics for the comparison between soil maps of PAH developed using reference chemical GC-MS data and vis-NIR spectral data. The soils were collected from a petroleum contaminated site at Bomu 1 (Site B) in Ogoniland, Niger Delta province of Nigeria.

		Vis-NIR classification							Total
		1	2	3	4	5	6	7	
GC-MS classification	1	376	548	78	1				1003
	2	20	1891	2554	160				4625
	3		722	13643	2288	21			16674
	4		1117	5846	16318	8476	39		31796
	5			11	3112	30515	119		33757
	6				179	3695	2248	10	6132
	7					121	771	181	1073
Total		396	4278	22132	22058	42828	3177	191	95060
Cross-tabulation									0.69
Kappa Coefficient									0.56

Table 6-6 Kappa statistics for the comparison between soil maps of PAH developed using reference chemical GC-MS data and vis-NIR spectral data. The soils were collected from a petroleum contaminated site at Bomu 2 (Site C) in Ogoniland, Niger Delta province of Nigeria.

		Vis-NIR classification							Total
		1	2	3	4	5	6	7	
GC-MS classification	1	66	789	1959	608	118			3540
	2		311	8418	3430	170			12329
	3		3	7205	24863	200			32271
	4		39	261	14052	4535			18887
	5		3	65	1355	9647	12		11082
	6					1748	254		2002
	7						62	4	66
Total		66	1145	17908	44308	16418	328	4	80177
Cross-tabulation									0.39
Kappa Coefficient									0.19

6.4.2 Potential management zones and hazard assessment of studied sites using GAC

The high spatial variability of PAH within the studied sites underscore the importance of management zones in the site-specific management of PAH-contaminated land. Its benefit in resource allocation and conservation during remediation projects cannot be overemphasized. In this study, management zones developed using measured and predicted data are compared in Figure 6-4. As shown in Figure 6-4, there is a near-perfect match of the geometry of the management zones between measured and predicted maps. The geometrical area of each zone within each site is summarised in Table 6-7. As shown in Table 6-7, most zones determined by vis-NIR method are almost equal in area to those determined by the conventional method. For instance for zone 1 and 2 in site C, their areas measured by vis-NIR method differ by a very small margin (about 1%) from those measured by the conventional method (Table 6-7). In this study, remarkably small zones were integrated into much larger zones, which in practice, allows for easy management (Halcro et al., 2013). This zones fusion and human cartographical error are partly responsible for differences in area between measured and predicted zones. Remarkably, vis-NIR method produces an equal number of management zones as the conventional method does in the studied sites. This result is remarkable because it leaves the area of the management zones as the main distinguishing feature between the two measurement methods. This means that other hazard assessment rubrics applicable to petroleum release sites based on GAC can be applied without distinction between the two measurement methods (Table 6-7).

As shown in Table 6-7, the degree of action for site-specific risk assessment and/or remediation differs among the different zones. These results suggest that vis-NIR method may be used for hazard assessment to inform site-specific risk assessment and/or remediation of petroleum release sites in the Niger Delta province of Nigeria.

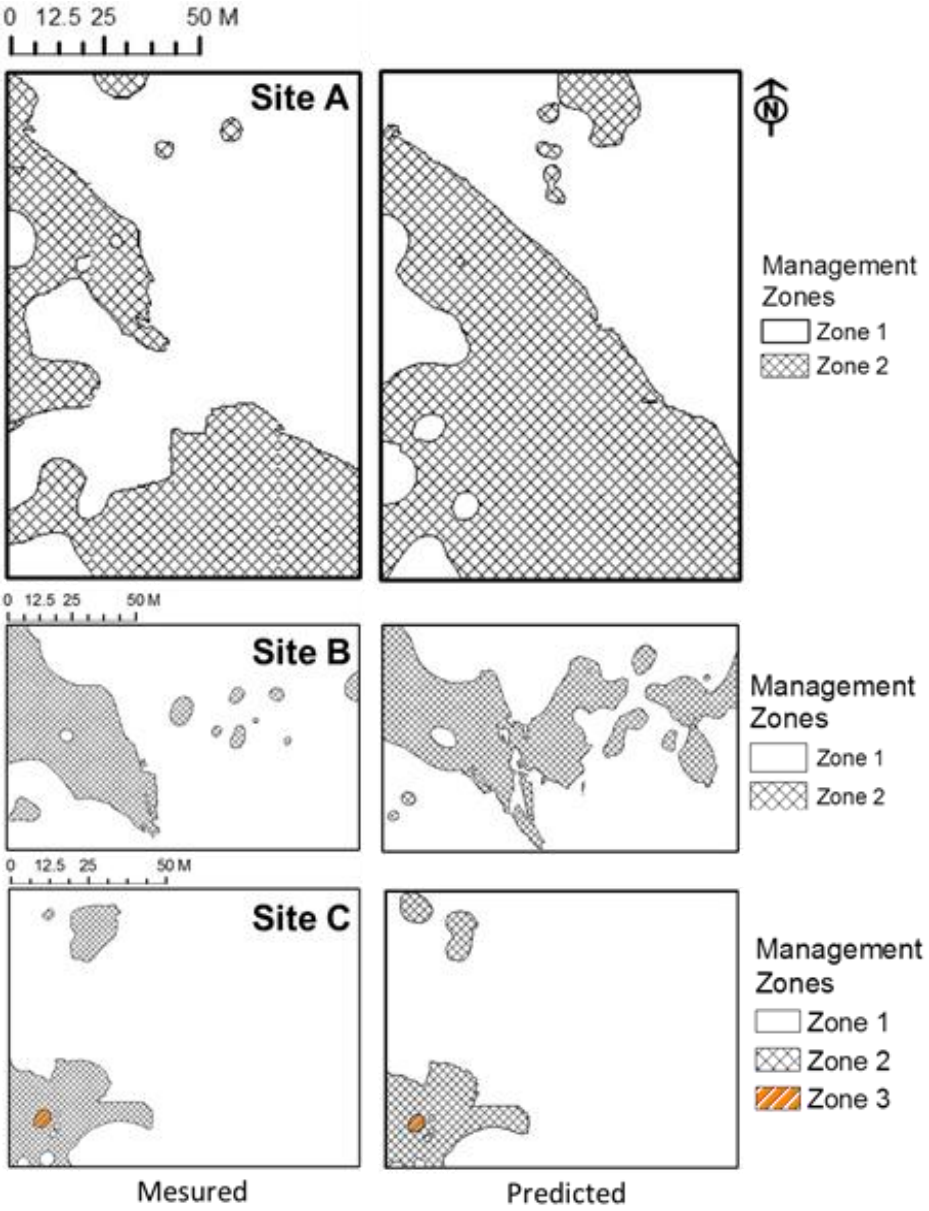


Figure 6-4 Comparison maps of potential management zones for three petroleum release sites in Ogoniland, Niger Delta province of Nigeria.

Table 6-7 Result of hazard assessment of three petroleum release sites in Ogoniland, Niger Delta province of Nigeria, using Generic Assessment Criteria (GAC) established in this study. The areal extent of contamination in the sites as determined by the conventional GC–MS and innovative vis-NIR methods is compared by management zone.

Site	MGZ	Hazard assessment rubrics				
		GAC (mg kg^{-1})	Risk impact	Action required	Areal extent of contamination (m^2)	
					GC–MS method	Vis-NIR method
A	1	< 0.24	Insignificant	None	7778	5473
	2	0.24–2.16	Significant and acceptable	Site-specific risk assessment	4421	6726
B	1	< 0.24	Insignificant	None	9303	8001
	2	0.24–2.16	Significant and acceptable	Site-specific risk assessment	2650	3952
C	1	< 0.24	Insignificant	None	9307	9433
	2	0.24–2.16	Significant and acceptable	Site-specific risk assessment	1164	1043
	3	> 2.16	Significant and unacceptable	Site-specific remediation	25	20

GAC, generic assessment criteria; GC–MS, gas chromatography–mass spectrometry; MGZ, management zone; vis-NIR; visible and near-infrared

6.5 Conclusions

This chapter investigated the possibility of using point vis-NIR reflectance spectroscopy for mapping of PAH and the level of toxicity in three petroleum release sites in Ogoniland in the Niger Delta province of Nigeria. Results obtained support the following conclusions:

- IDW soil maps of PAH and TTEC show high spatial variability of PAH and TTEC across the studied sites.
- There were non-significant (p<0.05) differences between measured and predicted soil maps of PAH and TTEC as the ANOVA revealed.
- There was a fair to good agreement between soil maps of PAH developed by reference GC-MS data and vis-NIR data (Kappa coefficient = 0.19–0.56).
- Potential management zones presented a near-perfect match between measured and predicted maps of high and low risk zones.
- The degree of action for site-specific risk assessment and/or remediation differs among the different zones. Only one zone covering an average of 22.5 m^2 in site C requires site-specific remediation as the PAH level exceeds the action value of 2.16 mg kg^{-1} set out as GAC for the studied sites.
- These results suggest that vis-NIR method may be a useful tool for hazard assessment to inform site-specific risk assessment and/or remediation of petroleum release sites in the Niger Delta, Nigeria. However, it should be emphasized that this study is nowhere close to

covering the numerous petroleum release sites up and down the region, and have not covered every likely source of variability in all petroleum release sites in the region. The large land area of the province may indicate large variability in soils, parent materials, colour, and other properties that have to be accounted for in the vis-NIR modelling procedure.

Chapter 7: General conclusions and future work

7.1 Conclusions

The overall goal of this study was to investigate the feasibility of using the vis-NIR method for diagnostic screening of soils contaminated with PAHs to inform site-specific risk assessment and/or remediation. To achieve this goal, the study was split into two parts – each designed to achieve a set of objectives. The first part of the study was laboratory-based while the second was field-based. The laboratory-based study was designed to assure that the methodology can produce reasonable prediction accuracy of PAHs under various soil conditions of moisture, texture, and oil concentration similar to field conditions. The individual and combined effects of these variables on the performance of vis-NIR spectroscopy (350–2500 nm) were investigated. On-site adaptive (OSA) trials at selected petroleum release sites in Ogoniland in the Niger Delta province of Nigeria was the original focus of the field study. But for logistics challenges, the OSA trials were conducted offsite in the UK with the contaminated soil samples from Nigeria.

7.1.1 Laboratory investigation

- The individual and combined effects of oil concentration, moisture and clay contents on the spectral characteristics of diesel-contaminated soils and calibration models developed by PLS regression analysis for PAH using vis-NIR spectroscopy were investigated. From results obtained, soil diffuse reflectance was found to decrease with increasing oil concentration, moisture content and clay content. This might be because of colour-associated variations in the visible region, and spectral

absorptions due to water, clay mineral and hydrocarbon in the NIR region.

- The individual and combined effects of oil concentration, clay content and soil moisture content were found to be significant ($p < 0.05$), although the combined effect of the three variables has the most significant ($p < 0.05$) influence on the accuracy of PAH determination with vis-NIR spectroscopy.

- Under laboratory conditions, the PLS calibration models for PAH prediction performed reasonably well. Excellent model accuracy was achieved at the lowest oil concentration and good accuracy at the least clay content. Excellent accuracy was also achieved at moisture content of 20% contrary to the notion that models developed with dry samples are always better than those developed with wet samples.

- Determination of PAH in soils without prior sample preparations might be possible with vis-NIR spectroscopy. This is because of the excellent model performance achieved for the field-moist intact soil samples.

- In the mean spectral reflectance curves of the contaminated soil samples, hydrocarbon absorption minima due to C–H stretching modes of ArCH chemical functional group apparent around 1647 nm in the first overtone region of the NIR spectrum might be linked to PAH.

- Of the five PAHs investigated, only phenanthrene could be predicted with reasonable accuracy. The concentrations of the poorly predicted PAHs including acenaphthylene, fluorene, anthracene, and pyrene were found to be very low (≤1.09 mg kg^{-1}) in the samples, which might be a reason they were not predicted well.

- The mechanism of prediction of PAH is seen to be dependent on the co-variation of PAH with other soil properties such as soil organic carbon and clay. This is because the inter-relationship between clay and soil organic carbon was found to influence the intensity of NIR spectral signal of sorbed PAH in soil. This clay and soil organic carbon interaction may be useful for evaluating the presence of PAH in soils by vis-NIR spectroscopy. Nonetheless, it is recommended that the unique spectral signal of sorbed phenanthrene observed in the current study should be investigated further in connection with decision-making for risk assessment and remediation of petroleum release sites.

7.1.2 Field investigation

- The results of the OSA trials revealed that soils diffuse reflectance decreases with increasing PAH concentration, which supports the finding of the laboratory study earlier.

- Positive regression coefficients observed around 1647 nm confirm a positive link to PAH, which also agrees with the result of the laboratory study.

- The PAH prediction accuracy of the generalized PLS model developed with combined data from the three sites was excellent demonstrating the possibility for quantitative application. However, it should be emphasized that the extrapolation of the model is limited to the three sites investigated in Nigeria. This is because, setting up a vis-NIR-based model for general application in the Niger Delta province of Nigeria would require a larger dataset covering both the concentration ranges and all the other sources of variability in oil spill sites in the region.

- IDW soil maps of PAH and TTEC of the PAH mixture developed with chemically-measured and vis-NIR-predicted data are comparable in terms of geometry and weight of interpolated surfaces with non-significant ($p < 0.05$) differences between measured and predicted soil maps. From the distribution of error, the frequency of samples with low margin of error was more than those with high error. There was a fair to good agreement between soil maps of PAH developed by reference GC-MS data and vis-NIR data (Kappa coefficient = 0.19–0.56)

- Potential management zones developed for hazard assessment based on GAC established in this study show a correspondence of high and low risk zones in the studied sites between measured and predicted maps. Results of the assessment using both measurement methods show that the impact of the contamination varied distinctly across the zones. Similarly, the type of action required for site-specific risk assessment and/or remediation varied correspondingly among the different zones.

Overall, the result suggests that vis-NIR method may be a useful tool for hazard assessment to inform site-specific risk assessment and/or remediation of petroleum release sites in the Niger Delta province of Nigeria.

7.2 Future work

Proposed work on the application of vis-NIR spectroscopy as a decision support tool for risk assessment and remediation of contaminated land as well as the quality assessment/quality control (QA/QC) of oily sludges and drill cuttings from crude oil production activities in Nigeria would consider the following:

1. On-line (tractor-mounted) vis-NIR system (Mouazen, 2006b) exists for on-line measurement of several soil properties. But because of the difficult terrain in which most oil spills occur in Nigeria, a tractor-mounted vis-NIR technology may not be realistic for application at the moment. Therefore, to further the implementation of the technology in Nigeria, the commercially available LabSpec®4 portable NIR spectrometer equipped with a GoLab® trolley (ASD Inc., USA) (Figure 7-1) would be employed.

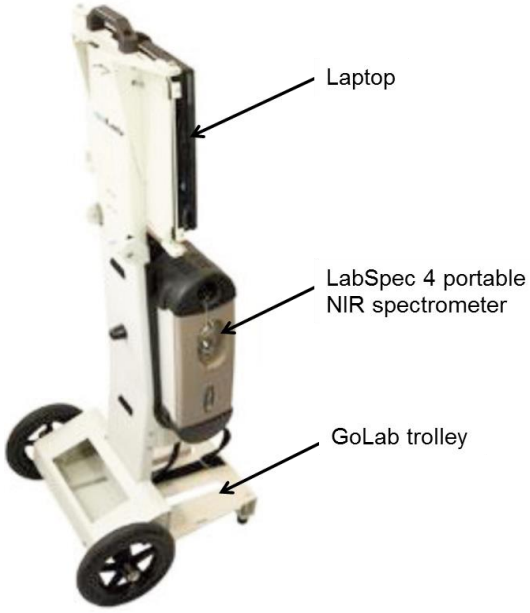


Figure 7-1 LabSpec®4 portable NIR spectrometer and GoLab® trolley (ASD Inc., USA) (Source: Analytik, CB, UK).

2. Application of the technology in the quality assessment/quality control (QA/QC) of hydrocarbon sludges from Terminal Improvement Projects (TIPs) and thermally desorbed oil-field drill cuttings for land application in Nigeria. This particular application is highly needed in Nigeria now more than ever before, not least because such wastes are now also known to be laden with heavy metals including chromium, lead, and barium (Ayotamuno et al., 2007). Again, covert investigation shows that most times such wastes are not characterised before final fate because of high analytical cost and the volume generated. Interestingly, some heavy metals like lead are spectrally active in the NIR band (Bray et al., 2010), and would most likely be detected by their spectral signature using vis-NIR spectroscopy.

3. Developing global calibration models based on vis-NIR spectroscopy for routine application in contaminated sites in Nigeria. To tackle calibration issues in Nigeria, the following principal steps proposed by Pasquini (2003) would be adopted:
    - Selection of calibration and test set of samples taking into account all sources of variability;
    - Determination of the concentration of PAH in the samples by reference GC–MS method;
    - Collection of NIR spectra in diffuse reflectance mode;
    - Development and optimisation of the mathematical calibration model;
    - Independent validation of the calibration model;
    - Application of the model in prediction of unknown samples; and
    - Regular maintenance of the model. Adopting methods like transformation, robust models, and model updating can reduce cost and time required for maintenance (Macho and Larrechi, 2002).
4. At the computer level, efforts would be made to test non-linear multivariate calibration techniques like artificial neural net-work or support vector machine for model development in hope that model prediction accuracy would be improved.

References

Abu, G. O., and Atu, N. D. (2008). An investigation of oxygen limitation in microcosm models in the bioremediation of a typical Niger Delta soil ecosystem impacted with crude oil. *Journal of Applied Science and Environmental Management*, 12, 13–22.

Adamchuk, V. I., Hummel, J. W., Morgan, M. T., and Upadhyaya, S. K. (2004). On-the-go soil sensors for precision agriculture. *Computer and Electronics in Agriculture*, 44, 71–91.

Adedokun, O. M. and Ataga, A. E. (2006). Effects of crude oil and oil products on growth of some edible mushrooms. *Journal of Applied Science and Environmental Management*, 10, 91–93.

Adesodun, J. K., and Mbagwu, J. S. C. (2008). Biodegradation of waste-lubricating petroleum oil in a tropical alfisol as mediated by animal droppings. *Bioresource Technology*, 99, 5659–5665.

Aldstadt, J., St. Germain, R., Grundl, T., and Schweitzer, R. (2002). *An in situ laser-induced fluorescence system for polycyclic aromatic hydrocarbon-contaminated sediments*. United States Environmental Protection Agency, Great Lakes National Program Office, Chicago, Illinois.

Aroh, K. N., Ubong, I. U., Eze, C. L., Harry, I. M., Umo-Otong, J. C., and Gobo, A. E. (2010). Oil spill incidents and pipeline vandalization in Nigeria – Impact on public health and negation to attainment of Millennium development goal: the Ishiagu example. *Disaster Prevention and Management*, 19(1), 70–87.

Askari, K., and Pollard, S. J. T. (2005). *The UK approach for evaluating human health risks from petroleum hydrocarbons in soils.* Science Report P5-080/TR3. Environment Agency, Bristol, UK. pp. 1–22.

Aske, N., Kallevik, H., and Sjoblom, J. (2001). Determination of saturate, aromatic, resin, and asphaltenic (SARA) components in crude oils by means of infrared and near-infrared spectroscopy. *Energy and Fuels*, 15, 1304–1312.

ASTM (1995). Standard Guide for Corrective Action Applied at Petroleum Release Sites. *ASTM E 1739-95.* American Society for Testing and Materials, West Conshohocken, PA, USA.

Atkins, P., and de Paula, J. (2002). *Atkins' Physical Chemistry*. 7th Edition. Oxford University Press, Great Clarendon Street, Oxford, UK.

Atlas, R. M. (1981). Microbial degradation of petroleum hydrocarbons: An environmental perspective. *Microbiology Reviews*, 45, 180–209.

Ayotamuno, M. J., Akor, A. J., and Igho, T. J. (2002). Effluent quality and wastes from petroleum drilling operations in the Niger Delta, Nigeria. *Journal of Environmental Management and Health*, 13(2), 207–216.

Ayotamuno, M. J., Okparanma, R. N., Ogaji, S. O. T., and Probert, S. D. (2007). Chromium removal from flocculation effluent of liquid-phase oil-based drill cuttings using powdered activated carbon. *Applied Energy*, 84, 1002–1011.

Bannert, D., Kuhn, F., and Horig, B. (1994). Ergebnisse spektroradiometrischer Untersuchungen an O lschieferaufschlu ssen in Pakistan. 14. Wissenschaftlich-Technische Jahrestagung der DGPF, Dresden. Oktober 5–10, pp. 171–178.

Barnes, B. (2009). *Framework for the use of rapid measurement techniques (RMT) in the risk management of land contamination*. Environment Agency, Bristol, UK. pp. 1-90.

Becker, R., Koch, M., Wachholz, S., Win, T. (2002). Quantification of total petrol hydrocarbons (TPH) in soil by IR-spectrometry and gas chromatography – conclusions from three proficiency testing rounds. *Accreditation and Quality Assurance*, 7, 286–289.

Benka-Coker, M. O., and Ekundayo, J. A. (1995). Effects of an oil spill on soil physico-chemical properties of a spill site in the Niger Delta area of Nigeria. *Environmental Monitoring and Assessment*, 36, 93–104.

Billets, S. (2001). *Innovative technology verification report – field measurement technologies for total petroleum hydrocarbons in soil*. United States Environmental Protection Agency (USEPA), Office of Research and Development, Washington DC.

Brassington, K. J., Hough, R. L., Paton, G. I., Semple, K. T., Risdon, G. C., Crossley, J,, Hay, I., Askari, K. and Pollard, S. J. T. (2007). Weathered hydrocarbon wastes: a risk management primer, *Critical Review in Environmental Science and Technology*, 37, 199–232.

Brassington, K. J., Pollard, S. J. T., and Coulon, F. (2010). Weathered hydrocarbon wastes: A risk management primer. In: K. N. Timmis (Editor). *Handbook of Hydrocarbon and Lipid Microbiology*, Springer-Verlag, Berlin Heidelberg. pp. 2488–2499.

Bray, J. G., Viscarra Rossel, R. A., McBratney, A. B. (2010). Diagnostic screening of urban soil contaminants using diffuse reflectance spectroscopy. In: R. A. Viscarra Rossel, A. B. McBratney, and B. Minasny (Editors), *Proximal Soil Sensing*, Springer-Verlag, Berlin Heidelberg. pp. 191–199.

British Standard Institute (2002). *Soil Quality – Gas Chromatographic Determination of the Content of Volatile Aromatic Hydrocarbons, Naphthalene and Volatile Halogenated Hydrocarbons – Purge-and-Trap Method with Thermal Desorption.* BS ISO 15009:2002. BSI, London.

British Standard Institute (2004). *Soil Quality – Determination of Content of Hydrocarbon in the Range C_{10} to C_{40} by Gas Chromatography*. BS ISO 16703:2004. BSI, London.

Brown, D. H. (2010). *Comparison of spectra from a Raman IdentiCheck versus an Ahura® TruScan® Raman Spectrometer*. PerkinElmer Inc., Shelton, CT.

Bujewski, G., and Rutherford, B. (1997). *The Rapid Optical Screening Tool (ROST™) laser-induced fluorescence (LIF) system for screening of petroleum hydrocarbons in subsurface soils.* U.S. Environmental Protection Agency, Office of Research and Development, Washington DC.

CAMO Software. (2012). Interpreting PLS plots. *The Unscrambler® X Version 10.2 User's Guide*. CAMO Software AS, Oslo, Norway.

Carlon, C. (2007). Derivation methods of soil screening values in Europe – A review and evaluation of national procedures towards harmonization. European Commission, Joint Research Centre, Ispra, EUR 22805-EN, 306 pp.

CEC (Commission of the European Communities) (2006). Proposal for a Directive of the European Parliament and the Council establishing a framework for the protection of soil and amending Directive 2004/35/Ec (presented by the Commission) 2006/0086/COD. Available at http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2006:0232:FIN:EN:PDF [accessed 29 Sep 2013].

Chakraborthy, S., Weindorf, D. C., Zhu, Li, Y., B., Morgan, C. L. S., Ge, Y., and Gulbraith, J. (2012). Spectral reflectance variability from soil physicochemical properties in oil contaminated soils. *Geoderma*, 177-178, 80–89.

Chakraborty, S., Weindorf, D. C., Morgan, C. L. S., Ge, Y., Galbraith, J. M., Li, B., and Kahlon, C. S. (2010). Rapid identification of oil-contaminated soils using visible near-infrared diffuse reflectance spectroscopy. *Journal of Environmental Quality*, 39, 1378–1387.

Chang, C. W., Laird, D. A., Mausbach, M. J., and Hurburgh Jr., C. R. (2001). Near-infrared reflectance spectroscopy—principal components regression analysis of soil properties. *Soil Science Society of America Journal*, 65, 480–490.

Cheng, Y., and Nathanail, P. C. (2009). Generic Assessment Criteria for human health risk assessment of potentially contaminated land in China. *Science of the Total Environment*, 408, 324–339.

Chuang, J. C., Van Emon, J. M., Chou, Y. L., Junod, N., Finegold, J. K., and Wilson, N. K. (2003). Comparison of immunoassay and gas chromatography-mass spectrometry for measurement of polycyclic aromatic hydrocarbons in contaminated soil. *Analytica Chimica Acta*, 486, 31–39.

Clark, R. N., King, T. V. V., Klejwa, M., Swayze, G., and Vergo, N. (1990). High spectral resolution reflectance spectroscopy of minerals. *Journal Geophysical Research*, 95, 12653–12680.

Cloutis, E. A. (1989). Spectral reflectance properties of hydrocarbons: remote-sensing implications. *Science*. 245, 87–89.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*. 20, 37–46.

Contaminated Land: Applications in Real Environments. (2010). Bioremediation of heavy hydrocarbons – reducing uncertainty in meeting risk-based targets: laboratory to field scale (PROMISE Project). *Contaminated Land: Applications in Real Environments (CL:AIRE) Research Bulletin*, RB 10, 1–8.

Creighton, K., and Richards, R. (1997). Field screening technique for total petroleum hydrocarbons in soils. *Practice Periodical of Hazardous, Toxic and Radioactive Waste Management*, 1, 78–83.

Current, R. W., and Tilotta, D. C. (1997). Determination of total petroleum hydrocarbons in soil by on-line supercritical fluid extraction-infrared spectroscopy using a fibre-optic transmission cell and a simple filter spectrometer. *Journal of Chromatography A*, 785, 269–277.

Daka, E. R. and Ekweozor, I. K. E. (2004). Effect of size on the acute toxicity of crude oil to the Mangrove Oyster (Carasostrea gasar). *Journal of Applied Science and Environmental Management*, 8, 19–22.

Daniel-Kalio, L. A. and Pepple, S. F. (2006). Effect of Bonny Light crude oil pollution of soil on the growth of dayflower (Commelina benghalensis L.) in the Niger Delta, Nigeria. *Journal of Applied Science and Environmental Management*, 10, 111–114.

Davies, A. M. C. (2005). Introduction to NIR spectroscopy. In: A. M. C. Davies (Editor). *NIR News*, 16(7), pp. 9–11.

De Meulenaer, B. (2006). Chemical hazards. In: P. A. Luning, F. Devlieghere, R. Verhe (Editors). *Safety in the Agri-food Chain.* Wageningen Academic Publishers, The Netherlands, pp.145–208.

Department of Petroleum Resources (2002). *Environmental Guidelines and Standards for the Petroleum Industry in Nigeria (EGASPIN).* Ministry of Petroleum and Natural Resources, Abuja, Nigeria. pp. 314.

Dexter, A. R., Richard, G., Arrouays, D., Czyz, E. A., Jolivet, C., and Duval, O. (2008). Complexed organic matter controls soil physical properties. *Geoderma*, 144, 620–627.

Dos Santos, M. L., Simoes, M. L., de Melo, W. J., Martin-Neto, L., and Pereira-Filho, R. (2010). Application of chemometric methods in the evaluation of chemical and spectroscopic data on organic matter from Oxisols in sewage sludge applications. *Deoderma*, 155, 121–127.

Douglas, G. S., McCaarthy, K. J., Dahlen, D. T., Seavey, J. A., Steinhauer, W. G., Prince, R. C., and Elmendorf, D. L. (1992). The use of hydrocarbon analyses for environmental assessment and remediation. Journal of Soil Contamination, 1, 197–216.

Ebuehi, O. A. T., Abibo, I. B., Shekwolo, P. D., Sigismund, K. T., Adoki, A. and Okoro, I. C. (2005). Remediation of crude oil polluted soil by enhanced natural attenuation. *Journal of Applied Science and Environmental Management*, 9, 103–106.

Ekundayo, E. O., and Obuekwe, C. O. (1997). Effects of an oil spill on soil physico-chemical properties of a spill site in a typic *Paleudult* of Midwestern Nigeria. *Environmental Monitoring and Assessment*, 45, 209–221.

EPA Method 1664. (1999b). n-Hexane extractable material (HEM; oil & grease) and silica gel treated n-hexane extractable material (SGT-HEM, non-polar material) by extraction and gravimetry. United States Environmental Protection Agency, Office of Water, Washington, DC, USA.

EPA Method 4030. (1996a). Soil screening for petroleum hydrocarbons by Immunoassay. *SW-846 Manual*, United States Environmental Protection Agency, Government Printing Office, Washington, DC, USA.

EPA Method 4035. (1996b). Soil screening for polynuclear aromatic hydrocarbons by immunoassay. *SW-846 Manual*, United States Environmental Protection Agency, Government Printing Office, Washington, DC, USA.

EPA Method 418.1. (1978). Total Recoverable Petroleum Hydrocarbons by IR. Government Printing Office, Washington, DC, USA.

EPA Method 625. (1984). Base/neutral, acids and pesticides using GC/MS. EPA 600/4-84-053, National Technical Information Service, PB84-206572, Springfield, Virginia, USA.

EPA Method 8015. (2000). Non-halogenated organics using GC/FID. *SW-846 Manual*, United States Environmental Protection Agency, Government Printing Office, Washington, DC, USA.

EPA Method 8270. (1999a). Semi volatile organic compounds (SVOCS) by gas chromatography/mass spectrometry (GC/MS). *SW-846 Manual*, United States Environmental Protection Agency, Government Printing Office, Washington, DC, USA.

ESRI (2012). *How inverse Distance Weighted Interpolation Works. ArcGIS® Version 10.1 – Help: Geostatistical Analyst.* ESRI Inc., USA.

Essington, M. E. (2004). *Soil and water chemistry: An integrative approach.* CRC Press Inc., Boca Raton, FL, USA.

Eyvazi, M. J., and Zytner, R. G. (2009). A Correlation to estimate the bioventing degradation rate constant. *Bioremediation Journal*, 13, 141–153.

Fan, C. Y., Krishnamurthy, S., and Chen, C. T. (1994). A critical review of analytical approaches for petroleum contaminated soil. In: T. A. O'Shay and K. B. Hoddinott (Editors). *Analysis of soil contaminated with petroleum constituents*. American Society for Testing and Materials ASTM STP 1221, Philadelphia, PA. pp. 61–74.

Ferguson, C., Darmendrail, D., Freier, K., Jensen, B. K., Jensen, J., Kasamas, H., Urzelai, A., and Vegter, J. (1998). *Risk assessment for contaminated sites in Europe – Scientific Basis*. LQM Press, Nottingham, United Kingdom.

Forrester, S. (2010). In-situ determination of total petroleum hydrocarbon (TPH) contamination: a quick infrared spectroscopic test for TPH at contaminated sites. Poster presented at the 19th world congress of soil science, soil solutions for a changing world, Brisbane, Australia. August 1–6.

Forrester, S. F., Janik, L. J., McLaughlin, M. J. (2011). Method of contaminant prediction. PCT Patent WO/2011/035391. Date issued: 31 March.

Forrester, S. T., Janik, L. J., McLaughlin, M. J., Soriano-Disla, J. M., Stewart, R., and Dearman, B. (2013). Total petroleum hydrocarbon concentration prediction in soils using diffuse reflectance infrared spectroscopy. *Soil Science Society of America Journal*, 77, 450–460.

Forrester, S., Janik, L., and McLaughlin, M. (2010). An infrared spectroscopic test for total petroleum hydrocarbon (TPH) contamination in soils. In: *Proceedings of the 19th world congress of soil science, soil solutions for a changing world*, Brisbane, Australia. August 1–6, pp. 13–16.

Frysinger, G., Gaines, R. B., Xu, L., and Reddy, C. M. (2003). Resolving the unresolved complex mixture in petroleum-contaminated sediments. *Environmental Science and Technology*, 37, 1653–1662.

Fuller, M. P., and Griffiths, P. R. (1978). Diffuse reflectance measurements by infrared Fourier transform spectrometry. *Analytical Chemistry*, 50(13), 1906–1910.

Gammoun, A., Moros, J., Tahiri, S., Garrigues, S., and de la Guardia, M. (2006). Partial least-squares near-infrared determination of hydrocarbons removed from polluted waters by using tanned solid wastes. *Analytical and Bioanalytical Chemistry*, 385, 766–770.

GE Security Inc. (2006). *Raman spectroscopy: the science behind the technology*. GE Homeland Protection Inc. Available at http://www.gesecurity.com [accessed 5 September 2012].

González, O. M., Kuang, B., Quraishi, M. Z., García, M. Á. M., and Mouazen, A. M. (2013). On-line measurement of soil properties without direct spectral response in near infrared spectral range. *Soil and Tillage Research*, 132, 21–29.

Graham, K. N. (1998). Evaluation of analytical methodologies for diesel fuel contaminants in soil. M.Sc. Thesis. University of Manitoba, Canada.

Greason, S. (2009). *Field screening petroleum hydrocarbons using ultraviolet fluorescence technology*. Sitelab Corporation, West Newbury, MA.

Halcro, G., Corstanja, R., and Mouazen, A. M. (2013). Fusion of proximal soil sensing and crop data for fertility zone delineation. The 3rd Global Workshop on Proximal Soil Sensing, Potsdam, Germany. May 26–29.

Harris, C. M. (2003). Today's chemist at work. *American Chemical Society*, 33–38.

Harris, D. C. (2010). *Quantitative Chemical Analysis*, 8th edition. W. H. Freeman and Company, New York.

Hites, R. A., and Gschwend, P. M. (1982). The ultimate fate of polycyclic aromatic hydrocarbons in marine and lacustrine sediments. In: M. Cooke, A. J. Dennis, and F. L. Fisher (Editors). *Poly-nuclear Aromatic Hydrocarbons – Physical and Biological Chemistry*. Bettelle Press, Columbus, Ohio, USA. pp. 357–365.

Horig, B., Kuhn, F., Oschutz, F., and Lehmann, F. (2001). HyMap hyperspectral remote sensing to detect hydrocarbons. *International Journal of Remote Sensing*, 22(8), 1413–1422.

Huang, W., Peng, P., Yu, Z. and Fu, J. (2003). Effects of organic matter heterogeneity on sorption and desorption of organic contaminants by soils and sediments. *Applied Geochemistry*, 18, 955–972.

ISO 16703. (2004). Soil quality – determination of content of hydrocarbon in the range C10 to C40 by gas chromatography. ISO, Geneva, Switzerland.

ISO/DIS 16703. (2001). Soil quality – determination of mineral oil content by gas chromatography. ISO, Geneva, Switzerland.

ISO/TR 11046. (1992). Soil quality – determination of mineral oil content – method A: infrared screening. ISO, Geneva, Switzerland.

Jack, I. R., Fekarurhobo, G. K., and Igwe, F. U. (2005). Determination of hydrocarbons levels in some marine organisms from some towns within the Rivers State of Nigeria. *Journal of Applied Science and Environmental Management*, 9, 59–61.

Janik, L. J., and Skjemstad, J. O. (1995). Characterization and analysis of soils using mid-infrared partial least-squares. II. Correlations with some laboratory data. *Australian Journal Soil Research*, 33, 637–650.

Janik, L. J., Skjemstad, J. O., and Raven, M. D. (1995). Characterization and analysis of soils using mid-infrared partial least-squares. I. Correlations with XRF-determined major element composition. *Australian Journal Soil Research*, 33, 621–636.

Jehlicka, J., Edwards, H. G. M., Villar, S. E. J. and Pokorny, J. (2005). Raman spectroscopic study of amorphous and crystalline hydrocarbons from soil, peats and lignite. *Spctrochimica Acta*, A61, 2390–2398.

Kadafa, A. A. (2012). Oil exploration and spillage in the Niger Delta of Nigeria. *Civil and Environmental Research*, 2(3), 38–51.

Katayama, Y., Oura, T., Iizuka, M., Orita, I., Cho, K. J., Chung, I. Y. and Okada, M. (2003). Effects of spilled oil on microbial communities in a tidal flat. *Marine Pollution Bulletin*, 47, 85–90.

Kim, I. S., Ritchie, L., Setford, S., Taylor, J., Allen, M., Wilson, G., Heywood, R., Pahlavanpour, B., and Saini, S. (2001). Quantitative immunoassay for determining polycyclic aromatic hydrocarbons in electrical insulating oils. *Analytica Chimica Acta*, 450, 13–25.

Kodaira, M., and Shibusawa, S. (2013). Using a mobile real-time soil visible-near infrared sensor for high resolution soil property mapping. *Geoderma*, 199, 64–79.

Krupcik, J., Oswald, P., Oktavec, D., and Armstrong, D. W. (2004). Calibration of GC–FID and IR-spectrometric methods for determination of high boiling petroleum hydrocarbons in environmental samples. *Kluwer Academic Publishers*, 153, 329–341.

Kuang, B., and Mouazen, A. M. (2011). Calibration of a visible and near infrared spectroscopy for soil analysis at field scales across three European farms. *European Journal of Soil Science,* 62(4), 629–636.

Kuang, B., and Mouazen, A. M. (2012). Influence of the number of samples on prediction error of visible and near-infrared spectroscopy of selected soil properties at the farm scale. *European Journal of Soil Science*, 63(3), 421–429.

Kuang, B., and Mouazen, A. M. (2013). Effect of spiking strategy and ratio on calibration of on-line visible and near infrared soil sensor for measurement in European farms. *Soil and Tillage Research*, 128, 125–136.

Kuang, B., Mahmood, H. S., Quraishi, Z., Hoogmoed, W. B., Mouazen, A. M., and van Henten, E. J. (2012). Sensing soil properties in the laboratory, in situ, and on-line: a review. *Advances in Agronomy*, 114, 155–224.

Kuhn, F., and Horig, B. (1995). Environmental remote sensing for military exercise places. *Remote Sensing and GIS for Site Characterizations: Applications and Standards,* ASTM STP 1279, American Society for Testing and Materials, West Conshohocken, PA, USA. pp. 5–16.

Kuhn, F., Oppermann, K., and Horig, B. (2004). Hydrocarbon Index – an algorithm for hyperspectral detection of hydrocarbons. *International Journal of Remote Sensing*, 25(12), 2467–2473.

Lambert, P., Fingas, M., and Goldthorp, M. (2001). An evaluation of field total petroleum hydrocarbon (TPH) systems. *Journal of Hazardous Materials*, 83, 65–81.

Landis, J. R., and Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33, 159–74.

Latimer, J. S., and J. Zheng. (2003). The sources, transport and fate of PAHs in the marine environment. In: P. E. T. Douben (Editor). *PAHs: An eco-toxicological perspective*. John Willy and Sons Ltd., USA. pp. 10–22.

Li, S. and Dai, L. (2012). Classification of gasoline brand and origin by Raman spectroscopy and a novel R-weighted LSSVM algorithm. *Fuel*, 96, 146–152.

Lohmannsroben, H. G. and Roch, T. (2000). In situ laser-induced fluorescence (LIF) analysis of petroleum product-contaminated soil samples. *Journal of Environmental Monitoring,* 2, 17–22.

Lorenzi, D., Cave, M., and Dean, J. R. (2010). An investigation into the occurrence and distribution of polycyclic aromatic hydrocarbons in two soil size fractions at a former industrial site in NE England, UK using in situ PFE-GC-MS. *Environmental Geochemistry and Health*, 32, 553–565.

Lyon, A., Keating, C. D., Fox, A. P., Baker, B. E., He, L., Nicewarner, S. R., Mulvaney, S. P., and Natan, M. J. (1998). Raman Spectroscopy. *Analytical Chemistry*, 70, 341R–361R.

Macho, S., and Larrechi, M. S. (2002). Near-infrared spectroscopy and multivariate calibration for the quantitative determination of certain properties in the petrochemical industry. *Trends in Analytical Chemistry*, 21(12), 799–806.

Maddams, W. F., and Royaud, I. A. M. (1990). The characterization of polycyclic aromatic hydrocarbons by Raman spectroscopy. *Spctrochimica Acta*, 46A, 309–314.

Maleki, M. R., Van Holm, L., Ramon, H., Merckx, R., De Baerdemaeker, J., and Mouazen, A. M. (2006). Phosphorus sensing for fresh soils using visible and near infrared spectroscopy. *Biosystems Engineering*, 95, 239–250.

Malle, H., and Fowlie, P. (1998). A Canadian interlaboratory comparison for analysis of petroleum hydrocarbons in soil. In: *Proceedings of the second biennial international conference on chemical measurement and monitoring of the environment, EnviroAnalysis '98 Conference*, Ottawa, Canada. May 11–14, pp. 321–322.

Malley, D. F., Hunter, K. N., and Barrie-Webster, G. R. (1999). Analysis of diesel fuel contamination in soils by near-infrared reflectance spectrometry and solid phase micro extraction-gas chromatography. *Journal of Soil Contamination*, 8, 481–489.

Martens, H., and Naes, T. (1989). *Multivariate Calibration,* 2nd Edition. John Wiley and Sons, Chichester, UK.

Masucci, J. A., and Caldwell, G. W. (2004). Techniques for gas chromatography/mass spectroscopy. In: R. L. Grob and E. F. Barry (Editors). *Modern Practice of Gas Chromatography*, 4th Edition. John Willy and Sons Inc., Hoboken, NJ. pp. 339–401.

Mitra, S. (2003). *Sample Preparation Techniques in Analytical Chemistry*. Wiley and Sons Inc., Publication, Hoboken, NJ, USA.

Mouazen, A. M. (2006b). Soil Survey Device. International publication published under the patent cooperation treaty (PCT). World Intellectual Property Organization, International Bureau. International Publication Number: WO2006/015463; PCT/BE2005/000129; IPC: G01N21/00; G01N21/00.

Mouazen, A. M., De Baerdemaeker, J., and Ramon, H. (2005a). Towards development of on-line soil moisture content sensor using a fibre-type NIR spectrophotometer. *Soil and Tillage Research*, 80, 171–183.

Mouazen, A. M., De Baerdemaeker, J., Ramon, H. (2006a). Effect of wavelength range on the measurement accuracy of some selected soil constituents using visual-near infrared spectroscopy. *Journal of Near Infrared Spectroscopy*, 14, 189–199.

Mouazen, A. M., Karoui, R., Deckers, J., De Baerdemaeker, J., Ramon, H. (2007b). Potential of visible and near-infrared spectroscopy to derive colour groups utilizing the Munsell soil colour charts. *Biosystems Engineering*, 97, 131–143.

Mouazen, A. M., Kuang, B., De Baerdemaeker, J., and Ramon, H. (2010). Comparison among principal component, partial least squares and back propagation neural network analyses for accuracy of measurement of selected soil properties with visible and near infrared spectroscopy. *Geoderma*, 158, 23–31.

Mouazen, A. M., Maleki, M. R., De Baerdemaeker, J., and Ramon, H. (2007a). On-line measurement of some selected soil properties using a VIS–NIR sensor. *Soil and Tillage Research*, 93(1), 13–27.

Murray, H. (2012). Overview and comparison of portable spectroscopy techniques: FTIR, NIR, and Raman. In: H. Murray (Editor), *Focus on Spectroscopy*. Labmate, St. Albans, Hertfordshire, UK.

Naes, T., Isaksson, T., Fearn, T., and Davies, T. (2002). *A user friendly guide to multivariate calibration and classification*. NIR Publications, Chichester, UK.

Nguyen, T. T., Janik, L. J., Raupach, M. (1991). Diffuse reflectance infrared Fourier transform (DRIFT) spectroscopy in soil studies. *Australian Journal of Soil Research*, 29, 49–67.

Niger Delta Environmental Survey (1995). *Background and Mission: Briefing Note 1.* Publication of the Steering Committee, NDES, Falomo, Lagos, Nigeria. pp. 1–7.

Nisbet, I., and LaGoy, P. (1992). Toxic equivalency factors (TEFs) for polycyclic aromatic hydrocarbons (PAHs). *Regulatory Toxicology and Pharmacology*, 16, 290–300.

Nocita, M., Stevens, A., Noon, C., and van Wesemael, B. (2013). Prediction of soil organic carbon for different levels of soil moisture using vis–NIR spectroscopy. *Geoderma*,199, 37–42.

Nording, M., Frech, K., Persson, Y., Forsman, M., and Haglund, P. (2006). On the semi-quantification of polycyclic aromatic hydrocarbons in contaminated soil by an enzyme-linked immunosorbent assay kit. *Analytica Chimica Acta*, 555, 107–113.

Ogri, O. R. (2001). A review of the Nigerian petroleum industry and the associated environmental problems. *The Environmentalist*, 21, 11–21.

Okoro, D., and Ikolo, A. O. (2007). Sources and compositional distribution of polycyclic aromatic hydrocarbons in soils of Western Niger Delta. *Journal of Applied Science and Technology*, 12(1-2), 35–40.

Okparanma, R. N., and Mouazen, A. M. (2013a). Determination of total petroleum hydrocarbon (TPH) and polycyclic aromatic hydrocarbon (PAH) in soils: a review of spectroscopic and non-spectroscopic techniques. *Applied Spectroscopy Reviews*, 48(6), 458–486.

Okparanma, R. N., and Mouazen, A. M. (2013b). Combined effects of oil concentration, clay and moisture contents on diffuse reflectance spectra of diesel-contaminated soils. *Water, Air and Soil Pollution*, 224(5), 1539–1556.

Okparanma, R. N., and Mouazen, A. M. (2013c). Visible and near-infrared spectroscopy analysis of a polycyclic aromatic hydrocarbon in soils. *The Scientific World Journal*. Article in press.

Okparanma, R. N., Coulon, F., and Mouazen, A. M. (2014a). Analysis of petroleum-contaminated soils by diffuse reflectance spectroscopy and sequential ultrasonic solvent extraction–gas chromatography. *Environmental Pollution*, 184, 298–305.

Okparanma, R. N., Coulon, F., Mayr, T., and Mouazen, A. M. (2014b). Mapping polycyclic aromatic hydrocarbon and total toxicity equivalent soil concentrations for hazard assessment by visible and near infrared spectroscopy. *Science of the Total Environment.* Article in review.

Osborne, B. G., Fearn, T., and Hindle, P. H. (1993). *Practical NIR spectroscopy – with applications in food and beverage analysis*. 2nd Edition. Longman Group UK Limited, Essex, England.

Pasquini, C. (2003). Near infrared spectroscopy: fundamentals, practical aspects and analytical applications. *Journal of the Brazilian Chemical Society*, 14(2), 198–219.

Patnaik, P. (1997). *Handbook of environmental analysis: chemical pollutants in air, water, soil and solid wastes*. CRC Press Inc., Boca Raton, FL.

Pérez-Caballero, G., Andrade, J., Muniategui, S., and Prada, D. (2009). Comparison of single-reflection near-infrared and attenuated total reflection mid-infrared spectroscopies to identify and monitor hydrocarbons spilled in the marine environment. *Analytical and Bioanalytical Chemistry*, 395(7), 2335–2347.

Peterson, G. S., Axler, R. P., Lodge, K. B., Schuldt, J. A. and Crane, J. L. (2002). Evaluation of a fluorometric screening method for predicting total PAH concentrations in contaminated sediments. *Environmental Monitoring and Assessment*, 78, 111–129.

Petts, J., Cairney, T., and Smith, M., (1997). *Risk-based Contaminated Land Investigation and Assessment.* John Willey and Sons Ltd., Chichester, England.

Pfannkuche, J., Lubecki, L., Schmidt, H., Kowalewska, G., and Kronfeldt, H. (2012). The use of surface-enhanced Raman scattering (SERS) for detection of PAHs in the Golf of Gdansk (Baltic Sea). *Marine Pollution Bulletin*, 64, 614–626.

Pollard, S. J. T., Kenefick, S. L., Hrudy, S. F., Fuhr, B. J., Holloway, L. R., and Rawluk, M. (1994). A tiered analytical protocol for the characterisation of heavy oil residues at petroleum-contaminated hazardous waste sites. In: T. A. O'Shay and K. B. Hoddinott (Editors). *Analysis of soil contaminated with petroleum constituents*. American Society for Testing and Materials, Philadelphia, PA. pp. 38–52.

Poster, D. L., Schantz, M. M., Sander, L. C., and Wise, S. A. (2006). Analysis of polycyclic aromatic hydrocarbons (PAHs) in environmental samples: a critical review of gas chromatographic (GC) methods. *Analytical and Bioanalytical Chemistry*, 386, 859–881.

Prommer, H., Davis, G. B., Barry, D. A., and Miller C. T. (2003). Modelling the fate of petroleum hydrocarbons in groundwater. In: A. Langley, M. Gilbey, and B. Kennedy (Editors). Proceedings of the 5th national workshop on the assessment of site contamination, Adelaide, Australia. May 2002.

Puppels, G. J., Colier, W., Olminkhof, J. H. F., Otto, C., de Mul, F. F. M., and Greve, J. (1991), Description and performance of a highly sensitive confocal Raman micro spectrometer. *Journal of Raman Spectroscopy*, 22, 217–225.

Quick Results On Site. (2012). *Hydrocarbon Analysis with QED*. Available at http://www.qros.co.uk/hydrocarbon_analysis.html [accessed 5 August 2012].

Quraishi, M. Z., and Mouazen, A. M. (2013). Development of a methodology for in situ assessment of topsoil dry bulk density. *Soil and Tillage Research*, 126, 229–237.

Rhodes, I. A. L., Olvera, R. Z., and Leon, J. A. (1990). Determination of gasoline range total petroleum hydrocarbon and approximate boiling point distribution in soil by gas chromatography. In: P. T. Kostecki, and E. J. Calabrese (Editors). *Hydrocarbon Contaminated Soils*, Lewis Publishers, MI, USA. pp. 273–290.

Risdon, G. C., Pollard, S. J. T., Brassington, K. J., McEwan, J. N., Paton, G. I., Semple, K. T., and Coulon, F. (2008). Development of an analytical procedure for weathered hydrocarbon contaminated soils within a UK risk-based framework. *Analytical Chemistry,* 80, 7090–7096.

Saari, E., Peramaki, P., and Jalonen, J. (2007). Effect of sample matrix on the determination of total petroleum hydrocarbons (TPH) in soil by gas chromatography-flame ionization detection. *Microchemical Journal*, 87, 113–118.

Saari, E., Peramaki, P., and Jalonen, J. (2010). Evaluating the impact of GC operating settings on GC–FID performance for total petroleum hydrocarbon (TPH) determination. *Microchemical Journal*, 94, 73–78.

Schneider, I., Nau, G., King, T. V. V., and Aggarwal, I. (1995). Fibre-optic near-infrared reflectance sensor for detection of organics in soils. *IEEE Photonics Technology Letters*, 7, 87–89.

Schultze, R. H., and Lewitzka, F. (2005). On-site and in-situ analysis of contaminated soils using laser induced fluorescence spectroscopy. In: M. Ehlers, and U. Michel (Editors). *Remote Sensing for Environmental Monitoring, GIS Applications, and Geology V*. Proceedings of SPIE, 5983, pp. 1–10.

Schwartz, G., Ben-Dor, E., and Eshel, G. (2012). Quantitative analysis of total petroleum hydrocarbons in soils: comparison between reflectance spectroscopy and solvent extraction by 3 certified laboratories. *Applied Environmental Soil Science*, 2012, 1–11.

Schwartz, G., Eshel, G., and Ben-Dor, E. (2011). Reflectance spectroscopy as a tool for monitoring contaminated soils. In: S. Pascucci (Editor). *Soil Contamination*, InTech., 67–90.

Sherma, J. (1972). Principles and techniques. In: G. Zweig and J. Sherma (Editors). *CRC Handbook of Chromatography: General Data and Principles*, 2, CRC Press Inc., Boca Raton, FL. pp. 1–101.

Sitelab Corporation (2010). *Polycyclic aromatic hydrocarbon application using Sitelab UVF-3100D*. Sitelab Corporation, West Newbury, MA, USA.

Snape, I., Harvey, P. M., Ferguson, S. H., Rayner, J. L., and Revill, A. (2005). Investigation of evaporation and biodegradation of fuel spills in Antarctica I. a chemical approach using GC–FID. *Chemosphere*, 61, 1485–1499.

Soares, A. A., Moldrup, P., Minh, L. N., Vendelboe, A. L., Schjonning, P., and de Jonge, L. W. (2013). Sorption of phenanthrene on agricultural soils. *Water, Air, and Soil Pollution*, 224, 1519–1531.

Soriano-Disla, J. M., Janik, L. J., Viscarra Rossel, R. A., MacDonald, L. M., and McLaughlin, M. J. (2014). The Performance of visible, near-, and mid-infrared reflectance spectroscopy for prediction of soil physical, chemical, and biological properties. *Applied Spectroscopy Reviews*, 49(2), 139–186.

SPDC (2006). *Environmental Impact Assessment (EIA) of Rumuekpe (OML 22) and Etelebou (OML 28) Area 3 Dimensional Seismic Survey*. Publication of the Shell Petroleum Development Company (SPDC) of Nigeria Limited, Abuja, Nigeria. Available at http://www.shell.com.ng/environment-society/environment-impact-assessments.html [accessed 25 February, 2013].

SPDC (2013). Oil spills in the Niger Delta – Monthly Data. Available at http://www.shell.com.ng/environment-society/environment-tpkg/oil-spills/monthly-data.html [accessed 13 November, 2013].

Stallard, B. R., Garcia, M. J., and Kaushik, S. (1996). Near-IR reflectance spectroscopy for the determination of motor oil contamination in sandy loam. *Applied Spectroscopy*, 50, 334–338.

State University of New York at Oswego. (2008). *Lecture Note for CHEMISTRY245L*. Available at http://www.oswego.edu/~kadima/CHE425/CHE425L/FLUORESCENCE_SPECTROSCOPY_08.pdf [accessed 31 July, 2012].

Stenberg, B. (2010). Effects of soil sample pre-treatments and standardized rewetting as interacted with sand classes on Vis-NIR predictions of clay and soil organic carbon. *Geoderma*, 158, 15–22.

Stenberg, B., Viscarra Rossel, R. A., Mouazen, A. M., and Wetterlind, J. (2010). Visible and near infrared spectroscopy in soil science. In: D. Sparks (Editor). *Advances in agronomy*, 107, 163–215.

Tanee, F. B. G., and Albert, E., (2011). Post-remediation assessment of crude oil polluted site at Kegbara-Dere Community, Gokana L.G.A. of Rivers State, Nigeria. *Bioremediation and Biodegradation*, 2, DOI: http://dx.doi.org/10.4172/2155-6199.1000122 [accessed 26 February, 2013].

Tekin, Y., Tumsavas, Z., and Mouazen, A. M. (2012). Effect of moisture content on prediction of organic carbon and pH using visible and near infrared spectroscopy. *Soil Science Society of America Journal*, 76(1), 188–198.

Umechuruba, C. I. (2005). Health impact assessment of mangrove vegetables in an oil spilled site at the Bodo West Field in Rivers State, Nigeria. *Journal of Applied Science and Environmental Management*, 9, 69–73.

UNEP (2011). *Environmental Assessment of Ogoniland*. United Nations Environment Programme (UNEP), Nairobi, Kenya. Available: http://www.unep.org (accessed October 29, 2012).

USDA (2005). Global soil region. US Department of Agriculture, Natural Resources Conservation Service, Soil Survey Division, Washington, D.C.

Vallejo, B., Izquierdo, A., Blasco, R., Perez del Campo, P., and Luque de Castro, M. D. (2001). Bioremediation of an area contaminated by a fuel spill. *Journal of Environmental Monitoring*, 3, 274–280.

Van-Dissel, J. P., and Omuka, P. S. (1994). Environmental impact of exploration and production operations on the Niger Delta mangrove. In: *Proceedings of the II international conference on health, safety and environment in oil and gas exploration and production*, Jakarta, Indonesia. pp. 374–445.

Vegter, J., Lowe, J., and Kasamas, H. (2002). Sustainable management of contaminated land: an overview. Austrian Federal Environment Agency on behalf of CLARINET.

Villalobos, M., Avila-Forcada, A. P., and Gutierrez-Ruiz, M. E. (2008). An improved gravimetric method to determine total petroleum hydrocarbons in contaminated soils. *Water, Air and Soil Pollution*, 194, 151–161.

Viscarra Rossel, R. A., and Behrens, T. (2010). Using data mining to model and interpret soil diffuse reflectance spectra. *Geoderma*, 158, 46–54.

Viscarra Rossel, R. A., and McBratney, A. B. (1998). Laboratory evaluation of a proximal sensing technique for simulta-neous measurement of soil clay and water. *Geoderma*, 85, 19–39.

Viscarra Rossel, R. A., McGlynn, R. N., and McBratney, A. B. (2006b). Determining the composition of mineral-organic mixes using UV–vis–NIR diffuse reflectance spectroscopy. *Geoderma*, 137, 70–82.

Viscarra Rossel, R. A., Walvoort, D. J. J., McBratney, A. B., Janik, L. J., and Skjemstad, J. O. (2006a). Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. *Geoderma*, 131, 59–75.

Wang, Z. D., Fingas, M., and Sigouin, L. (2002). Using multiple criteria for fingerprinting unknown oil samples having very similar chemical composition. In: *Proceedings of the 25th Arctic and Marine Oil Spill Program (AMOP) Technical Seminar: Environment Canada*, Ottawa, Canada. pp. 639–660.

Wang, Z. D., Fingas, M., Landriault, M., Sigouin, L., Grenon, S., and Zhang, D. (1999). Source identification of an unknown spilled oil from Quebec (1998) by unique biomarker and diagnostic ratios of source-specific marker compounds. *Environmental Technology*, 20, 851–862.

Wang, Z., and Fingas M. (1995). Differentiation of the sources of spilled oil and monitoring of the oil weathering process using gas chromatography – mass spectroscopy. *Journal of Chromatography A*, 712, 321–343.

Wang, Z., and Fingas, M. (2003). Development of oil hydrocarbon fingerprinting and identification techniques. *Marine Pollution Bulletin*, 47, 423–452.

Wei, M. Y., Wen, S. D., Yang, X. Q., Guo, L. H. (2009). Development of redox-labelled electrochemical immunoassay for polycyclic aromatic hydrocarbons with controlled surface modification and catalytic voltammetric detection. *Biosensors and Bioelectronics*, 24, 2909–2914.

Weisman, W. (1998). Analysis of petroleum hydrocarbons in environmental media. In: W. Weisman (Editor). *Total Petroleum Hydrocarbon Criteria Working Group (TPHCWG) Series, 1.* Amherst Scientific Publishers, Amherst, MA. pp. 1–98.

Whalley, W. R., and Stafford, J. V. (1992). Real-time sensing of soil water content from mobile machinery: options for sensor design. *Computer and Electronics in Agriculture*, 7, 269–358.

White, D. M., and Irvine, R. L. (1994). Analysis of bioremediation in organic soils. Paper jointly presented at the meeting of The Society of Chemical Industry and Royal Society of Chemistry, Sunbury on Thames, Middlesex, UK.

Whittaker, M., Pollard, S. J. T., and Fallick, T. E. (1995). Characterisation of refractory wastes at heavy oil-contaminated sites: A review of conventional and novel analytical methods. *Environmental Technology*, 16, 1009–1033.

Wikipedia, the Free Encyclopaedia. (2012). *Raman Spectroscopy*. Available at http://en.wikipedia.org/wiki/Raman_spectroscopy [accessed 31 July, 2012].

Willey, R. R. (1976). Fourier transform infrared spectrophotometer for transmittance and diffuse reflectance measurements. *Applied Spectroscopy*, 30(6), 593–601.

Williams, P. C., and Sobering, D. C. (1986). Attempts at standardization of hardness testing of wheat. II. The near-infrared reflectance method. *Cereal Foods World*, 31, 417–420.

Williams, P., and Norris, K. (1987). *Near infrared technology in the agriculture and food industries*. American Association of Cereals Chemists Inc., St. Paul, Minnesota, USA.

Workman Jr., J. (1999). Review of process and non-invasive near-infrared and infrared spectroscopy: 1993–1999. *Applied Spectroscopy Reviews*, 34, 1–89.

Workman Jr., J., Weyer, L. (2008). *Practical guide to interpretive near-infrared spectroscopy*. CRC Press, Taylor and Francis Group, Boca Raton, FL, USA.

WSDE (2007). Evaluating the Toxicity and Assessing the Carcinogenic Risk of Environmental Mixtures Using Toxicity Equivalency Factors. Washington State Department of Ecology (WSDE), Richland, WA, USA.

Yang, Z., Yang, C., Wang, Z., Hollebone, B., Landriault, M., and Brown, C. E. (2011). Oil fingerprinting analysis using commercial solid phase extraction (SPE) cartridge and gas chromatography-mass spectrometry (GC–MS). *Analytical Methods*, 3, 628–635.

Yunker, M. B., and MacDonald, R. W. (1995). Composition and origins of polycyclic aromatic hydrocarbons in the Mackenzie River and on the Beaufort Sea Shelf. *Journal of Arctic Institute of North America*, 48(2), 118–129.

Zhang, Y. F., Ma, Y., Gao, Z. X. and Dai, S. G. (2010). Predicting the cross-reactivities of polycyclic aromatic hydrocarbons in ELISA by regression and CoMFA methods. *Analytical and Bioanalytical Chemistry*, 397, 1551–2557.

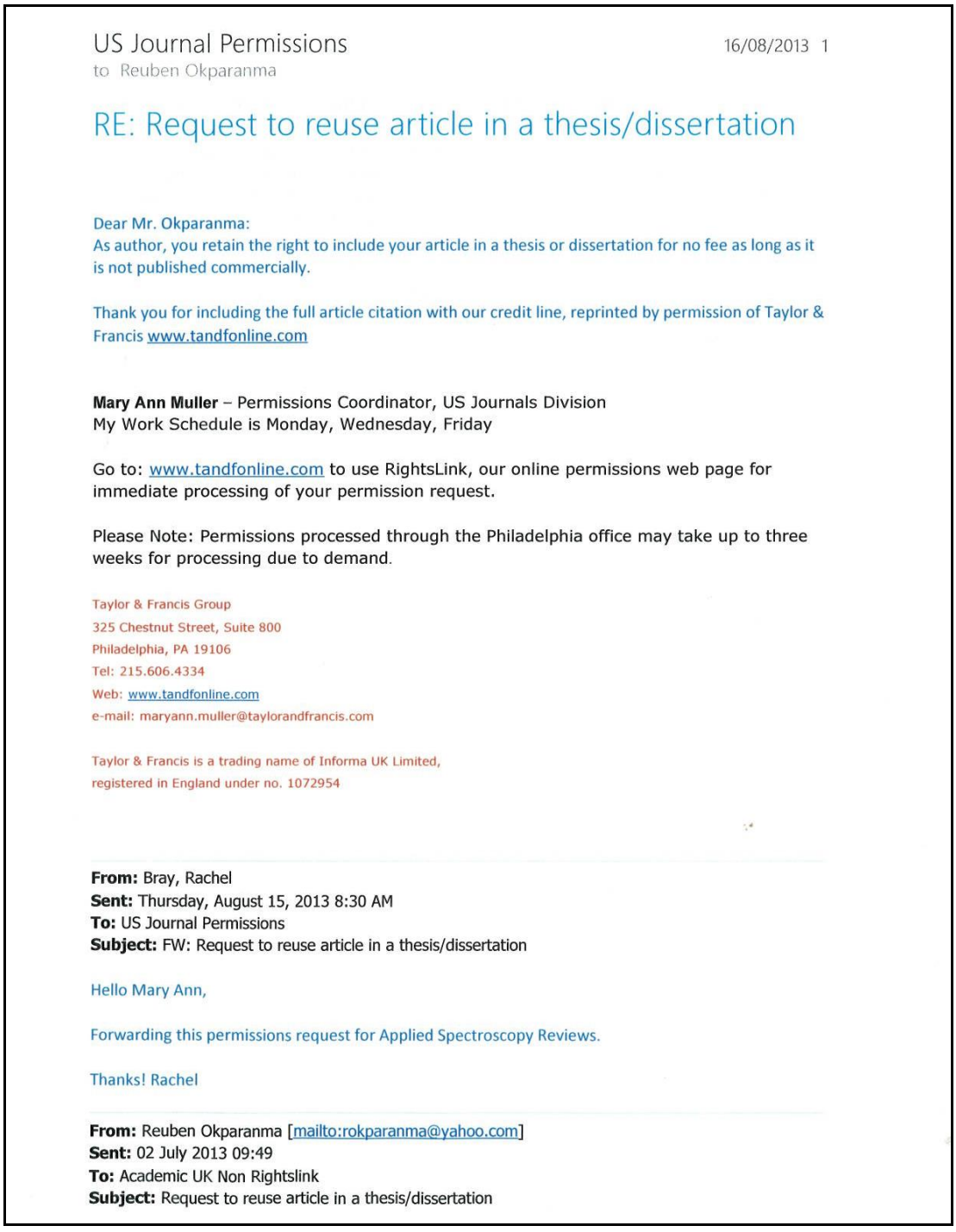
Zhang, Y., Zhu, Y. G., Houot, S., Qiao, M., Nunan, N., and Garnier, P. (2011). Remediation of polycyclic aromatic hydrocarbon (PAH) contaminated soil through composting with fresh organic wastes. *Environmental Science and Pollution Research*, DOI 10.1007/s11356-011-0521-5.

Zhou, C., Wang, Q. E., Gao, S. S., and Zhuang, H. S. (2009). Determination of naphthalene by competitive fluorescence immunoassay. *Environmental Monitoring and Assessment*, 154, 233–239.

Zwanziger, Z., and Foster, H. (1998). Near infrared spectroscopy of fuel contaminated sand and soil. I. Preliminary results and calibration study. *Journal of Near Infrared Spectroscopy*, 6, 189–197.

APPENDICES

Appendix A: Permission to reuse article in a thesis/dissertation – Taylor and Francis



US Journal Permissions 16/08/2013 1
to Reuben Okparanma

RE: Request to reuse article in a thesis/dissertation

Dear Mr. Okparanma:
As author, you retain the right to include your article in a thesis or dissertation for no fee as long as it is not published commercially.

Thank you for including the full article citation with our credit line, reprinted by permission of Taylor & Francis www.tandfonline.com

Mary Ann Muller – Permissions Coordinator, US Journals Division
My Work Schedule is Monday, Wednesday, Friday

Go to: www.tandfonline.com to use RightsLink, our online permissions web page for immediate processing of your permission request.

Please Note: Permissions processed through the Philadelphia office may take up to three weeks for processing due to demand.

Taylor & Francis Group
325 Chestnut Street, Suite 800
Philadelphia, PA 19106
Tel: 215.606.4334
Web: www.tandfonline.com
e-mail: maryann.muller@taylorandfrancis.com

Taylor & Francis is a trading name of Informa UK Limited,
registered in England under no. 1072954

From: Bray, Rachel
Sent: Thursday, August 15, 2013 8:30 AM
To: US Journal Permissions
Subject: FW: Request to reuse article in a thesis/dissertation

Hello Mary Ann,

Forwarding this permissions request for Applied Spectroscopy Reviews.

Thanks! Rachel

From: Reuben Okparanma [mailto:rokparanma@yahoo.com]
Sent: 02 July 2013 09:49
To: Academic UK Non Rightslink
Subject: Request to reuse article in a thesis/dissertation

Appendix B: Permission to reuse article in a thesis/dissertation – Springer

SPRINGER LICENSE
TERMS AND CONDITIONS

Jul 02, 2013

This is a License Agreement between Reuen Okparanma ("You") and Springer ("Springer") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Springer, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number	3180731498517
License date	Jul 02, 2013
Licensed content publisher	Springer
Licensed content publication	Water, Air, and Soil Pollution
Licensed content title	Combined Effects of Oil Concentration, Clay and Moisture Contents on Diffuse Reflectance Spectra of Diesel-Contaminated Soils
Licensed content author	Reuben N. Okparanma
Licensed content date	Jan 1, 2013
Volume number	224
Issue number	5
Type of Use	Thesis/Dissertation
Portion	Full text
Number of copies	1
Author of this Springer article	Yes and you are the sole author of the new work
Order reference number	
Title of your thesis / dissertation	RAPID MEASUREMENT OF HYDROCARBON CONTAMINATION IN SOILS BY VISIBLE AND NEAR-INFRARED SPECTROSCOPY
Expected completion date	Oct 2013
Estimated size(pages)	200
Total	0.00 GBP
Terms and Conditions	

Introduction
The publisher for this copyrighted material is Springer Science + Business Media. By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the Billing and Payment terms and conditions established by Copyright Clearance Center, Inc. ("CCC"), at the time that you opened your Rightslink account and that are available at any time at http://myaccount.copyright.com).

Limited License
With reference to your request to reprint in your thesis material on which Springer Science and Business Media control the copyright, permission is granted, free of charge, for the use indicated in your enquiry.

Appendix C: Permission to reuse article in a thesis/dissertation – Elsevier

ELSEVIER LICENSE
TERMS AND CONDITIONS

Sep 27, 2013

This is a License Agreement between Reuben Okparanma ("You") and Elsevier ("Elsevier") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Elsevier, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

Supplier	Elsevier Limited The Boulevard,Langford Lane Kidlington,Oxford,OX5 1GB,UK
Registered Company Number	1982084
Customer name	Reuben Okparanma
Customer address	16 Prince Philip Avenue, Bedford, Bedfordshire MK43 0SX
License number	3237161187385
License date	Sep 27, 2013
Licensed content publisher	Elsevier
Licensed content publication	Environmental Pollution
Licensed content title	Analysis of petroleum-contaminated soils by diffuse reflectance spectroscopy and sequential ultrasonic solvent extraction–gas chromatography
Licensed content author	Reuben N. Okparanma,Frederic Coulon,Abdul M. Mouazen
Licensed content date	January 2014
Licensed content volume number	184
Licensed content issue number	
Number of pages	8
Start Page	298
End Page	305
Type of Use	reuse in a thesis/dissertation
Intended publisher of new work	other
Portion	full article
Format	both print and electronic
Are you the author of this Elsevier article?	Yes
Will you be translating?	No
Order reference number	

Appendix D: Permission to use part of Table 4 (TPCWG 1998, vol. 2)

RE: Permission to use part of Table 4 (TPHCWG 1998, vol. 2)

FROM:Brenna Lockwood
TO:'Reuben okparanma'
Message flagged
Wednesday, 14 December 2011, 14:52

Message Body

Reuben,

You have our permission to use the table for your review paper.

Thank you,

Brenna Lockwood

AEHS Foundation, Inc.

150 Fearing St., Suite 21

Amherst, MA 01002

413-549-5170 T

413-549-0579 F

brenna@aehsfoundation.org

www.aehsfoundation.org

From: Reuben okparanma [mailto:rokparanma@yahoo.com]
Sent: Tuesday, December 13, 2011 2:21 PM
To: orders@aehs.com
Subject: Permission to use part of Table 4 (TPHCWG 1998, vol. 2)

Dear Editor,

I am emailing to seek permission to use part of Table 4 in pages 20 and 21 in one of your publications: TPHCWG 1998, vol. 2. for a review paper intended to be submitted to the Environmental Science and Technology Journal.

Best wishes.

Reuben.

Reuben N. Okparanma,
Researcher, School of Applied Sciences,
Cranfield University, MK43 0AL Bedfordshire, UK.
r.okparanma@cranfield.ac.uk
01234-750111 ext 2793

Appendix E: Chemical Analysis Results for Baraboo

Sample ID	Individual PAHs ($mgkg^{-1}$)												Total PAH ($mgkg^{-1}$)
	ACE	FLU	PHE	PYR	B[a]A	CHR	B[b]F	B(k)F	B(a)P	I[1,2,3-cd]	D(a,h)A	B(g,h,i)P	
BRB-01	0.73	0.15	0.00	0.59	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.46
BRB-02	1.46	0.58	0.29	0.15	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.48
BRB-06	0.29	0.44	0.15	0.15	0.15	0.15	0.00	0.00	0.00	0.00	0.00	0.00	1.46
BRB-07	0.58	1.02	0.15	0.15	0.15	0.15	0.00	0.00	0.15	0.00	0.00	0.00	2.48
BRB-08	0.87	0.00	0.15	0.15	0.44	0.44	0.00	0.00	0.15	0.00	0.00	0.00	2.19
BRB-09	0.44	0.44	0.00	0.15	0.15	0.15	0.00	0.00	0.00	0.00	0.00	0.00	1.31
BRB-10	0.87	1.16	0.00	0.00	0.15	0.15	0.00	0.00	0.44	0.00	0.00	0.00	2.77
BRB-11	1.16	0.29	0.00	0.29	0.00	0.00	0.00	0.00	0.15	0.00	0.00	0.00	1.90
BRB-12	1.75	0.15	0.15	0.59	0.00	0.00	0.00	0.00	0.15	0.00	0.00	0.00	2.77
BRB-14	0.29	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.29
BRB-16	0.15	0.15	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.29
BRB-18	0.29	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.29
BRB-21	0.07	0.15	0.22	0.95	0.37	0.37	0.00	0.00	0.44	0.22	0.00	0.22	3.45
BRB-22	0.15	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.15
BRB-23	0.29	0.15	0.00	0.15	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.58
BRB-24	0.29	0.00	0.15	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.58
BRB-25	0.29	0.29	0.15	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.73
BRB-26	1.60	0.29	0.15	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.18
BRB-31	0.58	0.15	0.00	0.73	0.00	0.29	0.15	0.15	0.29	0.00	0.00	0.00	2.49
BRB-32	0.73	0.00	0.15	0.29	0.15	0.15	0.00	0.00	0.15	0.00	0.00	0.00	1.61
BRB-33	1.02	0.73	0.15	0.15	0.00	0.00	0.00	0.00	0.29	0.00	0.00	0.00	2.33
BRB-34	0.44	0.44	0.15	0.73	0.73	0.73	0.00	0.00	0.59	0.15	0.00	0.00	4.10
BRB-35	0.44	0.29	0.15	0.29	0.44	0.44	0.15	0.15	0.29	0.00	0.00	0.15	2.93
BRB-36	3.06	0.58	0.73	0.59	0.29	0.29	0.15	0.15	0.44	0.00	0.00	0.00	6.57
BRB-37	0.58	0.00	0.00	0.88	0.15	0.00	0.00	0.00	0.29	0.00	0.00	0.15	2.05
BRB-38	1.31	0.29	0.58	0.29	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	3.21
BRB-39	1.16	1.60	0.15	1.03	0.88	0.88	0.00	0.00	0.29	0.15	0.15	0.15	6.58
BRB-40	0.58	0.29	0.15	0.29	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.46
BRB-41	0.87	0.00	0.15	0.15	0.29	0.29	0.15	0.15	0.29	0.00	0.00	0.00	2.49
BRB-43	1.16	0.29	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.60
BRB-44	0.29	0.00	0.15	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.58
BRB-45	0.15	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.29
BRB-46	0.58	1.31	0.58	0.73	0.44	0.44	0.00	0.00	0.00	0.00	0.00	0.00	4.67
BRB-47	2.33	2.18	2.48	0.59	0.15	0.15	0.00	0.00	0.44	0.15	0.00	0.00	10.20
BRB-48	0.58	0.87	0.87	0.29	0.15	0.15	0.00	0.00	0.00	0.00	0.00	0.00	3.64
BRB-49	0.58	0.87	0.87	0.15	0.29	0.15	0.00	0.00	0.15	0.00	0.00	0.00	3.94
BRB-50	0.73	1.60	1.16	0.73	0.15	0.15	0.15	0.15	0.00	0.00	0.00	0.00	6.42
BRB-51	0.73	1.16	0.58	0.29	0.00	0.00	0.15	0.15	0.29	0.00	0.00	0.00	4.23
BRB-53	1.16	2.04	1.75	0.44	0.29	0.29	0.00	0.00	0.29	0.00	0.00	0.00	7.29
BRB-54	1.02	4.80	3.64	0.00	0.88	1.17	0.15	0.15	0.59	0.00	0.00	0.00	16.62
BRB-57	0.87	1.46	1.31	0.00	0.29	0.29	0.15	0.15	0.73	0.00	0.00	0.00	5.25
BRB-58	1.46	2.33	1.16	0.59	0.15	0.15	0.00	0.00	0.29	0.00	0.00	0.00	7.29
BRB-60	1.89	2.77	1.60	0.44	0.59	0.44	0.15	0.15	0.29	0.00	0.00	0.15	9.92
Mean	0.83	0.73	0.47	0.30	0.18	0.18	0.03	0.03	0.17	0.02	0.00	0.02	3.38
Min.	0.07	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.15
Max.	3.06	4.80	3.64	1.03	0.88	1.17	0.15	0.15	0.73	0.22	0.15	0.22	16.62
Range	0.76	0.73	0.47	0.30	0.18	0.18	0.03	0.03	0.17	0.02	0.00	0.02	3.23

Appendix F: Chemical Analysis Results for Bomu 1

Sample ID	Individual PAHs ($mgkg^{-1}$)													Total PAH ($mgkg^{-1}$)
	ACE	FLU	PHE	ANT	PYR	B[a]A	CHR	B[b]F	B(k)F	B(a)P	I[1,2,3-cd]P	D(a,h)A	B(g,h,i)P	
BM1-01	0.18	0.20	0.18	0.18	0.44	0.00	0.00	0.00	0.00	0.25	0.00	0.00	0.00	1.43
BM1-02	0.22	0.51	0.73	0.71	1.11	0.38	0.39	1.33	1.23	0.25	0.56	0.37	0.56	8.36
BM1-03	0.30	0.07	0.92	0.50	0.07	0.09	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.95
BM1-04	2.40	2.70	3.15	0.18	0.22	0.00	0.00	0.00	0.00	1.80	0.00	0.00	0.00	10.45
BM1-05	0.06	0.06	0.12	0.12	0.29	0.22	0.22	0.08	0.08	0.08	0.38	0.14	0.38	2.22
BM1-06	0.18	0.20	0.12	0.12	0.87	0.80	0.08	0.11	0.11	0.05	0.07	0.46	0.07	3.24
BM1-07	0.12	0.37	0.06	0.06	0.58	0.29	0.29	0.22	0.22	0.09	0.49	0.38	0.49	3.67
BM1-08	0.06	0.12	0.12	0.12	0.07	0.00	0.00	0.46	0.46	0.05	0.00	0.00	0.00	1.47
BM1-09	0.31	0.55	0.92	0.92	0.51	0.22	0.22	0.14	0.14	0.28	0.36	0.23	0.36	5.14
BM1-10	0.31	0.18	0.18	0.18	0.15	0.07	0.07	0.07	0.07	0.09	0.00	0.00	0.00	1.38
BM1-12	0.37	1.10	0.18	0.18	1.74	0.87	0.87	0.67	0.67	0.26	1.48	1.13	1.48	11.02
BM1-13	0.18	0.55	0.80	0.80	0.73	0.36	0.36	0.12	0.12	0.10	0.13	0.49	0.13	4.89
BM1-14	1.00	0.50	0.06	0.06	0.07	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.70
BM1-15	0.18	0.18	0.18	0.18	2.83	1.96	0.20	0.59	0.59	0.00	0.65	1.15	0.65	9.36
BM1-16	0.06	0.06	0.12	0.12	0.15	0.00	0.00	0.08	0.08	0.06	0.23	0.15	0.23	1.34
BM1-17	0.11	0.25	0.37	0.36	0.56	0.19	0.19	0.66	0.61	0.13	0.28	0.19	0.28	4.18
BM1-18	0.12	0.12	0.12	0.12	0.58	0.29	0.29	0.06	0.06	0.00	0.31	0.33	0.31	2.73
BM1-19	1.00	0.40	0.06	0.06	0.07	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.60
BM1-21	1.00	0.50	0.12	0.12	0.36	0.07	0.07	0.12	0.12	0.08	0.00	0.00	0.00	2.58
BM1-22	0.18	0.06	0.12	0.12	0.87	0.07	0.07	0.19	0.19	0.32	0.31	0.24	0.31	3.07
BM1-23	0.12	0.12	0.12	0.12	0.29	0.29	0.29	0.40	0.06	0.08	0.20	0.12	0.20	2.43
BM1-24	0.12	0.12	0.24	0.24	0.29	0.15	0.15	0.14	0.14	0.07	0.33	0.00	0.33	2.30
BM1-25	0.43	0.06	0.06	0.06	0.44	0.07	0.07	0.06	0.06	0.12	0.44	0.18	0.44	2.48
BM1-26	0.60	0.80	1.50	0.06	0.98	0.56	0.40	0.25	0.09	0.00	0.00	0.00	0.00	5.24
BM1-27	0.02	0.02	0.02	0.09	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.15
BM1-28	0.31	0.06	0.12	0.12	0.73	0.65	0.07	0.20	0.20	0.11	0.06	0.20	0.06	2.88
BM1-29	0.06	0.15	1.00	0.00	0.29	0.22	0.22	0.08	0.08	0.08	0.00	0.24	0.00	2.41
BM1-30	0.12	0.80	0.31	0.31	0.29	0.15	0.15	0.00	0.00	0.07	0.00	0.00	0.00	2.18
BM1-31	0.90	0.80	1.80	0.06	0.70	0.60	0.50	0.00	0.00	0.07	0.00	0.00	0.00	5.43
BM1-32	0.06	0.12	0.06	0.06	0.29	0.29	0.29	0.06	0.06	0.10	0.24	0.23	0.24	2.11
BM1-33	0.02	0.30	1.06	0.06	0.15	0.07	0.07	0.00	0.00	0.00	0.00	0.00	0.00	1.73
BM1-34	0.06	0.03	1.06	0.06	0.36	0.00	0.00	0.18	0.18	0.21	0.38	0.15	0.00	2.67
BM1-35	0.06	0.02	1.06	0.06	0.36	0.36	0.36	0.00	0.00	0.06	0.00	0.18	0.00	2.53
BM1-36	0.37	0.43	1.06	0.06	0.80	0.87	0.09	0.10	0.10	0.08	0.08	0.00	0.08	4.11
BM1-37	0.06	0.42	1.31	0.31	0.65	0.36	0.36	0.07	0.07	0.06	0.08	0.30	0.08	4.13
BM1-38	0.12	0.06	1.06	0.06	0.65	0.07	0.07	0.14	0.14	0.28	0.33	0.17	0.33	3.48
BM1-39	0.06	0.06	1.06	0.06	0.94	0.65	0.07	0.20	0.20	0.00	0.22	0.38	0.22	4.12
BM1-40	0.12	1.06	1.12	0.12	0.51	0.15	0.15	0.12	0.12	0.17	0.26	0.25	0.26	4.41
BM1-41	0.24	0.12	1.18	0.18	0.51	0.94	0.09	0.11	0.11	0.05	0.00	0.27	0.00	3.83
BM1-42	0.31	0.12	1.18	0.18	0.22	0.15	0.15	0.06	0.06	0.11	0.20	0.15	0.20	3.09
BM1-43	0.12	0.12	1.18	0.18	0.29	0.22	0.22	0.15	0.15	0.09	0.05	0.34	0.05	3.17
BM1-44	0.06	0.00	1.06	0.06	0.36	0.22	0.22	0.30	0.30	0.23	0.09	0.24	0.09	3.23
BM1-45	0.37	0.12	1.24	0.24	0.36	0.15	0.15	0.08	0.08	0.06	0.38	0.16	0.44	3.82
BM1-46	0.71	0.09	1.00	0.00	0.07	0.00	0.00	0.45	0.45	0.07	0.00	0.00	0.00	2.84
BM1-47	1.10	1.29	1.18	0.18	2.40	2.61	0.26	0.30	0.30	0.24	0.24	0.00	0.24	10.34
BM1-48	0.55	0.31	1.10	1.10	0.51	0.29	0.29	0.24	0.11	0.31	0.14	0.23	0.14	5.32
BM1-49	0.61	0.49	1.98	0.98	0.58	0.51	0.05	0.08	0.08	0.15	0.33	0.00	0.33	6.16
BM1-50	0.98	0.55	1.29	1.29	0.44	0.29	0.29	0.25	0.12	0.11	0.31	0.12	0.31	6.34
BM1-51	1.31	1.18	1.30	1.20	1.22	1.10	0.98	1.08	1.08	0.13	0.41	0.00	0.41	11.40
BM1-52	1.96	1.12	2.57	2.57	0.87	0.58	0.58	0.51	0.24	0.23	0.62	0.23	0.62	12.70
BM1-53	0.18	0.12	0.06	0.06	0.07	0.29	0.29	0.11	0.11	0.15	0.27	0.12	0.27	2.09
BM1-54	1.24	1.67	2.30	1.80	0.51	0.51	0.05	0.09	0.09	0.11	0.25	0.06	0.25	8.93
BM1-55	0.18	0.31	1.06	0.06	0.36	0.15	0.15	0.15	0.15	0.05	0.45	0.13	0.45	3.65
BM1-56	0.55	0.55	1.10	1.10	0.87	0.87	0.87	0.73	0.73	0.16	1.29	1.15	1.29	11.26
BM1-57	0.37	0.06	0.18	0.18	0.44	0.15	0.15	0.06	0.06	0.11	0.00	0.23	0.00	1.97
BM1-58	1.06	1.20	1.06	1.06	0.29	0.22	0.22	0.49	0.49	0.16	0.19	1.82	0.19	8.45
BM1-59	1.50	1.50	1.06	1.06	1.20	1.10	0.00	0.07	0.07	0.07	0.00	0.00	0.00	7.64
BM1-60	1.06	1.00	1.12	0.12	0.15	1.07	1.07	1.00	0.00	0.07	0.00	0.00	0.00	6.67
Mean	0.46	0.45	0.81	0.36	0.56	0.39	0.22	0.23	0.19	0.14	0.23	0.22	0.22	4.47
Min.	0.02	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.15
Max.	2.40	2.70	3.15	2.57	2.83	2.61	1.07	1.33	1.23	1.80	1.48	1.82	1.48	12.70
Range	0.44	0.45	0.79	0.36	0.56	0.39	0.22	0.23	0.19	0.14	0.23	0.22	0.22	4.32

Appendix G: Chemical Analysis Results for Bomu 2

Sample ID	Individual PAHS ($mgkg^{-1}$)													Total PAH ($mgkg^{-1}$)
	ACE	FLU	PHE	ANT	PYR	B[a]A	CHR	B[b]F	B(k)F	B(a)P	I[1,2,3-cd]P	D(a,h)A	B(g,h,i)P	
BM2-01	0.60	0.78	0.87	0.30	0.14	0.07	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.78
BM2-02	0.91	0.30	0.76	0.76	1.44	0.29	0.24	0.16	0.16	0.30	0.00	0.00	0.00	5.34
BM2-03	3.51	0.30	0.30	0.30	0.58	1.01	0.38	0.18	0.18	0.12	0.78	0.00	0.33	7.98
BM2-04	0.30	0.00	0.15	0.15	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.61
BM2-05	0.30	0.00	0.30	0.30	0.29	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.20
BM2-06	0.76	3.66	1.68	1.68	0.86	0.14	0.55	0.11	0.11	0.52	0.82	0.00	0.32	11.21
BM2-07	2.29	2.90	0.76	0.76	0.14	0.00	0.00	0.00	0.00	8.35	0.00	0.00	0.00	15.20
BM2-08	0.00	1.22	0.61	0.61	0.29	0.14	0.00	0.15	0.15	0.94	0.00	0.00	0.00	4.10
BM2-09	0.61	0.00	2.44	2.44	0.86	0.14	0.00	0.00	0.00	0.00	0.00	0.00	0.00	6.49
BM2-10	0.61	0.61	1.68	1.68	0.43	0.29	0.00	0.00	0.00	0.00	0.00	0.00	0.00	5.29
BM2-11	4.57	0.46	0.30	0.30	0.86	0.14	0.00	0.14	0.14	0.14	0.00	0.00	0.00	7.06
BM2-12	1.37	0.00	0.46	0.46	0.43	0.14	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.86
BM2-13	0.00	0.00	0.30	0.30	0.29	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.90
BM2-18	1.22	6.25	0.61	0.61	0.58	0.43	0.00	0.13	0.13	0.40	0.00	0.00	0.00	10.35
BM2-19	0.91	0.91	1.22	1.22	0.86	0.43	0.29	0.00	0.00	0.31	0.00	0.00	0.00	6.16
BM2-20	2.29	0.46	0.30	0.30	1.73	0.14	0.53	0.16	0.16	0.11	0.69	0.00	0.00	6.89
BM2-21	1.07	1.68	0.00	0.00	0.58	0.14	0.00	0.00	0.00	0.13	0.00	0.00	0.00	3.59
BM2-22	3.05	1.37	0.15	0.15	0.72	0.00	0.46	0.11	0.11	0.69	0.00	0.00	0.00	6.81
BM2-23	1.37	1.83	0.91	0.91	0.86	0.43	0.40	0.10	0.10	0.00	0.00	0.00	0.00	6.93
BM2-37	1.83	0.00	0.15	0.15	0.14	0.00	0.00	0.00	0.00	9.36	0.00	0.00	0.00	11.64
BM2-40	2.29	0.00	0.15	0.15	0.29	0.14	0.00	0.00	0.00	0.15	0.00	0.00	0.00	3.17
BM2-41	1.37	1.98	0.61	0.61	0.29	0.00	0.28	0.00	0.00	0.21	0.00	0.00	0.00	5.36
BM2-42	8.53	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8.53
BM2-44	2.44	3.66	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.12	0.00	0.00	0.00	6.21
BM2-45	2.44	0.00	0.30	0.30	0.14	0.29	0.00	0.00	0.00	0.88	0.00	0.00	0.00	4.36
BM2-46	0.00	0.00	0.46	0.46	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.91
BM2-49	0.46	0.46	1.07	1.07	0.72	0.14	0.00	0.00	0.00	0.13	0.00	0.00	0.00	4.04
BM2-50	0.76	0.61	0.30	0.30	0.58	0.29	0.00	0.00	0.00	0.11	0.00	0.00	0.00	2.96
BM2-51	1.07	0.00	0.30	0.30	0.86	0.14	0.00	0.00	0.00	0.81	0.00	0.00	0.00	3.49
BM2-52	0.46	0.00	0.46	0.46	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.37
BM2-53	0.46	0.00	0.15	0.15	0.14	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.91
BM2-54	0.91	0.00	1.83	1.83	0.72	0.14	0.00	0.00	0.00	0.14	0.00	0.00	0.00	5.58
BM2-55	0.30	0.76	6.86	6.86	1.44	0.00	0.23	0.15	0.15	0.32	0.00	0.00	0.00	17.09
BM2-56	0.46	0.91	0.46	0.46	0.72	0.14	0.00	0.00	0.00	0.11	0.00	0.00	0.00	3.26
BM2-57	0.30	0.00	0.15	0.15	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.61
BM2-60	0.91	0.61	0.30	0.30	0.58	0.14	0.00	0.00	0.00	0.10	0.00	0.00	0.00	2.96
Mean	1.41	0.88	0.76	0.75	0.52	0.15	0.09	0.04	0.04	0.68	0.06	0.00	0.02	5.39
Min.	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.61
Max.	8.53	6.25	6.86	6.86	1.73	1.01	0.55	0.18	0.18	9.36	0.82	0.00	0.33	17.09
Range	1.41	0.88	0.76	0.75	0.52	0.15	0.09	0.04	0.04	0.68	0.06	0.00	0.02	4.78