



ELSEVIER

Contents lists available at ScienceDirect

Talanta

journal homepage: www.elsevier.com/locate/talanta

Rapid prediction of total petroleum hydrocarbons in soil using a hand-held mid-infrared field instrument



Grant T. Webster^{a,*}, José M. Soriano-Disla^{b,c}, Joel Kirk^a, Leslie J. Janik^{b,c}, Sean T. Forrester^b, Mike J. McLaughlin^{b,c}, Richard J. Stewart^a

^a Ziltek Pty Ltd., PO Box 373, Kent Town, SA, 5071 Australia

^b CSIRO Land and Water, Environmental Contaminant Mitigation and Technologies, Waite Campus, Waite Road, Urrbrae, 5064 Australia

^c University of Adelaide, PMB 2, Glen Osmond, South Australia, 5064 Australia

ARTICLE INFO

Article history:

Received 1 June 2016

Received in revised form

21 July 2016

Accepted 22 July 2016

Available online 25 July 2016

Keywords:

DRIFT, diffuse reflectance infrared Fourier transform

Hand-held MIR

TPH, total petroleum hydrocarbons

ABSTRACT

This manuscript reports on the performance of a hand-held diffuse reflectance (mid)-infrared Fourier transform (DRIFT) spectrometer for the prediction of total petroleum hydrocarbons (TPH) in three different diesel-contaminated soils. These soils include: a carbonate dominated clay, a kaolinite dominated clay and a loam from Padova Italy, north Western Australia and southern Nigeria, respectively. Soils were analysed for TPH concentration using a standard laboratory methods and scanned in DRIFT mode with the hand-held spectrometer to determine TPH calibration models. Successful partial least square regression (PLSR) predictions, with coefficient of determination (R^2) ~ 0.99 and root mean square error (RMSE) < 200 mg/kg, were obtained for the low range TPH concentrations of 0 to $\sim 3,000$ mg/kg. These predictions were carried out using a set of independent samples for each soil type. Prediction models were also tested for the full concentration range (0–60,000 mg/kg) for each soil type model with R^2 and RMSE values of ~ 0.99 and $< 1,255$ mg/kg, respectively. Furthermore, a number of intermediate concentration range models were also generated for each soil type with similar R^2 values of ~ 0.99 and RMSE values < 800 mg/kg. This study shows the capability of using a portable mid-infrared (MIR) DRIFT spectrometer for predicting TPH in a variety of soil types and the potential for being a rapid in-field screening method for TPH concentration levels at common regulatory thresholds.

A novel hand-held mid-infrared instrument can accurately detect TPH across different soil types and concentrations, which paves the way for a variety of applications in the field.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Total petroleum hydrocarbons (TPH) comprise a mixture of short and long-chain aliphatic hydrocarbons and aromatic compounds (mixture of alkyl and mono/polycyclic) derived from crude oil. However, the aromatics group is a minor component in crude oil. The environmental occurrence of aliphatic short-chain fraction and mid-range fraction (particularly C_{10} – C_{36} compounds) in soil, sediment and water can adversely affect biological and human health [1,2]. The main sources of pollution by TPH are spills from the production, storage and distribution of petroleum products [1]. Assessment of TPH contamination in soil requires rapid, accurate, on-site and cost-effective methods to facilitate risk assessment processes and to assist with pollutant management and

remediation. However, the traditional laboratory method for the determination of TPH concentration using the capillary GC/FID method is time consuming and expensive. There are some portable methods available for the analysis of TPH [3], but they are still time consuming, and most require the extraction of the contaminant using chemical reagents.

A recent study by Forrester et al. [4] demonstrated that diffuse reflectance (mid)-infrared Fourier transform (DRIFT) spectroscopy using benchtop instrumentation, coupled with multivariate modelling such as partial least squares regression (PLSR) methods, could provide accurate and much more rapid predictions of TPH concentrations. In the article by Forrester et al. [4], the authors used the mid-infrared (MIR) region 3000 – 2600 cm^{-1} , which was identified to be sensitive to TPH concentration and capable of minimising interferences from natural soil organic matter and carbonates. While that study showed accurate and viable models for the prediction of TPH using a benchtop instrument, prediction of TPH using a portable hand-held instrument would be even

* Corresponding author at: Ziltek Pty Ltd., PO Box 373, Kent Town, SA 5071, Australia.

E-mail address: Grant.Webster@ziltek.com (G.T. Webster).

more advantageous. However, there may be concerns about the performance of such instruments (e.g. noise level) relative to laboratory-based benchtop instrumentation, considered to be a crucial factor for the development of reliable and precise models [5].

As discussed in a recent review [6], only a small percentage of the publications on the use of visible and infrared (IR) spectroscopy in soil analysis refer to hand-held instruments. Most studies reporting regression models using hand-held instruments for soil analysis were based on the visible-near infrared (vis-NIR), with MIR based devices rarely used [7,8]. This can be partially attributed to the historically poor portability of MIR instruments in the field [9]. This issue has now been solved with the emergence of true hand-held DRIFT MIR spectrometers suitable for in-field analysis.

Successful models have been reported using portable/hand-held vis-NIR to predict increasing concentrations of oil spiked into standard soils [10,11]. Conversely, worse results are obtained when using field samples [12,13], in agreement with Okparanma and Mouazen [14] with regard to MIR outperforming vis-NIR for the determination of TPH. Hitherto, there has only been one attempt to develop models for the prediction of TPH in soils using an MIR-DRIFT hand-held spectrometer [7]. However, those samples were sourced from a single unique site and comprised a relatively low TPH concentration range (0–11,000 mg/kg). In that study, an independent set of samples was not collected for validation and did not report on the influence of different soil types and TPH concentration ranges on calibration development. As previously determined by Forrester et al. [4], soil type and different TPH concentration ranges have been shown to influence the performance of MIR calibrations.

The present study reports on the performance of a portable hand-held DRIFT spectrometer for the prediction of TPH in a range of contaminated soils. This would allow for the practical/routine use of the MIR technique for accurately predicting TPH concentrations in the field. Thus, the practical use of this device in the field is discussed, as well as the impact of different soil types and TPH concentrations on the calibration models.

2. Materials and methods

2.1. Total petroleum hydrocarbon contaminated soils

Three sets of contrasting soils, contaminated by diesel spills or spiked in the laboratory, were used in this study. These soils were chosen to address site contamination issues and because they represented a good spread of TPH with which to build calibration models. The three soil sample sets can be described as follows: Set 1 comprised 67 uncontaminated samples collected from a site in Padova (Italy) and were individually spiked in the laboratory with diesel covering a range of concentrations (0–60,000 mg/kg). These samples were individually spiked at nominal concentrations covering an even spread throughout the concentration range. Set 2 comprised 67 samples sourced from the Pilbara in Western Australia at an actual contaminated site following a major diesel spill. These latter soils had been stocked into piles and samples collected to cover a range of concentrations (0–40,000 mg/kg). Set 3 comprised 60 samples sourced from an undisclosed (due to confidentiality restrictions) location in Nigeria. Similar to set 1, these samples (uncontaminated) were collected at the site and were spiked individually with diesel covering a range of concentrations (0–30,000 mg/kg). These samples were equally spiked at high (~30,000 mg/kg), medium (~10,000 mg/kg) and low (~1,000 mg/kg) TPH concentrations. The samples were air dried (24 h) and sieved to <2 mm, then split into two even duplicate sets using a riffle apparatus for infrared scanning and TPH

laboratory analysis. All samples were stored in a fridge at 4 °C to limit biodegradation of the TPH and avoid change of the TPH profile.

2.2. DRIFT-MIR spectra collection

The IR spectra of the soil samples used in this study were scanned in DRIFT mode using a hand-held 4100 ExoScan FTIR spectrometer (Agilent Technologies, USA) in the frequency range 6000–650 cm⁻¹, with a resolution of 8 cm⁻¹ and a 15 s scan time. The 4100 ExoScan instrument is now re-badged as RemScan™ and marketed by Ziltek Pty. Ltd., Australia for the quantitative analysis of TPH pollution in soil. Before acquiring a spectrum with the 4100 ExoScan FTIR spectrometer, the soil samples were mixed, tamped flat and the instrument placed square to the surface and in direct contact with the soil. Spectra were scanned in five replicates, with samples thoroughly mixed between each scan to capture the extent of heterogeneity of the soil. A silicon carbide reference disc (Perkin Elmer Inc., USA) was used as a background (assumed to have a reflectance $R_0=1$) and the spectra converted from reflectance (R_s) into pseudo absorbance (A) units (where $A=\text{Log}_{10}R_0 \cdot R_s^{-1}$). Spectra were further converted from the 4100 ExoScan (.ASP) format into GRAMS (.SPC) format using the GRAMS-AI converter software (GRAMS, Thermo Fisher Scientific, USA). To account for baseline variations, the spectra were pre-processed with the standard GRAMS “Auto-baseline” pre-processing application by fitting a first order polynomial function to each individual spectrum and then removing this to apply the baseline correction. The spectral range was restricted to 5000–750 cm⁻¹ in order to avoid excessive spectral noise at both ends of the range.

2.3. Laboratory determination of total petroleum hydrocarbons

Field and spiked samples were analysed for TPH C₁₀–C₄₀ by an accredited analytical reference laboratory (Australian Laboratory Services Pty Ltd) following the USEPA-SW 846-8015 A method. Sample extracts were analysed by Capillary GC/FID and quantified against alkane standards over the range C₁₀–C₄₀.

2.4. Multivariate models

Principal components analysis (PCA) using the spectral data only were developed by leave-one-out cross validation for each site. Briefly, PCA is an unsupervised data reduction method with which the spectral information is reduced to a few orthogonal latent variables (Principal Components, PC's). The first component (PC 1) explains most of the variability. The values of each sample in the new PC's are called scores, calculated through the loadings. The PCA loadings and score distributions were used to discern the spectral features contributing to the spectral variance and spectral variability [15].

Partial least squares regression (PLSR), previously described by Wold [16], is a multivariate method similar to PCA but which includes the correlations between the spectral information with the corresponding analytical data (TPH in this study). The PLSR models produce calibrations that are used for the prediction of the analytical values in unknown samples. Previously, predictive models were developed by using the MIR region 3000–2600 cm⁻¹ as discussed in Forrester et al. [4]. In this manuscript, both the 3000–2600 cm⁻¹ and 4500–4200 cm⁻¹ ranges were included in the MIR and NIR regions, respectively. The NIR region was used, in combination with the MIR range, to more accurately model high TPH concentrations, rather than with only the 3000–2600 cm⁻¹ region. The PLSR models were developed in GRAMS and Unscrambler v10.3 (Camo, Norway). Models were trained by leave-one-out cross validation using the calculation of the root mean square

Table 1
Comparison of FTIR^a TPH cross-validation models for sites 1, 2 and 3.

Site	Sub-model	TPH min (mg/kg)	TPH max (mg/kg)	Model cutoff (mg/kg)	Samples	Spectra	Factors	RMSECV (mg/kg)	R ²
1	1	0	60,000	20,000	50	250	7	899	0.997
	2	0	20,000	10,000	43	215	7	510	0.993
	3	0	10,000	5,000	37	185	7	448	0.98
	4	0	5,000	3,000	29	145	8	192	0.98
	5	0	3,000	N/A	29	145	4	123	0.990
2	1	0	40,000	20,000	48	265	6	1255	0.990
	2	0	20,000	5,000	43	240	4	739	0.97
	3	0	10,000	6,000	41	225	3	491	0.97
	4	0	6,000	3,000	39	210	3	283	0.96
	5	0	3,000	N/A	37	190	4	137	0.98
3	1	0	30,000	10,000	43	215	4	930	0.993
	2	0	10,000	5,000	34	170	3	306	0.996
	3	0	5,000	N/A	25	125	6	190	0.98

^a Individual FTIR measurements taken with 4100 ExoScan spectrometer.

error of cross validation (RMSECV) and the coefficient of determination (R²) based on the average of these replicate measurements. Models using both the full and restricted concentration ranges are compared in Table 1. Each model has a concentration cut-off (lower limit), which defines the concentration range of the model. The lower limits of each model were selected based on the relative standard deviation (RSD) of each respective model not exceeding 15% and is typically < 10%. Additionally, samples cover the maximum concentration of each concentration range model.

As a further, more stringent, test of prediction potential, independent validation was carried out for each site by splitting the initial set into two sets; one for derivation of a calibration model, trained by leave-one-out cross validation, and the remaining samples for an independent validation used to test the prediction of the TPH values. The independent validation samples for each set was selected to ensure an even spread of concentrations similar to their respective calibration set. For set 1, 50 samples were used for the calibration and 17 for the validation test. Set 2 comprised 48 samples for calibration and 19 used for validation. In set 3, 43 samples were used for calibration and 17 for validation.

The performance of each calibration cross-validation was reported in terms of RMSECV and R² for each site. In addition to the statistics of the calibration models, we also show the root mean square error of the prediction (RMSEP) and residual predictive deviation (RPD) for the validation samples. The RPD is used for quality assessment of PLSR for prediction of independent validation samples and provides an assessment of the performance across different data sets. Typically, models with RPD values higher than 3 are considered to provide analytical quality predictions [4].

3. Results and discussion

3.1. Spectra of soils

Soil samples, collected from the three different site locations, were characterised by the PCA loadings as well as the mean spectra and standard deviations (Figs. 1–3). Each score on a given scores plot represents a spectrum in multi-dimensional space. The individual variances calculated for each spectrum, based on the variability in the molecular vibrations detected in each sample scanned, contribute to the position of each score (spectrum) on a scores plot. The loadings plots were then used, in combination with the scores plot, to identify the variables (loadings at different wavenumber values) that contribute to the variability in the data set. By inspecting the relative contribution of the loadings to the scores, we can then shed light on the soil chemistry mechanisms

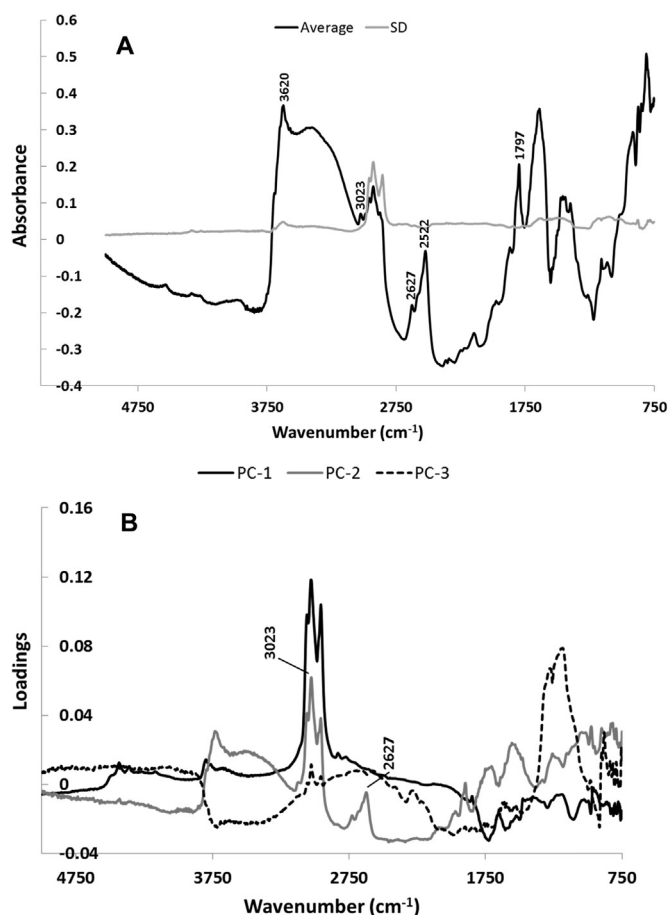


Fig. 1. (a) Auto-baseline corrected average and standard deviation (SD) spectra calculated from DRIFT spectra taken with 4100 ExoScan spectrometer of samples from site 1; (b) PCA loadings for principal component (PC) 1, 2 and 3 over the spectral range 5000–750 cm⁻¹ accounting for 50%, 28% and 10% variability, respectively.

responsible for the reported variation.

3.2. Site 1

The TPH concentration range of site 1 samples (50) spanned from 0 to 60,000 mg/kg (average 10,359 mg/kg, standard deviation (SD) 15,696 mg/kg). The variation amongst the soils is illustrated through the standard deviation spectra and mostly the first two PCA loadings (Fig. 1a and b respectively). The standard deviation spectrum shows that the soil variability of the sample set across

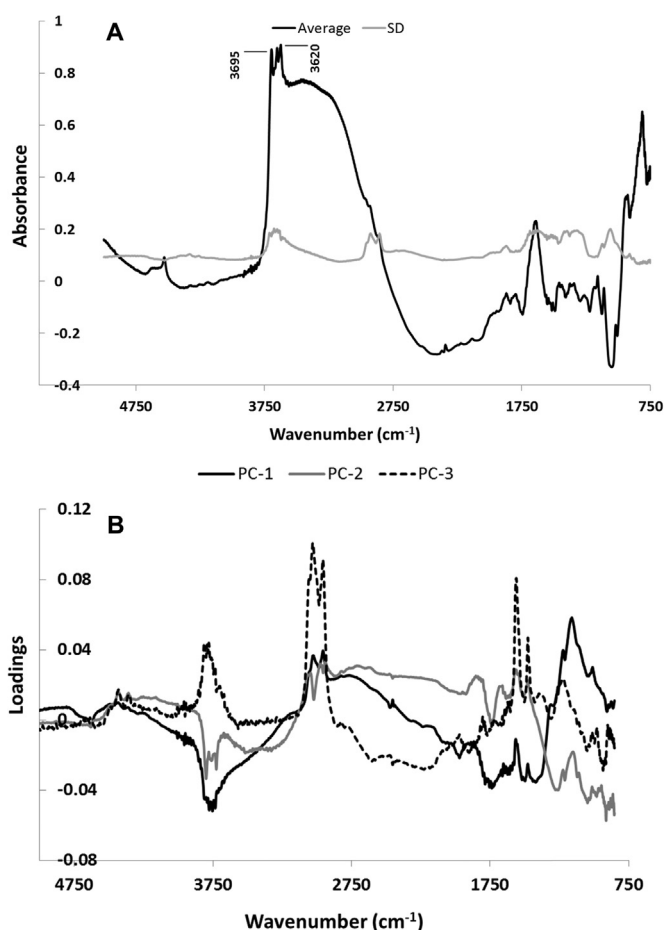


Fig. 2. (a) Auto-baseline corrected average and standard deviation (SD) spectra calculated from DRIFT spectra taken with 4100 ExoScan spectrometer of samples from site 2; (b) PCA loadings for PC 1, 2 and 3 over the spectral range 5000–750 cm^{-1} accounting for 80%, 9% and 5% variability, respectively.

site 1 was mainly due to changes in TPH. A generic feature of the average spectrum was a peak at 3620 cm^{-1} due to 2:1 layer silicates such as smectite and illite clays, a broad water band associated with such clays at $3600\text{--}3300\text{ cm}^{-1}$, and strong peaks due to carbonates (typical of dolomite) featured at 3023 cm^{-1} , 2627 cm^{-1} and 2523 cm^{-1} .

The PCA loadings plot (Fig. 1b) shows peaks associated with TPH concentration (peaks from $3000\text{--}2800\text{ cm}^{-1}$) and are proposed to be the main components accounting for variations along PC 1 and PC 2 (which accounted for 78% of the variation: 50% and 28%, respectively). In addition, loadings along PC 2 were associated with carbonates, in particular dolomite, showing peaks at 3023 cm^{-1} and 2627 cm^{-1} . Loadings along PC 3 were related to particle size (negative correlation with quartz).

3.3. Site 2

The TPH concentration range of site 2 samples (48) spanned from 0 to $40,000\text{ mg/kg}$ (average $5,603\text{ mg/kg}$, SD $10,332\text{ mg/kg}$), with the average TPH concentration similar to Site 1. The first principal component (PC 1) accounted for 80% of the variability in the soil data set with only 9% variability assigned to PC 2. According to the average and standard deviation spectra (Fig. 2a), samples showed a little more variability in the soil profile with variations mainly due to changes in kaolinite clay, TPH content and particle size. The main chemistry of samples in site 2 seems to be dominated by a mixture of 1:1 and 2:1 clays (peaks at 3695 and 3620 cm^{-1} typically found for kaolinite, 3620 cm^{-1} for smectite/

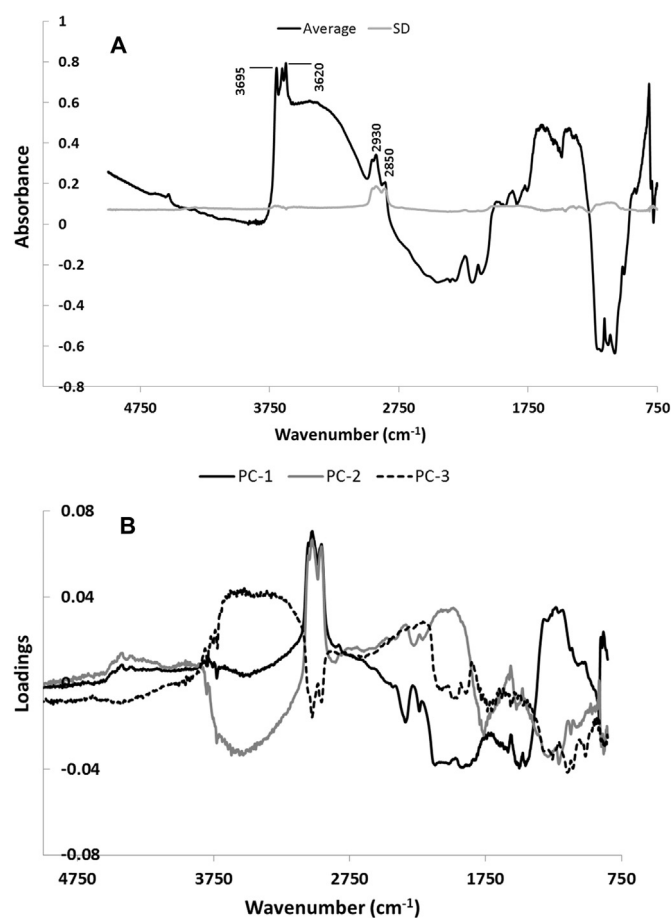


Fig. 3. (a) Auto-baseline corrected average and standard deviation (SD) spectra calculated from DRIFT spectra taken with 4100 ExoScan spectrometer of samples from site 3; (b) PCA loadings for PC 1, 2 and 3 over the spectral range 5000–750 cm^{-1} accounting for 68%, 11% and 9% variability, respectively.

illite, and a broad water band at $3600\text{--}3300\text{ cm}^{-1}$). In this case, carbonates were not observed.

The PCA loadings plot (Fig. 2b) confirms that the main variations in the chemical composition of the samples were due to clay (possibly kaolinite and pyrophyllite) and TPH content along PC 3. It was interesting to note that loadings 1 and 2 showed little evidence of TPH contribution suggesting that the soils from Site 1 may have been more variable in their mineralogy.

3.4. Site 3

The TPH concentration range of site 3 samples (43) spanned from 0 to $30,000\text{ mg/kg}$ (average $9,242\text{ mg/kg}$, SD $11,013\text{ mg/kg}$). The scores for PC 1 account for 68% of variability in the soil data set, with only 11% variability assigned to PC 2. According to the average and standard deviation spectra (Fig. 3a), most of the variability in the sample set comes from the TPH content alone. The main chemistry of samples in site 3 was dominated by 1:1 clays (peaks at 3695 and 3620 cm^{-1}), soil organic matter (very strong broad peak due to organic --OH at about 3400 cm^{-1} and --CH between $2930\text{--}2850\text{ cm}^{-1}$), TPH ($3000\text{--}2800\text{ cm}^{-1}$), and sand (strong peak inversion at $1100\text{--}1000\text{ cm}^{-1}$). There was a high degree of overlap between the organic matter and TPH peaks in the $3000\text{--}2800\text{ cm}^{-1}$ region.

The PCA loadings plots (Fig. 3b) confirmed that the main variations in the chemical composition of the samples were due to organic matter --OH , TPH and quartz (sand) content.

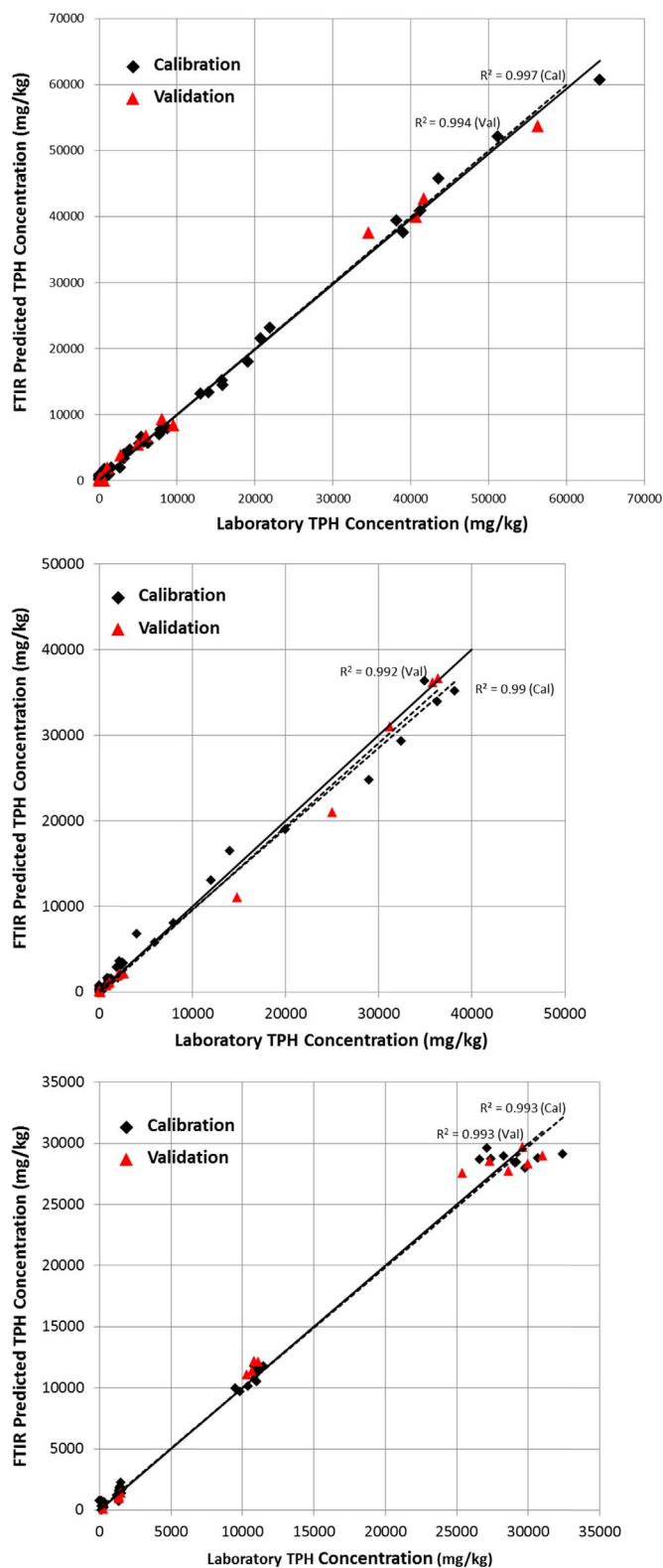


Fig. 4. (a) Laboratory versus FTIR predicted TPH concentration over 0–60,000 mg/kg range showing cross-validation ($n=50$, \blacklozenge) and “independent” validation samples ($n=17$, \blacktriangle) for Site 1. The solid line is the $x=y$ line and the dashed line is line of best fit (calibration and validation lines overlap). (b) Laboratory versus FTIR predicted TPH concentration over 0–40,000 mg/kg range showing cross-validation ($n=48$, \blacklozenge) and “independent” validation samples ($n=19$, \blacktriangle) for Site 2. The solid line is the $x=y$ line and the dashed lines is line of best fit for both calibration and validation samples. (c) Laboratory versus FTIR predicted TPH concentration over 0–30,000 mg/kg range showing cross-validation ($n=43$, \blacklozenge) and “independent” validation samples ($n=17$, \blacktriangle) for Site 3. The solid line is the $x=y$ line and the dashed lines is line of best fit for both calibration and validation samples.

3.5. Partial least squares models and predictions

Fig. 4a–c show the cross-validation and independent validation plot for Sites 1, 2 and 3, respectively, comparing the FTIR predicted TPH values with the laboratory TPH values over the full concentration range. As shown in Table 1, all calibration models were highly successful ($R^2 > 0.96$), irrespective of the concentration range and site considered. The RMSECV values decreased with reduced concentration range, confirming that the modelling errors were influenced by the concentration range. For similar concentration ranges, higher RMSECV values were found for the Site 2 models, in spite of the wide concentration ranges (0–40,000 mg/kg and 0–20,000 mg/kg) of the other sites. This was likely caused by errors in sampling due to the higher variability in the soil profile for this set compared with the other two sites. Models generated from contaminated field samples are more challenging to model due to the difficulty of collecting a spread of concentration from a single soil profile at site. This is a likely reason for the observed difference between the established models. However, the number of PLSR factors required for the optimum models in set 1 (on average 7) was higher than for set 2 (4) and set 3 (4). The increased number of factors may be related to the high level of background soil information from carbonates present within Site 1.

The soils that were in the low concentration range for Site 1 (0–3,000 mg/kg) were ground using a mortar and pestle to minimise interference from carbonates in the soil background. This reduced the number of factors from 8 to 4, making the model more robust for predicting TPH at these lower concentrations. Additionally, the RMSECV of the model reduced from 180 mg/kg to 123 mg/kg when ground.

Table 2 presents the statistics for the independent validation for Sites 1, 2 and 3. The Site 1 full range prediction model showed an RMSEP of 1,225 mg/kg for 17 samples, which was higher than the RMSECV of 899 mg/kg for the full range Site 1 cross-validation. This is likely due to the adverse background interference from carbonates in Site 1. When scanning low concentration ($< 3,000$ mg/kg) samples, it was observed that the RMSEP was comparable to the RMSECV, which indicates that grinding the soil effectively reduced the error from heterogeneity of the background soil carbonate matrix and minimised the degree of interference (data not shown).

Validation of Site 2 TPH concentration resulted in an RMSEP of 1,293 mg/kg for 19 samples, comparable to a cross-validation RMSECV of 1,255 mg/kg. In this case, there was no underlying background soil interference to overcome. Hence, the statistics for

Table 2

Comparison of FTIR TPH model statistics for independent validation samples for sites 1, 2 and 3 for the full concentration range model.

	Site 1	Site 2	Site 3
$\text{Sum}(x - x')$	23,995,987	30,105,132	19,044,359
n	17	19	17
RMSEP (mg/kg)	1,225	1,293	1,091
RMSECV (mg/kg)	899	1,255	930
R^2 (val)	0.994	0.992	0.993
TPH Average (mg/kg)	10,359	5,603	9,242
TPH SD (mg/kg)	15,696	10,332	11,013
RPD	13	8	10

x =actual laboratory TPH concentration (mg/kg).

x' =4100 ExoScan predicted TPH concentration (mg/kg).

n =number of samples.

RMSEP=root mean square error of prediction.

RMSECV=root mean square error of cross validation (as per Table 1).

R^2 =coefficient of determination.

SD=standard deviation.

RPD=residual predictive deviation.

the blind validation samples were comparable to the statistics of the cross-validation model. Similarly, for Site 3 the RMSEP was 1,091 mg/kg for 17 samples, which was comparable to the RMSECV of 930 mg/kg.

As previously indicated by the high R^2 and low RMSECV, a very close correlation was observed in Fig. 4a–c between predicted and laboratory TPH values for both cross-validation and independent validation samples. The RPD values for sites 1, 2 and 3 show values of 13, 8 and 10, respectively. These are high RPD values and are considered to provide analytical quality predictions.

4. General discussion

4.1. Comparison with previous studies

It can be challenging to put these results into context with other studies since, to the best of the authors' knowledge, there have been no reported previous studies available on the use of hand-held MIR instrumentation for the prediction of TPH across various soil types and concentration ranges. If we compare with previous studies using vis-NIR for the prediction of TPH in field samples (Chakraborty et al. [12], $R^2=0.63$ and $RPD=1.94$; Chakraborty et al. [13], $R^2=0.78$ and $RPD=2.19$) our results can be regarded as more acceptable for prediction modelling. Although researchers have reported [12] that random loss of TPH in air-dried samples is the principal contributor for poor performance of TPH models, a previous study (data not shown) undertaken by the authors showed minimal loss of TPH up to C_{40} . Whilst there may be some loss of the semi-volatile fractions ($\sim C_{10}-C_{14}$), the main contaminant being targeted is diesel range organic compounds, which do not contain much of this fraction especially if the soils have been weathered. Furthermore, the air-dried TPH calibrations we present in this manuscript show excellent performance so any loss in TPH did not affect the accuracy. Additionally, samples were kept chilled until they were received in the laboratory. Immediately after the samples were air-dried and sub-sampled, each sub-sample was scanned with the FTIR instrument and sent directly on ice to the laboratory for TPH analysis. If needed, the samples were stored in a fridge at 4 °C to limit biodegradation of the TPH and avoid change of the TPH profile.

According to previous research [5,17] the presence of free water affects the NIR and MIR spectra for the prediction of TPH by reducing the intensity of peaks related to these contaminants. At high moisture contents the non-linearity effect dominates and is a much stronger effect in the MIR. However, the NIR is also affected since the response to the TPH presence is based on overtones and combination of fundamental vibrations occurring in the MIR region [18]. As a result, air drying can reduce this effect, thus improving the accuracy of TPH predictions.

All models were developed with the samples being previously dried. This minimises the error from water in order to obtain the required accuracy for the technology to be used for routine regulatory applications. In order to account for the effects of potentially high moisture contents, an algorithm has been incorporated in the prediction software which prevents the user from obtaining a result if the level of free moisture in the soil is high and may negatively affect the accuracy of the predicted values. In such conditions, drying a very thin layer of soil sample is required. Thus, an alternative solution would be to provide a rapid method to dry the samples in the field in order to obtain accurate results for field screening wet samples by the PLSR DRIFT method. In dry weather conditions a direct measurement can be taken of the soil.

4.2. Effect of soil type in the PLSR models

Throughout the previous section we have shown models for the prediction of TPH using a hand-held IR instrument across different soils. The soil sets discussed in this manuscript represent distinctive particle sizes and mineralogy. As described above, soil Set 1 is characterised by 2:1 clays and carbonates (dolomite). Soil set 2 is characterised by a mixture of 1:1 and 2:1 clays, and soils from set 3 were dominated by the presence of sand (quartz). Particle size is known to be important for explaining, in general terms, the IR response of contaminant concentration, and, in particular, for TPH [5,17,18]. The presence of soil aggregates associated with clay soil types may have a shielding effect on the incident infrared radiation on the adsorbed TPH. This shielding reduces the IR radiation absorbed by the TPH and results in weaker TPH peaks [4]. In contrast, TPH accumulates differently in sandy soils where the TPH may occur as films on or between the sand particles. This enhances access to the IR radiation and thus presents stronger peaks in the spectrum. However, this manuscript shows that we can compensate for the different infrared response from various soil types by generating soil specific models. Additionally, it was observed that similar model performance using soil specific calibrations was achievable for soil sets containing carbonates. It is well known that carbonates can have a negative effect on the IR response of the TPH peaks [4], which was partially reduced by grinding the soils. This was thought that for unground soils (sieved to < 2 mm) carbonate may have a heterogeneous distribution throughout the sample at the < 2 mm scale, resulting in an underlying influence in the signal of TPH and a more variable IR signature. However, grinding the samples would homogenise the distribution of carbonate particles so the IR response to the background carbonate signal is effectively constant. This results in a consistent IR pattern/spectra, which is sensitive to changes in TPH concentrations alone and is not affected by carbonates.

For the prediction of any “unknown” samples, we need to make sure that the variability of these samples is well represented within the calibration set [19,20]. It is thus expected that the accuracy of TPH predictions will likely be degraded for soils that are not (well) represented in the calibration set. Hence, if the intention is to use the FTIR instrument for the prediction of TPH in any unknown soil worldwide, then there is a need for a more universal spectral library. Consequently, a procedure to build a variety of soil specific calibration models with the outlook of creating a library of calibration models that covers a wide range of soil matrices would be highly advantageous. Having access to a library of soil specific calibration models would allow environmental chemists, risk assessors and site managers to accurately predict TPH concentration for an unknown sample by allocating the sample to the best suited calibration model from the library by using the spectral information. This would ensure that the highest accuracy can be achieved. Currently, the instrument does not yet automatically select soil models for prediction improvement, but is the subject of the development of a universal calibration outside the scope of the present article.

4.3. Effect of TPH concentration range in the PLSR models

We have shown how the RMSECV increases with higher concentration ranges. This is something that is expected from high concentration compared with low concentration samples. In addition, and as previously shown by Forrester et al. [4], modelling different concentrations of TPH result in different PLSR loadings, possibly indicating a different chemistry of low and high TPH concentration samples. This basically means that some wavenumbers, selected as important for a large concentration range

model, may not be important for a narrow concentration and vice versa. Thus, errors for the prediction of low TPH samples may be increased by the inclusion of high TPH samples in the models. Errors can be also increased by having relatively few samples which are not representative of the rest of the set, but leverage the model. This suggests that the use of a large TPH concentration range model for the prediction of low TPH samples may not be advisable. Within this study we propose a cascade method, consisting first in the prediction of an unknown sample using the full concentration range model. If the sample is predicted high (i.e. close to the upper limit) we would take this value as acceptable. However, if the sample is predicted low, a model containing samples with similar low concentrations would be selected.

As discussed above, modelling high TPH concentration samples can be challenging. For high TPH, in this article we propose the use of the NIR (4500–4200 cm^{-1}) spectral region, in combination with the MIR, to help model such high concentrations. This is because very high concentrations of TPH can occur as reflective films on the sample particles, resulting in specular reflectance and interference fringing in the MIR, distorting the TPH peaks. Such distortions are not observed in the NIR due to the far lower absorption index in the NIR. By including the NIR range in high TPH models, high concentrations can be thus more accurately characterised and predicted.

5. Conclusions

In this article we have demonstrated the successful performance of a hand-held MIR spectrometer for the accurate prediction of TPH concentrations in soils, irrespective of the soil type and TPH concentration. Models with $R^2 \sim 0.99$ and $\text{RMSE} < 200 \text{ mg/kg}$ were obtained for the prediction of $< 3,000 \text{ mg/kg}$ TPH. High TPH concentration samples (0–60,000 mg/kg) were also accurately predicted ($R^2 \sim 0.99$, $\text{RMSEP} 1,225 \text{ mg/kg}$) when combining NIR and MIR regions in the model. The prediction of TPH using the proposed technology is cost-effective, accurate and rapid, and is suitable for commercial routine use. Possible applications would include site assessment, emergency spill response, earthworks validation, bioremediation monitoring and potentially, petroleum exploration.

Acknowledgements

The authors gratefully acknowledge Agilent Technologies for their technical support and Ziltek Pty Ltd for funding of this work.

References

- [1] Agency for Toxic Substances and Disease Registry (ATSDR), Toxicological

- profile for total petroleum hydrocarbons (TPH), Atlanta, GA, U.S. Department of Health and Human Services, Public Health Service. (<http://www.atsdr.cdc.gov/toxfaqs/faq.asp?id=423&tid=75>), 1999 (accessed 12.08.14).
- [2] G.D. Ritchie, K.R. Still, W.K. Alexander, A.F. Nordholm, C.L. Wilson, J. Rossi III, D. R. Mattie, A review of the neurotoxicity risk of selected hydrocarbon fuels, *J. Toxicol. Environ. Health Part B: Crit. Rev.* 4 (2001) 223–312.
- [3] Environmental Protection Agency, Field measurement technology for total petroleum hydrocarbons in soil, Document prepared by Tetra Tech EM Inc, Chicago, Illinois, 2000.
- [4] S.T. Forrester, L.J. Janik, M.J. McLaughlin, J.M. Soriano-Disla, R. Stewart, B. Dearman, Total petroleum hydrocarbon concentration prediction in soils using diffuse reflectance infrared spectroscopy, *Soil Sci. Soc. Am. J.* 77 (2013) 450–460.
- [5] D.F. Malley, K.N. Hunter, G.R.B. Webster, Analysis of diesel fuel contamination in soils by near-infrared reflectance spectrometry and solid phase micro-extraction–gas chromatography, *J. Soil Contam.* 8 (1999) 481–489.
- [6] J.M. Soriano-Disla, L.J. Janik, R.A. Viscarra Rossel, L.M. McDonald, M. J. McLaughlin, The performance of visible, near and mid-infrared spectroscopy for prediction of soil physical, chemical and biological properties, *Appl. Spectrosc. Rev.* 49 (2014) 139–186.
- [7] S.T. Forrester, L.J. Janik, M.J. McLaughlin, In-situ determination of total petroleum hydrocarbon (TPH) contamination: a quick infrared spectroscopic test for TPH at contaminated sites, in: Proceedings of the 19th World Congress of Soil Science, Brisbane, Australia, August 1–6, 2010.
- [8] J.B. Reeves, G.W. McCarty, W.D. Hively, Mid-versus near-infrared spectroscopy for on-site analysis of soil, in: R.A. Viscarra-Rossel, A.B. McBratney, B. Minasny (Eds.), *Proximal Soil Sensing*, Springer Science+Business Media, New York, 2010.
- [9] B. Kuang, H.S. Mahmood, M.Z. Quraishi, W.B. Hoogmoed, A.M. Mouazen, E. J. van Henten, Sensing soil properties in the laboratory in situ, and on-line: a review, *Adv. Agron.* 114 (2012) 155–223.
- [10] N.A. O'Brien, C.A. Hulse, D.M. Friedrich, F.J. Van Milligen, M.K. Von Gunten, F. Pfeifer, H.W. Siesler, Miniature near-infrared (NIR) spectrometer engine for hand-held applications, *Proc. SPIE* 8374 (2012).
- [11] S. Altinpinar, D. Sorak, H.W. Siesler, Near infrared spectroscopic analysis of hydrocarbon contaminants in soil with a hand-held spectrometer, *J. Infrared Spectrosc.* 21 (2013) 511–521.
- [12] S. Chakraborty, D.C. Weindorf, C.L.S. Morgan, Y. Ge, J.M. Galbraith, B. Li, C. S. Kahlon, Rapid identification of oil-contaminated soils using visible near-infrared diffuse reflectance spectroscopy, *J. Environ. Qual.* 39 (2010) 1378–1387.
- [13] S. Chakraborty, D.C. Weindorf, B. Li, A.A. Ali Aldabaa, R.K. Ghosh, S. Paul, M. Nasim Ali, Development of a hybrid proximal sensing method for rapid identification of petroleum contaminated soils, *Sci. Total Environ.* 514 (2015) 399–408.
- [14] R.N. Okparanma, A.M. Mouazen, Determination of total petroleum hydrocarbon (TPH) and polycyclic aromatic hydrocarbon (PAH) in soils: a review of spectroscopic and nonspectroscopic techniques, *Appl. Spectrosc. Rev.* 48 (2013) 458–486.
- [15] K. Varmuza, P. Filzmoser, *Introduction to Multivariate Statistical Analysis in Chemometrics*, Taylor & Francis, Boca Raton, FL, 2009.
- [16] H. Wold, Partial least squares, in: S. Kotz, N.L. Johnson (Eds.), *Encyclopaedia of Statistical Sciences*, John Wiley and Sons, New York, 1985, pp. 581–591.
- [17] G. Hazel, F. Bucholtz, I.D. Aggarwal, G. Nau, K.J. Ewing, Multivariate analysis of mid-IR FT-IR spectra of hydrocarbon-contaminated wet soils, *Appl. Spectrosc.* 51 (1997) 984–989.
- [18] R.N. Okparanma, A.M. Mouazen, Combined effects of oil concentration, clay and moisture contents on diffuse reflectance spectra of diesel-contaminated soils, *Water Air Soil Pollut.* 224 (2013) 1–8.
- [19] P. Dardenne, G. Sinnaeve, V. Baeten, Multivariate calibration and chemometrics for near infrared spectroscopy: which method? *J. Infrared Spectrosc.* 8 (2000) 229–237.
- [20] D.J. Brown, R.S. Bricklemeyer, P.R. Miller, Validation requirements for diffuse reflectance soil characterization models with a case study of VNIR soil C prediction in Montana, *Geoderma* 129 (2005) 251–267.