Infrared spectroscopy and chemometric modelling of organic carbon measured by Rock-Eval pyrolysis of UK shale rock

Darren J. Beriro* and Christopher H. Vane

British Geological Survey, Keyworth, Nottingham NG12 5GG, UK

 DJB, 0000-0003-2680-1124; CHV, 0000-0002-8150-3640

*Correspondence: darrenb@bgs.ac.uk

Abstract: Shale rock core from the Bowland Shale Formation, UK, was analysed in the laboratory using Rock-Eval pyrolysis and Fourier transform infrared spectroscopy (FTIR). These methods are used to characterize the organic constituents of soil and rock. This research is a proof-of-concept study to investigate whether regression models developed using FTIR and Rock-Eval data for the same length of core can be used to estimate selected Rock-Eval parameters. The accuracy of the regression models was assessed using statistical methods, the results of which were used to choose preferred models for each Rock-Eval parameter. The models produced were shown to have an acceptable level of uncertainty for total organic carbon and S1, S2 and S3 outputs, which led us to conclude that these are potentially suitable for estimating unknown down-core Rock-Eval parameter values. Conversely, the model for the temperature of the maximum rate of hydrocarbon generation (T_{max}) had higher variability in the cross-validation data above the acceptable level of uncertainty, which could lead to erroneous estimates. Down-core interpolations of selected Rock-Eval parameters could be practically achieved by modelling FTIR data by maintaining standard sample frequencies for Rock-Eval while supplementing with higher frequencies for FTIR and chemometric analysis.

Unconventional shale gas systems comprise black shale mudrocks that are self-contained hydrocarbon source and reservoir because they potentially support large quantities of methane that can be recovered at surface using hydraulic fracturing. Alongside inorganic mineral content and physical properties, the hydrocarbon generation capacity, organic matter (OM) type and thermal maturity are the determining factors used to estimate the gas-in-place and producible gas resources (Jarvie 2012; Andrews 2013; Scotchman 2014; Whitelaw *et al.* 2019; Li *et al.* 2021). During hydrocarbon exploration, understanding of the main organic geochemical factors is usually elicited from the interpretation of Rock-Eval pyrolysis and total organic carbon (TOC %) analyses, validated using optical methods such as vitrinite reflection (Whitelaw *et al.* 2019). However, new approaches to analysis and data interrogation such as infrared spectroscopy and chemometrics have been recently shown to provide additional utility (Leach *et al.* 2008; Fletcher *et al.* 2014; Chen *et al.* 2014*a*).

Rock-Eval is routinely employed during the appraisal of shales, coals and other sediments because it yields key information on kerogen type and maturity (Table 1), requires minimal sample preparation and uses just 10–60 mg of rock (Behar *et al.* 2001; Marshall *et al.* 2015; Newell *et al.* 2016; Waters *et al.* 2020). This traditional

geochemical screening approach is valuable to the exploration community because of the large body of literature, including: (1) data from basins in North and South America, Europe (UK, France, Poland), Africa and Australia; (2) published information on the technical limitations of the analyses (e.g. mineral matrix effects, suppression of the temperature of the maximum rate of hydrocarbon generation (T_{max})); and (3) shale gas assessment-specific methodologies and unified data processing/presentation approaches (Espitalie *et al.* 1987; Cornford *et al.* 1998; Jarvie *et al.* 2007; Dembicki 2009; Fleury and Romero-Sarmiento 2016). However, recent studies of the main unconventional hydrocarbon play in the UK, namely the Carboniferous Bowland Shale Formation, has shown fairly thin organic rich intervals interbedded with calcareous mudstones and turbidites (Andrews 2013; Słowakiewicz *et al.* 2015; Könitzer *et al.* 2016; Hennissen *et al.* 2017; Emmings *et al.* 2019; Waters *et al.* 2020). Given this cyclical and laterally variable pattern of OM quantity and type, there is a need for cost-effective, high-resolution geochemical screening methods.

One complementary approach to evaluate the hydrocarbon generation capacity and characterize OM types in shale mudrocks is infrared spectroscopy, coupled with statistical methods collectively referred to as chemometrics. Chemometrics can establish whether relationships exist between

Table 1. *Acceptance criteria used for model evaluation*

Method	Description	Acceptance criteria
Scatter plot	Scatter plot comparing observed Rock-Eval and modelled data	Conformity to one-to-one line
Goodness-of-fit statistics	MAE RMSE R^2 MPAE	MAE: low RMSE: low R^2: high MPAE: low
Residuals analysis	Scatter plot of the residual difference between measured and modelled data	Absence of pattern in the scatter of the residual points on plot and very low R^2
Absolute error analysis	Line plot comparing measured observed modelled data	Broad congruence and trend of peaks
Percentage error analysis	Plot showing percentage error	Low

MAE, mean absolute error; MPAE, mean percentage absolute error; RMSE, root-mean-squared error.

properties e.g. using FTIR and OM data such as lignin, organic carbon and polycyclic aromatic hydrocarbon contents and selected physico-chemical properties (Vane *et al.* 2003; Leach *et al.* 2008; Mas *et al.* 2010; Mostert *et al.* 2010; Washburn and Birdwell 2013; Tinti *et al.* 2015; Craddock *et al.* 2017; Varma *et al.* 2018; El-Rub *et al.* 2019).

IR spectroscopy quantifies the transitions between energy states after molecules that possess dipole moments are excited (Chen *et al.* 2015). The advantage of this technique over Rock-Eval is that it is non-destructive, it can be run expediently in the laboratory or field and has a low cost per sample (Washburn and Birdwell 2013). The IR causes the molecules to vibrate – the movements include stretching, bending, twisting, rocking, wagging and out-of-plane deformation. Each vibration type relates to the corresponding molecular bonds e.g. C – H, O – H and C = O. The intensity of the response is proportional to the abundance of the functional groups present in the sample (Chen *et al.* 2014*a*), although quantification in geological samples is difficult because of intrinsic sample heterogeneity. Characteristics of hydrocarbon source rocks, including the amount and type of organic carbon and mineral composition, have been evaluated by IR spectroscopy (Leach *et al.* 2008; Fletcher *et al.* 2014; Chen *et al.* 2014*a*). Mid-IR spectroscopy quantifies the molecular response of a sample to infrared radiation, typically from wavenumbers 4000 to 400 cm^{-1} (a wavenumber is a measure of the spatial frequency of waves per cm). The response provides information on the functional group chemistry of samples, which in this research is being used to supplement analysis by Rock-Eval.

This study investigates whether the application of FTIR chemometrics to Carboniferous shales from the Edale Formation, UK, (a lateral equivalent of the Bowland Shale Formation) can estimate Rock-Eval parameter values, which form the traditional basis of characterizations of hydrocarbon generation capacity, OM type and thermal maturity. The approach may therefore provide a useful supplement to Rock-Eval pyrolysis where high-resolution stratigraphic control is needed on organic carbon type, but economics preclude substantial analytical programmes.

Materials and methods

Karenight core from the Bowland Shale Formation

The Karenight-1 (SK36NWBJ13) borehole used for this study was drilled for mineral exploration by Drilling and Prospecting International in 1973 and stored in temperature- and moisture-controlled conditions at the UK Research and Innovation National Geological Repository hosted at the British Geological Survey. Karenight is located 5 km NE of Matlock, Derbyshire (1.53° W, 53.18° N), and is situated on the boundary of the southern extent of the Derbyshire structural high and northern extent of the Edale Gulf (Fig. 1: (Waters *et al.* 2009)). The rock core was taken from the Edale Shales Group from the Carboniferous period (Pendleian substage and the *Eumorphoceras* zone) (Wilson and Stevenson 1973). The core interval analysed spans 17 m and was located between 235 m and 252 m below ground level (m bgl). The stratigraphy comprises grey-coloured dark mudstone with carbonaceous layers with thin laminar siltstone between 235 and 245 m bgl and limestone between 245 and 250 m bgl (Fig. 2) (Hennissen *et al.* 2017). Seventy-two samples were extracted from the core to reflect the main lithological units. Each sample was freeze-dried and ground to a fine powder.
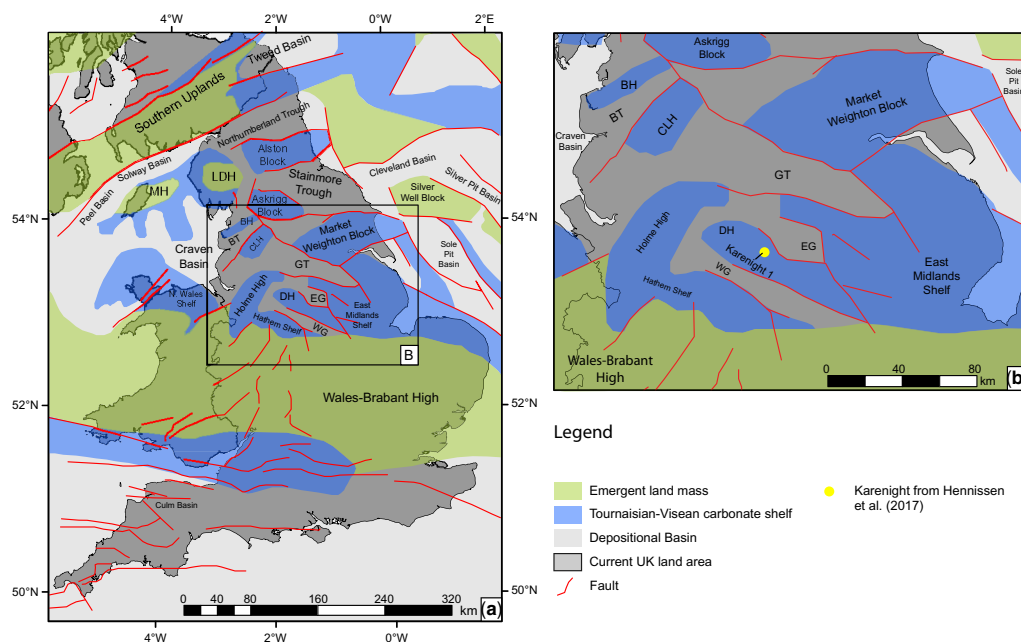
Fig. 1. Maps of (**a**) the study area and (**b**) the location of the Karenight-1 borehole. Source: based on Waters *et al.* (2009) © 2009 BGS (NERC). Contains Ordnance Survey data © Crown copyright and database rights 2021. Ordnance Survey Licence No. 100021290.

Rock-Eval. Each sample was analysed using a Rock-Eval(6) pyrolyser, configured in standard mode (pyrolysis and oxidation as a serial process). Each sample (*c.* 60 mg dry wt) was heated isothermally at 300°C for 3 min and then from 300 to 650°C at 25°C min^{-1} in an inert atmosphere of N_2. The residual carbon was oxidized at 300°C (isothermal 1 min) and then from 300 to 850°C at 20°C min^{-1} (hold 5 min). The performance of the instrument was checked every ten samples against the accepted values of the Institut Français du Pétrole (IFP) standard (IFP 160 000, S/N15-081840). Rock-Eval parameters were calculated by integration of the amounts of thermally vaporized free hydrocarbons (HC) expressed in mgHC g^{-1} rock (S1) and hydrocarbons released from cracking of bound OM expressed in mgHC g^{-1} rock (S2) (Engelhart *et al.* 2013). The hydrogen index (HI) was calculated from S2 × 100/TOC and the oxygen index (OI) from S3 × 100/TOC.

FTIR. Each sample was subject to mid-IR spectroscopy using FTIR (wavenumbers: 4000 to 400 cm^{-1}). A Bio-Rad FTS 3000 FTIR instrument was operated in diffuse reflectance mode using Resolutions Pro software (Version 5.1). Approximately 300 mg of freeze-dried and milled sample was transferred to a stainless steel cup and gently tamped and levelled using a stainless steel spatula and rod. The cup and sample were subsequently placed into Pike AutoDiff autosampler. Prior to analysis the alignment of the interferometer was checked and adjusted to optimize the signal reaching the detector. Background FTIR measurements were made using finely ground potassium bromide (KBr) prior to each run. Each sample, including background, was analysed in replicate. The first analysis was made in the position the sample was placed in the AutoDiff, the second following a 90° rotation of the cup and sample. Samples were scanned 40 times at 4 cm^{-1} resolution.

Data evaluation and predictive modelling

Data evaluation and predictive modelling was conducted using the R programming language (R Development Core Team 2016). Down-core profiles of the Rock-Eval data were produced to visualize geochemical variations and allow a qualitative interpretation of the parameters measured.

FTIR spectra were visually assessed for any obvious differences between 'a' and 'b' duplicate samples. If there were no differences, mean spectra were calculated and baseline corrected using asymmetric least-squares trend estimation function in the 'ptw' R package (Gerretzen *et al.* 2014) followed by smoothing using a FIVE-point running mean.
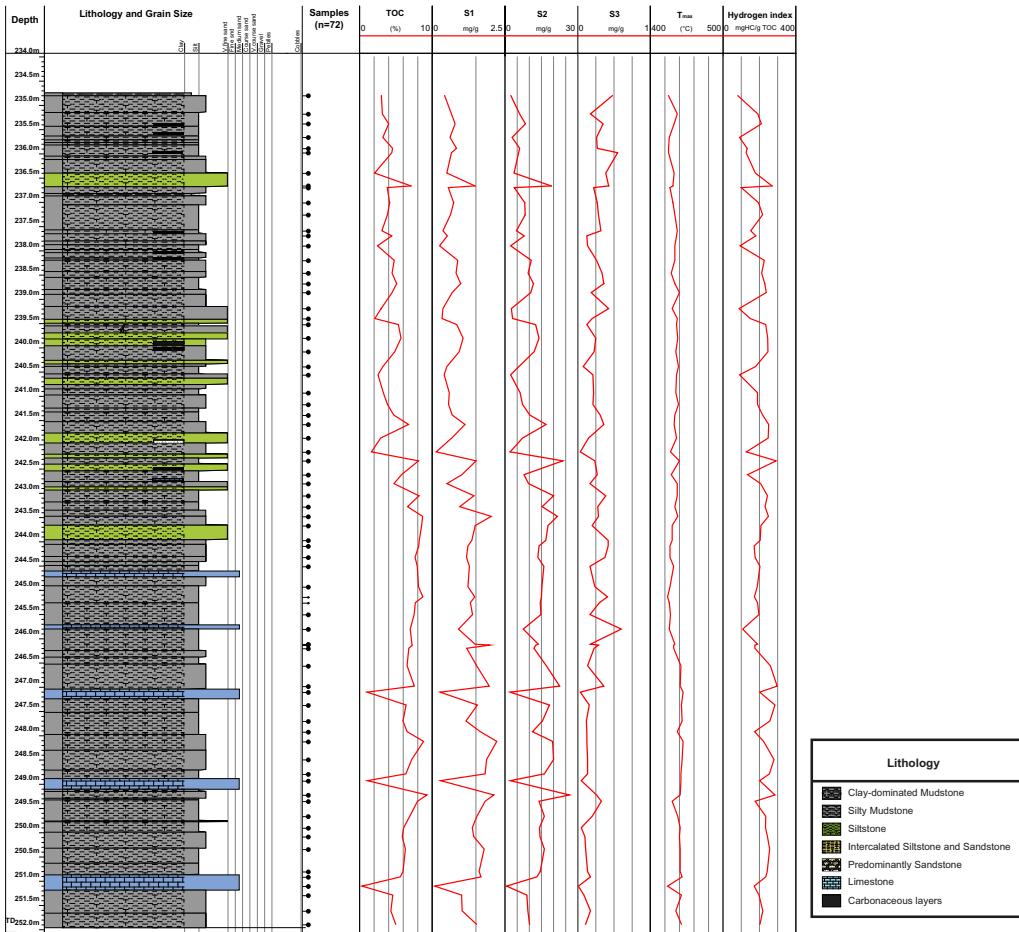
Fig. 2. Karenight lithology and Rock-Eval down-core plots. Source: based on Hennissen *et al.* (2017) © 2017 BGS (NERC), available under CC BY 4.0.

Principal components analysis (PCA) was conducted using a centred and scaled co-variance matrix of all FTIR wavenumbers and samples. These data were used to produce a number of uncorrelated vectors (loadings). The linear equation for the loadings data produces coefficients called principal component scores for each sample. The scores were plotted to assess whether there was any visual association with Rock-Eval data. The PC loadings and coefficients were plotted to characterize the influence wavenumber.

Partial least-squares regression (PLSR) modelling of the of the baseline corrected mean smoothed FTIR data and selected Rock-Eval parameters (T_{max}, S1, S2 and S3) was completed using the orthogonal scores algorithm in the 'R caret' Package (Kuhn *et al.* 2015). HI and OI were calculated using the equations in Table 1. Three models were produced:

(1) all data (wavenumbers 4000 to 570 cm^{-1}); (2) fingerprint region (wavenumbers 570 to 2400 cm^{-1}); and (3) non-fingerprint region (wavenumbers 2400 to 4000 cm^{-1}). Selected dependent Rock-Eval parameters included: (1) T_{max}; (2) HI; (3) S1; (4) S2; and (5) S3. Multiple PLSR model runs were completed, each used a different number of PCs as an input parameter, ranging from 1 to 15 PCs. The optimum PC number was selected using 100-fold cross validation and the minimum root mean square error of predictions using the train function in the R caret package. The models developed using the optimized PC value were subject to 10-fold cross-validation to train and test the models by a leave-one-out (LOO) method using the mvr_dcv function in the R Chemometrics package. Each model was subject to detailed evaluation using the multiple acceptance criteria (Table 1). Beriro *et al.* (2013) reported that

the multiple selection criteria approach to model evaluation is valid to select best performance models from a number of potentially preferred models.

Results and discussion

Rock-Eval

The Rock-Eval pyrolysis data ($n = 72$) for S1, S2, S3, HI and T_{max} values are summarized by down-core plots (Fig. 2) and descriptive statistics (Table 2). The down-core profiles of the Rock-Eval parameters show broadly parallel trends for hydrocarbon composition indicated by TOC, S1 and S2 and HI. Conversely, the amount of oxygen and thermal maturity determined by S3 and T_{max} follow independent trends.

Weight percentage TOC provides a general indication of organic richness and is indirectly related to the hydrocarbon generation potential of a source rock. Proceeding down-core through the Karenight succession, three TOC depth interval zones are clear: 235 to 242 m bgl characterized by low but variable TOC ranging from 2.54 to 7.08% with a mean of 4.02% ($n = 31$), a second transitional zone between 242 and 243 m bgl with TOC of 4.74 to 8.18 and a mean of 6.69 wt % ($n = 6$) and a higher and less variable TOC interval from 244 to and including 252 m bgl, ranging from 0.36 to 8.20 with a mean of 6.4% ($n = 35$).

The Rock-Eval data presented in the down-core plots show that OM corresponds with the stratigraphic and lithological profile of the core. The OM content is generally higher in the mudstone bands between 245 and 250 m bgl than the thin laminar siltstone between 235 and 245 m bgl (Fig. 2) (Hennissen et al. 2017). The changes in lithology in each part of the core clearly affect the Rock-Eval parameter values. The OM in the limestones is low, whereas the siltstone presents higher values. The Rock-Eval HI also corresponds with the same change in lithology where the values are higher in the mudstone and interbedded limestone. Further

comparison of the TOC content suggests the highest values correspond to nine carbonaceous units, whereas the TOC content in siltstone, silty mudstone and finer-grained clay-dominated mudstone varies independently of lithology (Fig. 2). In contrast, to the siltstones, the limestones at 247, 249, 251 m have low TOC in the range of 0.36 to 1.09%. Taken individually, the TOC suggests the source rock potential of Karenight is good to very good (Peters et al. 2005). The TOC data presented here are higher than those reported for drill-cores of Namurian and Arnsbergian ages from Widmerpool Gulf (range 0.4–7.3%, mean TOC 3%, $n = 32$), those for the Namurian upper Bowland-Hodder unit from the Cleveland Basin (range 0.37–2.45%, mean 1.9, $n = 6$) and the Craven Basin (range 1.76–3.72%, mean 2.3% ($n = 36$) (Hough et al. 2014; Słowakiewicz et al. 2015; Könitzer et al. 2016). However, it should be noted that not all OM has equal hydrocarbon generative potential.
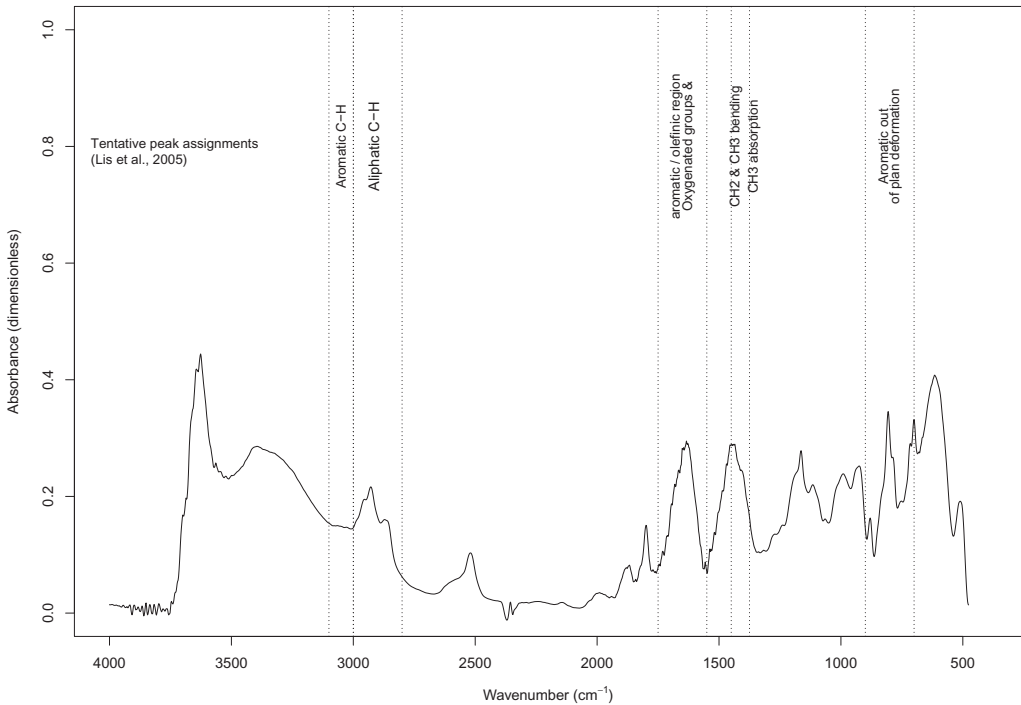
FTIR

The FTIR analysis of the Karenight samples is summarized by a mean spectrum for all samples (Fig. 3). The plot shows distinct peaks in the aliphatic (c. 2750–3000 cm^{-1}), carbonyl (c. 1700 cm^{-1}), CH_3 & CH_4 (c. 1400–1550 cm^{-1}) and aromatic deformation (c. 800 cm^{-1}). These wavenumber assignments are tentative and based on values published by Lis et al. (2005) and Chen et al. (2015). The data from these studies were derived from samples contained in a pressed potassium bromide (KBr) disc, whereas our FTIR data are produced using diffuse reflectance using the pure milled samples. The two approaches can produce slightly different peak assignments; for this reason, all assignments referred to herein are tentative.

The aliphatic region (c. 2750–3000 cm^{-1}) peaks produced for the Karenight core are similar to those produced for US shale samples originating from Anvil Pots mine and Garden Gulch Member (Washburn and Birdwell 2013). Again, the method (attenuated total reflectance-FTIR) is different from our

Table 2. *Descriptive statistics for Rock-Eval data*

	TOC (%)	S1 (mg HC g^{-1})	S2 (mg HC g^{-1})	S3 (mg CO_2 g^{-1})	T_{max} (°C)	HI (mg HC g^{-1} TOC)
Mean	5.43	1.07	11.4	6.48	435	201
Maximum	9.29	2.22	26.5	23.4	445	298
Minimum	0.36	0.09	0.62	1.60	424	82.0
Standard deviation	2.17	0.54	6.02	4.06	5.94	52.6
Relative standard deviation	40.0	50.3	52.8	62.6	1.37	26.2

$n = 72$.

Fig. 3. FTIR mean data spectrum.

study as it measures one specific sample point, rather than providing a diffuse reading, so direct comparison is subject to instrumental uncertainties. Further discussion of the Karenight core FTIR mean spectrum, especially in the fingerprint region (400–2000 cm^{-1}), is not possible due to the background noise caused by its inorganic (mineral) constituents. For example, the peak at 1400–1500 cm^{-1} could be CH_2, CH_3 bending or CH_3 absorption (Chen *et al.* 2014*b*) or be due to the presence of carbonates (Calderón *et al.* 2011). Further discrimination of the importance of this noise may be possible using complementary techniques such as X-ray diffraction or potentially pre-treatment of the sample to remove the inorganic constituents.

Baseline corrected and mean-smoothed FTIR spectra for all samples were reviewed in depth order (Fig. 4). There is no obvious depth-based trend in the spectra, although there are differences in the size of the aliphatic peaks (*c.* 2750–3000 cm^{-1}). Plots ordered by increasing Rock-Eval parameter values (S1, S2, S3 and T_{max}) show that for S1 and S2 there is an increase in the aliphatic region and a reduction in the aromatic and oxygenated functional group region (1850–1600 cm^{-1}). Peaks present in the 950–1050 cm^{-1} region show an increase in area and a peak at wavenumber *c.* 1400 cm^{-1} represents an increase in relative

height (CH_2 and CH_3). Reliable interpretation of these features is difficult because of the presence of the mineral component and the complexity of the OM.

There are a number of ways to overcome the issue of converging IR signals, including sample clean-up to isolate the organic constituents (Calderón *et al.* 2011; Fletcher *et al.* 2014), micro-FTIR (Chen *et al.* 2014*b*) and pyrolysis gas FTIR (Fletcher *et al.* 2014). Any additional sample preparation or analytical work on shale samples is likely to provide valuable information about their hydrocarbon characteristics. However, these approaches are time consuming and expensive, rendering them beyond the scope of this paper. This is because this study is designed to demonstrate the value of FTIR in rapid low-cost evaluation of shales to supplement hydrocarbon industry standard screening analysis, namely, Rock-Eval.

Chemometric modelling

PCA was undertaken using different parts of the FTIR data. The results were evaluated in the context of selected Rock-Eval trends and interpretation. T_{max} is defined as the temperature of the maximum rate of hydrocarbon generation measured on the S2 peak and is derived during the pyrolysis of non-volatile
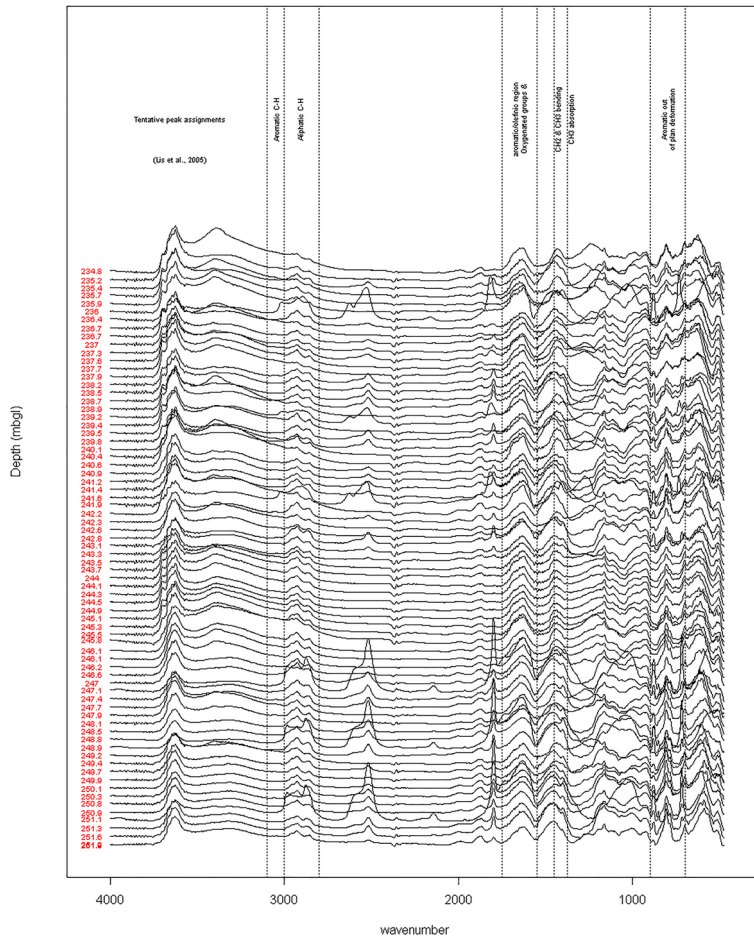
Fig. 4. FTIR sample data from all depths.

hydrocarbons contained within the sample (Behar *et al.* 2001). The oil–gas window of a shale is typically defined by an approximate cut-off point 435°C, below which the kerogen is considered immature and commercially unproductive (Słowakiewicz *et al.* 2015). The HI is also affected by thermal maturity and tends to decrease with maturity. An indicative HI cut-off point for a change in residual kerogen type (Type II/III oil prone to Type III to gas prone) is 200 mg HC g^{-1} TOC. The T_{max} and HI values do not lend themselves to direct identification using conventional FTIR data analysis because they are derived from either thermal cracking of bound hydrocarbons (in the case of T_{max}) or calculated parameters (in the case of HI). However, they are important indicators of thermal maturity and kerogen type that correspond with the organic composition of the shale rock, which for this reason mean they were chosen to aid with PCA interpretation.

A PCA bi-plot of the PC1 and PC2 scores with data points colour coded using the T_{max} and HI cut-off points (435°C and 200 mg HC g^{-1} TOC respectively) shows to two tentative clusters of samples and a limited number of outliers (Fig. 5).

The bottom right-hand cluster in Figure 5 (red data points: $T_{max} < 435$°C and HI < 200 mg HC g^{-1} TOC) comprises samples that are generally greater than 244 m bgl and correspond to mudstone with limestones lithology (Fig. 2). The top left-hand cluster (blue data points: $T_{max} > 435$°C and HI > 200 mg HC g^{-1} TOC) are shallower and match the mudstone with thin laminar siltstone (Fig. 2). The outliers are indicated on Figure 5 and, apart from one or two unexplained examples, relate to anomalies that appear to be associated with the core lithology and stratigraphy. These results show that the PCA is able to distinguish characteristics related to
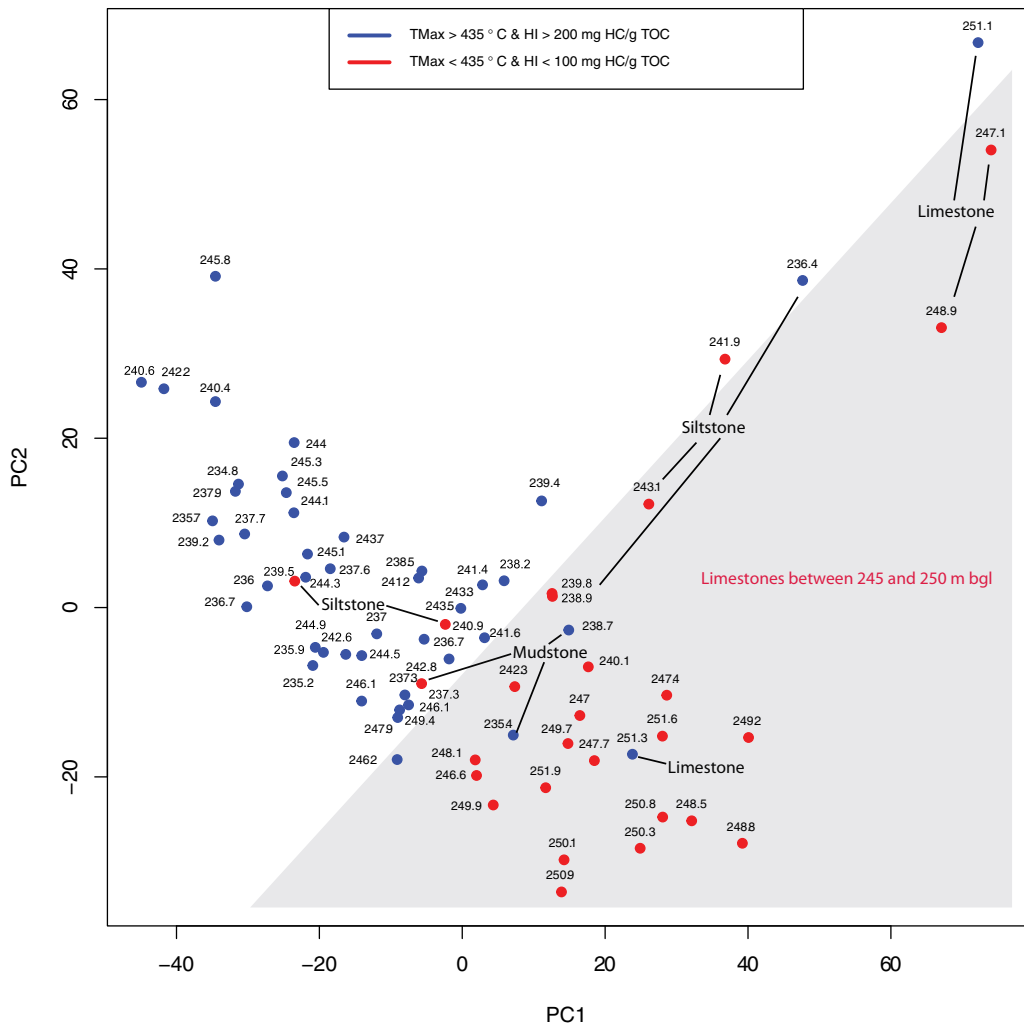
Fig. 5. FTIR principal components analysis bi-plot for PC1 and PC2.

the OM present in the mudstone not visible purely using the FTIR spectra. These characteristics correspond to the stratigraphy of the core, rather than to the thermal maturity of the OM.

The relative contribution to the loadings and coefficients for PC1 to PC4 made by each wavenumber was calculated and visualized (Fig. 6). The results suggest both the loadings and regression coefficients are strongly influenced by wavenumbers 3800–3400, 3300–2800, 2700–2400 and 2000–500 cm^{-1}. These wavenumbers reflect the trends in the mean spectrum discussed earlier (Fig. 3). The contributions of the wavenumbers to the loadings and coefficients are higher when the PCA is restricted to fingerprint and non-fingerprint region data. This suggests that the PLSR modelling might

provide better predictive models for the truncated datasets because of the stronger signals represented by the loadings and coefficients derived from the respective wavenumbers.

PLSR models using all data, with fingerprint and non-fingerprint regions as independent variables, were created using Rock-Eval parameters SI, S2, S3, T_{max} and TOC as dependent variables. HI and OI were calculated from modelled S2, S3 and TOC, respectively, using the equations shown in Table 1. Each of the PLSR models was evaluated using four goodness-of-fit statistics for the median cross-validation predicted values (Table 3). The most accurate models, defined using set criteria (Table 1), are: (1) S3 using all data; (2) T_{max}, S1 and S2 using the non-fingerprint region; and (3)
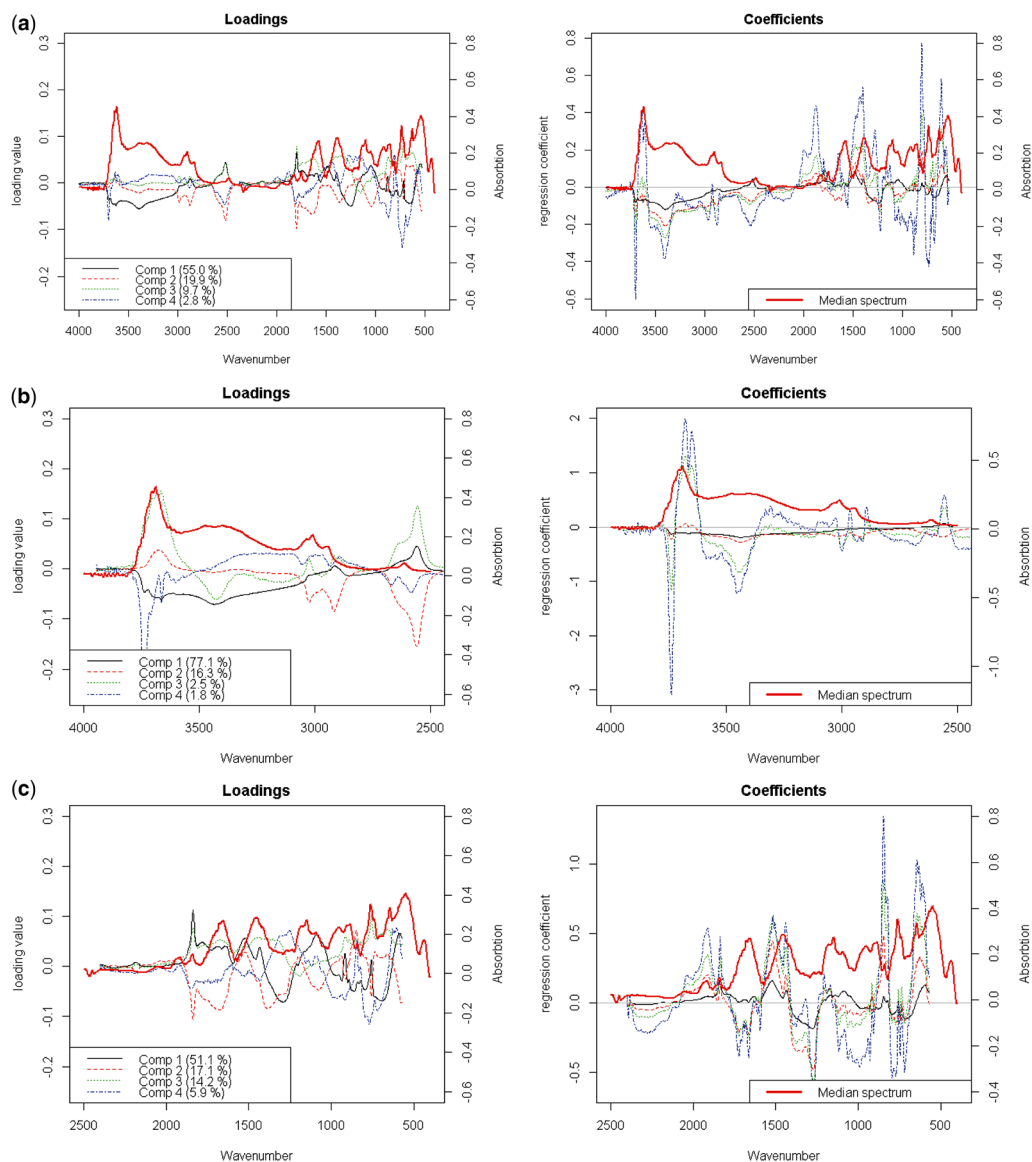
Fig. 6. PC loadings, regression coefficients and median spectra for PC1 to PC4 using all data for (**a**) all wavenumbers (4000–400 cm^{-1}), (**b**) the non-fingerprint (4000–2500 cm^{-1}) and (**c**) the fingerprint region (2500–500 cm^{-1}).

TOC using the fingerprint region. The fact that different sections of the FTIR spectra were used to produce the preferred models supports the assertion that that specific wavenumber regions reflect distinct organic characteristics of the Karenight core measured using Rock-Eval.
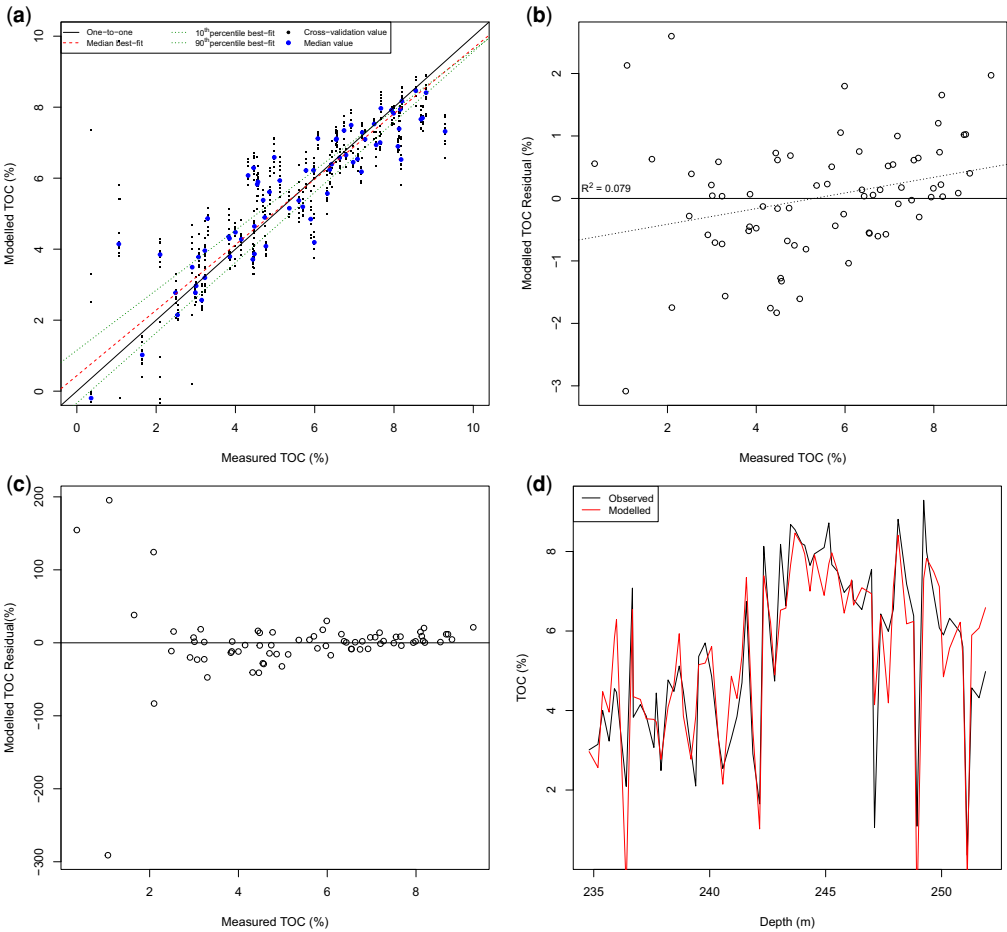
FTIR fingerprint region PCs produced the best model for predicting TOC (Table 3). The median cross-validation values conform well to the

one-to-one line with minimal variation and extreme values (Fig. 7). Lines of best fit for 10th and 90th percentiles fall within a small range but converge slightly at higher TOC, suggesting some heteroscedasticity (Fig. 7a). Analysis of the residuals shows that the model accounts for the variation in the data with no apparent systematic trend and low R^2 (Fig. 7b). The percentage error is quite high for low TOC values (Fig. 7c).

Table 3. *PLSR goodness-of-fit statistics calculated for median cross-validation data*

Dependent variable	Independent variables	MAE	RMSE	R^2	MPAE	MPRSE
All FTIR data	T_{max}	2.98	4.52	0.45	0.68	14.20
	S1	0.22	0.28	0.72	31.37	10.22
	S2	2.53	3.45	0.67	41.94	10.16
	S3	**0.06**	**0.07**	**0.65**	**42.64**	**9.77**
	TOC	0.78	0.96	0.8	22.59	8.78
Non-Fingerprint region FTIR data	**T_{max}**	**3.26**	**4.45**	**0.45**	**0.75**	**15.51**
	S1	**0.20**	**0.26**	**0.77**	**35.57**	**9.20**
	S2	**1.80**	**2.34**	**0.85**	**39.85**	**6.94**
	S3	11.17	12.67	0.02	96.83	43.11
	TOC	0.85	1.06	0.76	30.43	9.53
Fingerprint FTIR data	T_{max}	3.18	4.8	0.4	0.73	15.12
	S1	0.3	0.4	0.43	42.97	14.29
	S2	2.57	3.43	0.68	43.09	9.91
	S3	0.09	0.12	0.12	53.54	15.39
	TOC	**0.71**	**0.96**	**0.81**	**23.10**	**7.98**

Models that meet the goodness-of-fit acceptance criteria shown in bold. MPRSE, mean percentage range standardized error.
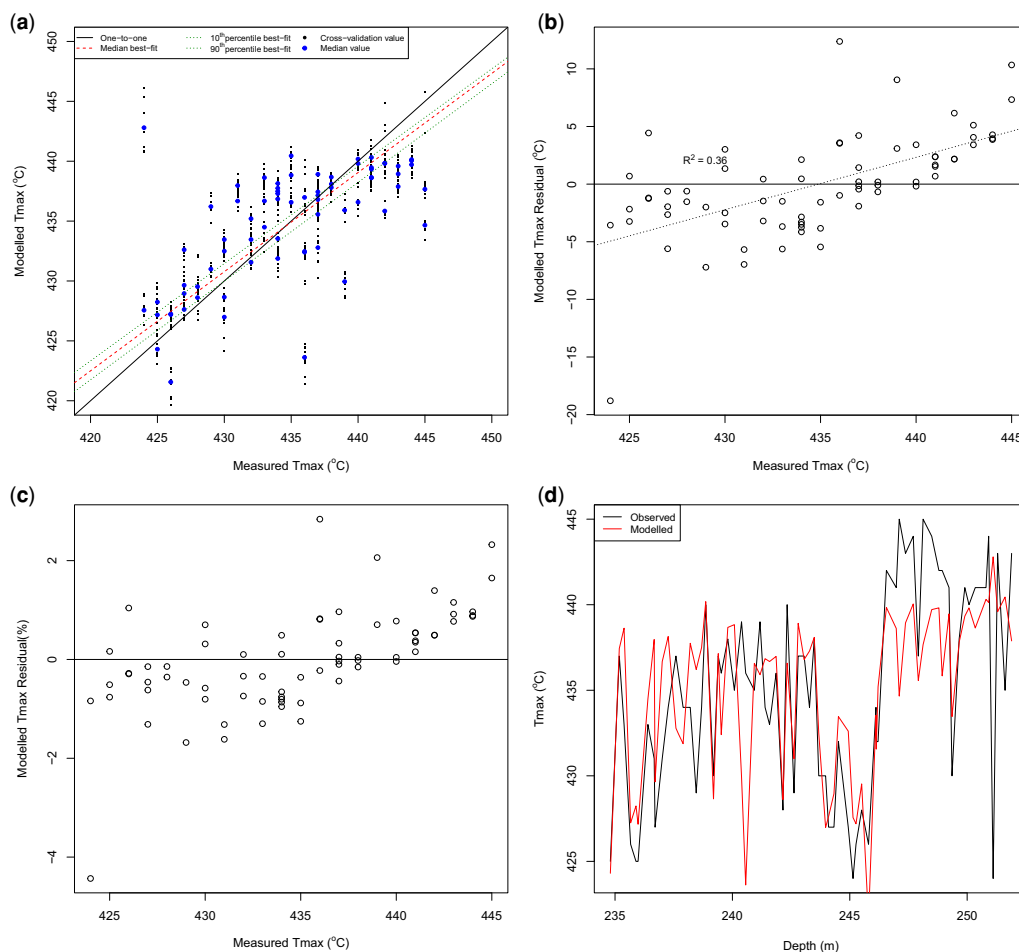

Fig. 7. Preferred T_{max} PLSR model derived from non-fingerprint data showing (**a**) repeated LOO PLSR, (**b**) absolute residuals, (**c**) percentage residuals and (**d**) absolute error.

Non-FTIR fingerprint PCs produced the best model for predicting T_{max} (Table 3). The median cross-validation values broadly conformed to the one-to-one line, notwithstanding a limited number of extreme values (Fig. 8a). The model tends to under predict low T_{max} values and overpredict high values (Fig. 8a, d). Lines of best fit for 10th and 90th percentiles fall within a small range but converge slightly at higher T_{max}, suggesting some heteroscedasticity (Fig. 8a). Repeated cross-validation data ranges presented in the same plot varied widely. Analysis of the residuals shows that the model accounts for the variation in the data with no apparent systematic pattern and low R^2 (Fig. 8b). The percentage error is generally high (Fig. 8c).

Non-fingerprint FTIR PCs produced the best model for predicting S1 (Table 3). The median cross-validation values conformed well to the one-to-one line at the high and low S1 values, (Fig. 9a). There were also a limited number of extreme values. As with TOC, lines of best fit for 10th and 90th percentile data show that most of the cross-validation data fall within a small range. The 10th and 90th percentile lines of best fit converge slightly at higher SI values, indicating possible heteroscedasticity. Repeated cross-validation of the S1 data presented in the same plot showed more variation for low and high values of the data, rather than the mid-range. Analysis of the residuals suggests that the model accounts for the variation in the data with no apparent systematic pattern, confirmed by a low R^2 value for the linear model line of best fit (Fig. 9b). The percentage error is well constrained throughout the data (Fig. 9c).

Non-fingerprint FTIR PCs produced the best model for predicting S2 (Table 3). The median

Fig. 8. Preferred HI PLSR model derived from non-fingerprint data showing (**a**) repeated LOO PLSR, (**b**) absolute residuals, (**c**) percentage residuals and (**d**) absolute error.
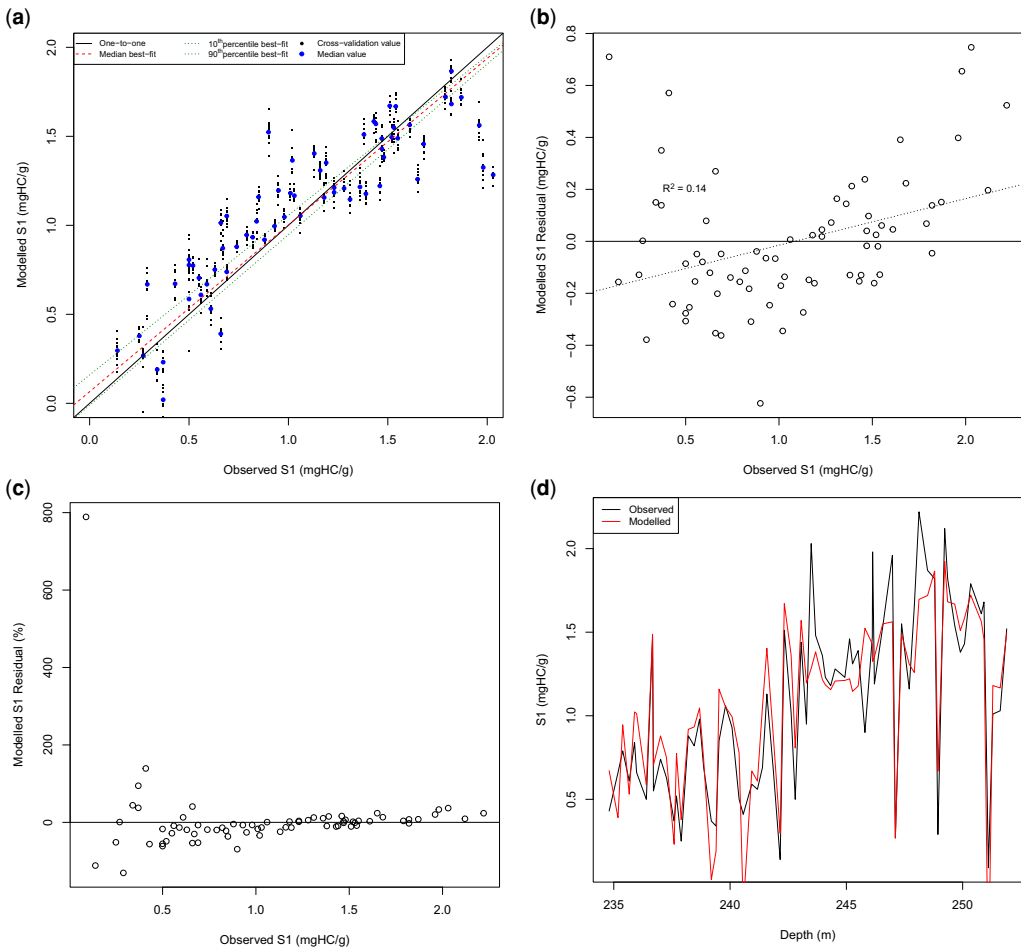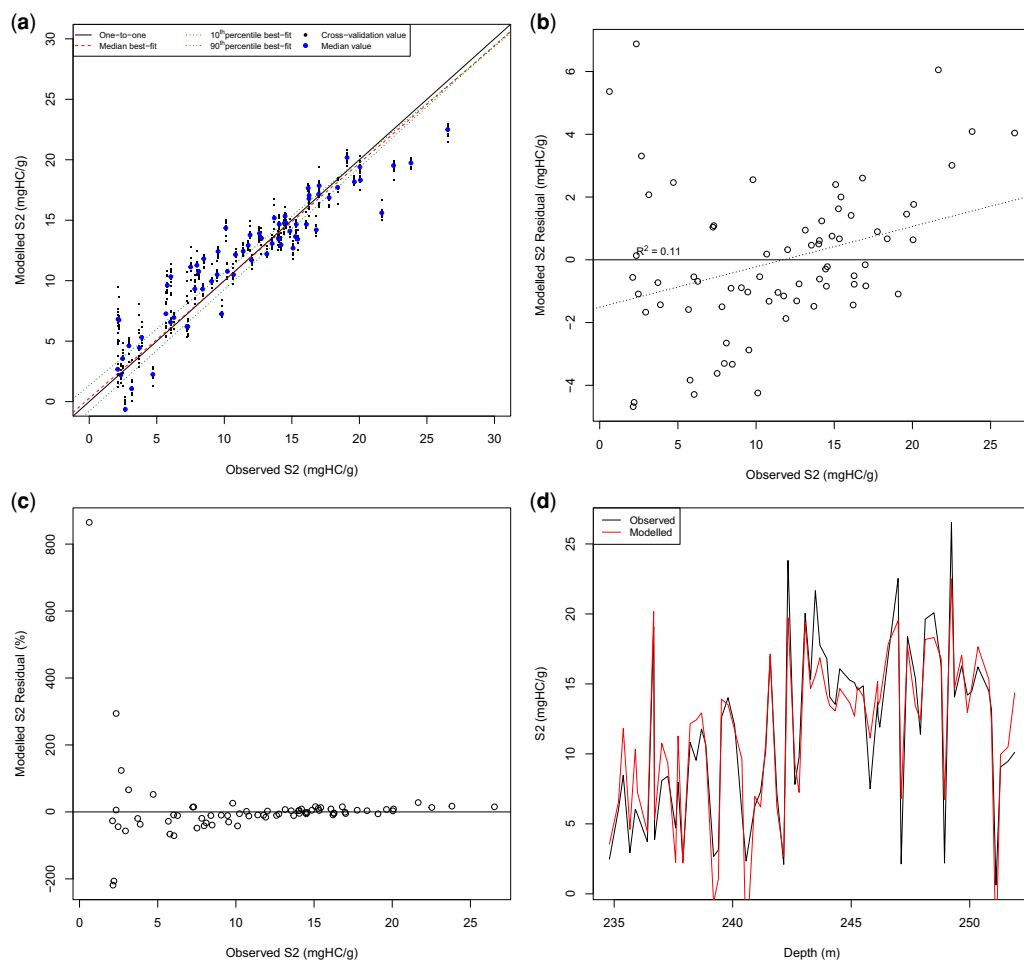
Fig. 9. Preferred S1 PLSR model derived from all data showing (**a**) repeated LOO PLSR, (**b**) absolute residuals, (**c**) percentage residuals and (**d**) absolute error.

cross-validation values conformed very well to the one-to-one line (Fig. 10a). the cross-validation data for S2 has a very small range in comparison to other parameters. The 10th and 90th percentile lines of best fit also converge at higher S2 values, indicating likely heteroscedasticity. These lines are slightly below the one-to-one line at higher S2 values. Repeated cross-validation data revealed minimal variation, suggesting repeatability good throughout the S2 range. Analysis of the residuals suggests that the model accounts for the variation in the data with no apparent systematic pattern and low R^2 (Fig. 10b). The percentage error was high (Fig. 10c).

All-data FTIR PCs produced the best model for predicting S3 (Table 3). The median cross-validation values conformed well to the one-to-one line, although a possible sigmoidal shape is present. The

model over predicts in the mid to upper range of the data (Fig. 11a). As with other Rock-Eval parameters, lines of best fit for 10th and 90th percentile data show that most of the data falls within a small but diverge slightly at higher S3 values, indicating possible heteroscedasticity. Repeated cross-validation data present low variation in comparison of other Rock-Eval parameters, suggesting repeatability is moderate. Analysis of the residuals suggests that the model accounts for the variation is the data but with a slight systematic pattern show also by the R^2 (Fig. 11b). The percentage error was high (Fig. 11c).

Overall, the most accurate models with the lowest cross-validation goodness-of-fit statistics are for S1, S2, S3 and TOC. The percentage error results are all higher than T_{max} despite lower mean absolute errors, root-mean-squared errors and higher R^2. The residuals plots for all Rock-Eval parameters suggest none
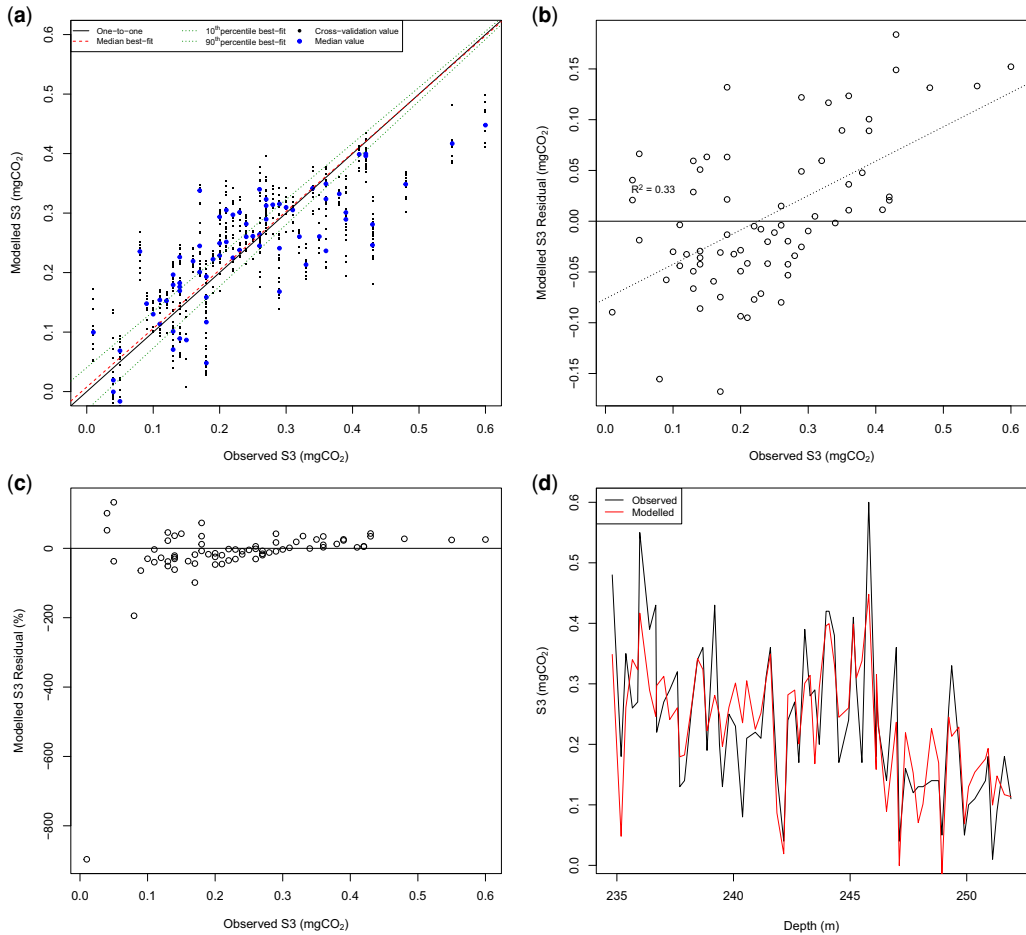
Fig. 10. Preferred S2 PLSR model derived from non-fingerprint data showing (**a**) repeated LOO PLSR, (**b**) absolute residuals, (**c**) percentage residuals and (**d**) absolute error.

of the models are over-fitted although there is an indication of heteroscedasticity in 10th and 90th percentile best fit lines. The absolute error plots show that S1, S2, S3, TOC and HI models conform well to the measured data trend line in comparison to T_{max}.

For this study we created a new goodness-of-fit metric called mean percentage range standardized error (MPRSE) to describe the model residuals as a percentage of the range. We did this because it was difficult to distinguish between goodness-of-fit statistics using mean absolute errors, root-mean-squared errors, R^2 and mean percentage absolute errors (MPAE). The low MPRSE indicates that the modelled values are closer to the measured values than high MPRSE. Using multiple acceptance criteria in environmental modelling is well-documented

(Jakeman *et al.* 2006; Bennett *et al.* 2012) but it is rarely applied rigorously (Beriro *et al.* 2013). For example, (Washburn and Birdwell 2013) report 'excellent models' based on R^2 being >0.9. We argue that although the published models produce high R^2 (Washburn and Birdwell 2013), such a definitive conclusion should be supported using congruent model acceptance criteria. Washburn and Birdwell (2013) show similar percentage error values as this research (*c.* 20%), which we would argue is not as excellent as a high R^2 might suggest. However, the authors acknowledge their sample set is small and not fully representative of the variation of shales present in the UK (Washburn and Birdwell 2013).

MPRSE for the PLSR models show that: (1) T_{max} has the highest values (>10%); and (2) S1, S2, S3

Fig. 11. Preferred S3 PLSR model derived from fingerprint data showing (**a**) repeated LOO PLSR, (**b**) absolute residuals, (**c**) percentage residuals and (**d**) absolute error.

and TOC MPRSE lowest (<10%) (Table 3). These findings are important because Rock-Eval pyrolysis is used as screening tool with selected parameter cut-off points or thresholds are often being applied to the data e.g. $T_{max} \geq 435°C$ is indicative of an oil-prone shale rock based on its thermal maturity. In practice, it is important to ensure that any decisions made using modelled values minimize uncertainties, especially false positives i.e. a modelled T_{max} of 440°C that is actually 410°C, given that this scenario may lead to the assumption that the shale is of sufficient thermal maturity to generate oil and/or oil and gas when it is not. This situation could occur if MPRSE was not used to evaluate the T_{max} or S1 models. Where T_{max} MPAE is 0.75% and MPESR is 15.51%, MPAE appears low but is caused by high absolute values >400°C in comparison to a small range *c.* 50°C (Fig. 14); conversely, S1 has a

high MPAE of 35.57% and a low MPRSE 9.2% (Table 3). When modelling is undertaken to predict one variable only, MPAE can be used to discuss the predictions relative to the magnitude of the parameter unit and its range, but when multiple units are being modelled than MPRSE is more appropriate. We suggest that the models presenting MPRSE *c.* 10% have an acceptable level of uncertainty in the context of the Rock-Eval ranges measured for the Karenight core (Table 3). However, this assumption needs to be considered in the context of the other of lines of evidence including the measured-modelled scatters plots. We suggest that the S1, S2, S3 and TOC models are potentially suitable for predicting down-core values. Modelled T_{max} has a low MPRSE (the cross-validation data in the scatter plots is widely spread) and therefore may not be suitable for making reliable predictions. The

relative ranges used to draw these conclusions are for the Karenight borehole only. Using a larger dataset with ranges that are geographically/geologically constrained to the wider basin(s) may be more useful. Such data were unfortunately not available for this research and may help establish whether this technique is useful to quantify geochemical parameters on a more basin-wide scale.

HI and OI were calculated using modelled TOC, S2 and S3 and compared to calculations made using measured values (Figs 12 & 13). HI correlations do not appear particularly strong whereas OI look reasonable. Further analysis was conducted by incorporating the data into the van Krevelen diagrams.
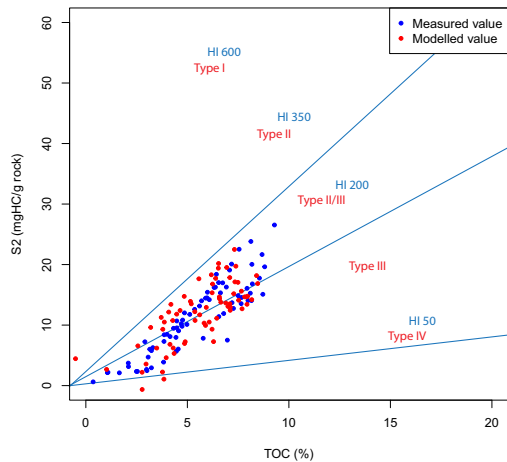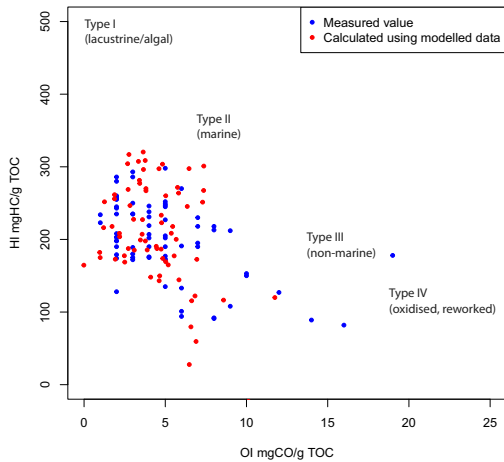
Van Krevelen diagrams and S2 to TOC bi-plots are used as an indication of the maturity of the shale and are also is known be used as crude estimate of kerogen type as well as generative-potential (Słowakiewicz *et al.* 2015; Könitzer *et al.* 2016; Monaghan *et al.* 2017; Waters *et al.* 2020). Comparison of the measured and modelled values suggest a good correspondence and similar clustering (Fig. 14). For example, the majority of modelled values fall within the expected HI range of *c.* 150–300 and OI *c.* 2–7 mg g^{-1} TOC, which in turn elicits a similar kerogen classification to that using measured values. Inspection of the modelled values on the van Krevelen diagram confirms that the modelled HI
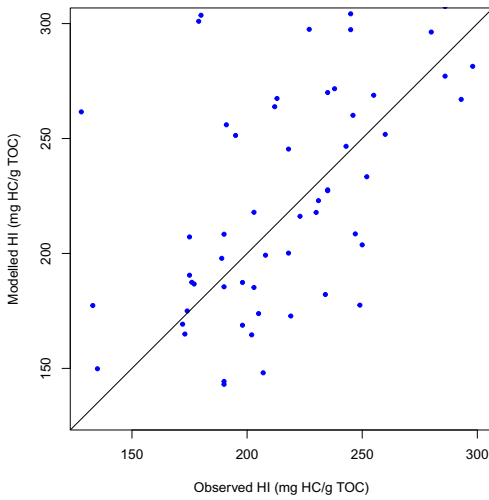
Fig. 12. Comparison between HI calculated using measured and modelled values.
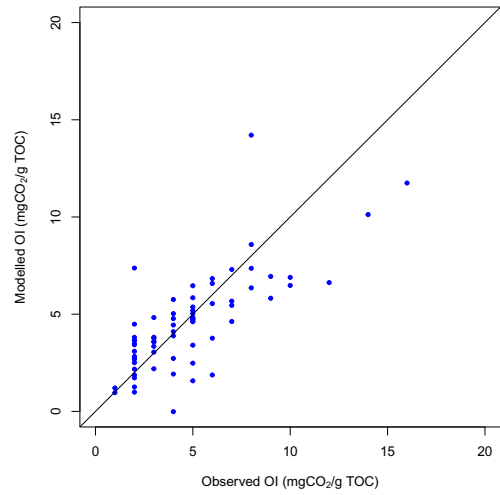
Fig. 13. Comparison between OI calculated using measured and modelled values.

Fig. 14. Van Krevelen diagram and S2 to TOC cross-plot for measured and modelled values.

values are conservative. However, three limestone samples with low measured HI values appear to be underestimated by the FTIR modelling approach. Nevertheless, when both bi-plots are considered together, as would be standard procedure during a prospect appraisal, it is clear the modelled values provide a similar indication of kerogen type and generative potential to that from the actual Rock-Eval pyrolysis derived HI, OI and TOC values.

The shales examined in this current study pose a similar problem in that they are from one core of limited thermal maturity range, kerogen types and palaeogeographic position in the basin. In addition, this research used one chemometric modelling method. There are many contemporary techniques that also produce transparent model outputs, e.g. evolutionary algorithms (Beriro *et al.* 2014), that could form the basis of a comparative study. Doing so is expected enhance the utility of applied chemometrics in understanding shales in an unconventional hydrocarbon context and optimize both the modelling method and model evaluation steps basin, as well as individual core assessments. Overall, this research provides encouraging proof-of-concept results that should be developed further using a range of shales that cover a range of burial conditions and maturity.

Conclusions

This study shows that FTIR analysis of UK shale from the Bowland Shale Formation can be used in conjunction with PCA and PLSR modelling to estimate selected organic characteristics derived using Rock-Eval pyrolysis. PCA explains more about the organic signature of the rock than spectral analysis alone. This is because the FTIR peak resolution is complicated by the presence of the inorganic (mineral) constituents of the rock. More detailed quantification and characterization of the organic fraction of the rock would be informative, but moves the focus of the research away from a rapid, low-cost technique for increasing the resolution of down-core OM characterizations. However, to provide greater confidence in the conclusions drawn here, it may be required in future studies. The PLSR models showing MPESR *c.* 10% and a small range associated with the cross-validation data present an acceptable level of uncertainty (TOC, S1, S2, S3, HI and OI) and are suitable interpolating down-core Rock-Eval parameter values. Conversely, FTIR modelled T_{max} has an MPESR > 10% and a high cross-validation data range is above the acceptable level of uncertainty. Down-core interpolations of selected Rock-Eval parameters could be practically achieved by modelling FTIR data by maintaining standard sample frequencies for Rock-Eval samples while supplementing with higher frequencies for FTIR and chemometric analysis. The approach has the potential to provide greater spatial characterization of the shale rock where high-resolution stratigraphic control is needed on organic carbon type, but economics preclude substantial analytical programmes. The application of the techniques shown here may also have wider value to other geological research pertaining to mudrock successions, including radioactive waste disposal.

Competing interests The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Author contributions DJB: conceptualization (equal), formal analysis (lead), investigation (equal), methodology (equal), project administration (equal), visualization (lead), writing – original draft (lead); **CHV**: conceptualization (equal), formal analysis (equal), project administration (equal), supervision (lead), writing – original draft (equal).

Data availability The datasets generated during and/or analysed during the current study are available from the corresponding author on request.

References

Andrews, I.J. 2013. *The Carboniferous Bowland Shale Gas Study: Geology and Resource Estimation*. British Geological Survey for Department of Energy and Climate Change, London, UK.

Behar, F., Beaumont, V., Penteado, D.B. and L, H. 2001. Rock-Eval 6 technology: performances and developments. *Oil & Gas Science and Technology*, **56**, 111–134, https://doi.org/10.2516/ogst:2001013

Bennett, N.D., Croke, B.F.W. *et al.* 2012. Characterising performance of environmental models. *Environmental Modelling and Software*, **40**, 1–20, https://doi.org/10.1016/j.envsoft.2012.09.011

Beriro, D.J., Abrahart, R.J., Nathanail, C.P., Moreno, J. and Bawazir, A.S. 2013. A typology of different development and testing options for symbolic regression

modelling of measured and calculated datasets. *Environmental Modelling and Software*, **47**, 29–41, https://doi.org/10.1016/j.envsoft.2013.03.020

Beriro, D.J., Abrahart, R.J. and Diplock, G. 2014. *Genetic Programming in GeoComputation*. 2nd edn. CRC Press.

Calderón, F.J., Mikha, M.M., Vigil, M.F., Nielsen, D.C., Benjamin, J.G. and Reeves, J.B., III 2011. Diffuse-reflectance mid-infrared spectral properties of soils under alternative crop rotations in a semi-arid climate. *Communications in Soil Science and Plant Analysis*, **42**, 2143–2159, https://doi.org/10.1080/00103624.2011.596243

Chen, Y., Furmann, A., Mastalerz, M. and Schimmelmann, A. 2014*a*. Quantitative analysis of shales by KBr-FTIR and micro-FTIR. *Fuel*, **116**, 538–549, https://doi.org/10.1016/j.fuel.2013.08.052

Chen, Y., Mastalerz, M. and Schimmelmann, A. 2014*b*. Heterogeneity of shale documented by micro-FTIR and image analysis. *Journal of Microscopy*, **256**, 117–189, https://doi.org/10.1111/jmi.12169

Chen, Y., Zou, C., Mastalerz, M., Hu, S., Gasaway, C. and Tao, X. 2015. Applications of micro-fourier transform infrared spectroscopy (FTIR) in the geological sciences – a Review. *International Journal of Molecular Sciences*, **16**, 30223–30250, https://doi.org/10.3390/ijms161226227

Cornford, C., Gardner, P. and Burgess, C. 1998. Geochemical truths in large data sets. I: geochemical screening data. *Organic Geochemistry*, **29**, 519–530.

Craddock, P.R., Prange, M.D. and Pomerantz, A.E. 2017. Kerogen thermal maturity and content of organic-rich mudrocks determined using stochastic linear regression models applied to diffuse reflectance IR Fourier transform spectroscopy (DRIFTS). *Organic Geochemistry*, **110**, 122–133, https://doi.org/10.1016/j.orggeochem.2017.05.005

Dembicki, H. 2009. Three common source rock evaluation errors made by geologists during propsect or play appraisals. *AAPG Bulletin*, **93**, 341–356, https://doi.org/10.1306/10230808076

El-Rub, Z.A., Kujawa, J., Albarahmieh, E., Al-Rifai, N., Qaimari, F. and Al-Gharabli, S. 2019. High throughput screening and characterization methods of Jordanian oil shale as a case study. *Energies*, **12**, 3148, https://doi.org/10.3390/en12163148

Emmings, J.F., Hennissen, J.A.I. *et al.* 2019. Controls on amorphous organic matter type and sulphurization in a Mississippian black shale. *Review of Palaeobotany and Palynology*, **268**, 1–18, https://doi.org/10.1016/j.revpalbo.2019.04.004

Engelhart, S.E., Horton, B.P. *et al.* 2013. Testing the use of microfossils to reconstruct great earthquakes at Cascadia. *Geology* , **41**, 1067–1070, https://doi.org/10.1130/G34544.1

Espitalie, J., Joubert, L., Kumer, R.K., Dwiedi, P., Banerijie, V. and Gupta, V. 1987. Use of T_{max} as a maturation index in petroleum exploration. *In*: *Petroleum Geochemistry and Exploration in the Afro-Asian Region*. AA Balkema Rotterdam, 67–74.

Fleury, M. and Romero-Sarmiento, M. 2016. Characterization of shales using T_1–T_2 NMR maps. *Journal of Petroleum Science and Engineering*, **137**, 55–62, https://doi.org/10.1016/j.petrol.2015.11.006

Fletcher, T.H., Gillis, R., Adams, J., Hall, T., Mayne, C.L., Solum, M.S. and Pugmire, R.J. 2014. Characterization of macromolecular structure elements from a Green River Oil shale, II. Characterization of pyrolysis products by 13C NMR, GC/MS, and FTIR. *Energy and Fuels*, **28**, 2959–2970, https://doi.org/10.1021/ef500095j

Gerretzen, J., Paul Eilers, P., Wouters, H., Bloemberg, T. and Ron Wehrens, R. 2014. *Parametric Time Warping* version 1.0-7. R.

Hennissen, J.A.I., Hough, E., Vane, C.H., Leng, M.J., Kemp, S.J. and Stephenson, M.H. 2017. The prospectivity of a potential shale gas play: an example from the southern Pennine Basin (central England, UK). *Marine and Petroleum Geology*, **86**, 1047–1066, https://doi.org/10.1016/j.marpetgeo.2017.06.033

Hough, E., Vane, C.H., Smith, N.J.P. and Moss-Hayes, V. 2014. The Bowland Shale Formation from Lancaster fells sub-basin of the Craven basin, UK: a potential shale play. SPE/EAGE European Unconventional Conference Exhibition, Vienna, Austria, Society of Petroleum Engineers, 11.

Jakeman, A.J., Letcher, R.A. and Norton, J.P. 2006. Ten iterative steps in development and evaluation of environmental models. *Environmental Modelling and Software*, **21**, 602–614, https://doi.org/10.1016/j.envsoft.2006.01.004

Jarvie, D.M., Hill, R.J., Ruble, T.E. and Pollastro, R.M. 2007. Unconventional shale-gas systems: the Mississippian Barnett Shale of north-central Texas as one model for thermogenic shale-gas assessment. *AAPG Bulletin*, **91**, 475–499, https://doi.org/10.1306/12190606068

Jarvie, D.M. (ed.). 2012. Shale resource systems for oil and gas: Part 1 – Shale-gas resource systems. *AAPG Memoirs*, **97**, 69–87.

Könitzer, S.F., Stephenson, M.H., Davies, S.J., Vane, C.H. and Leng, M.J. 2016. Significance of sedimentary organic matter input for shale gas generation potential of Mississippian Mudstones, Widmerpool Gulf, UK. *Review of Palaeobotany and Palynology*, **224**, 146–168, https://doi.org/10.1016/j.revpalbo.2015.10.003

Kuhn, M., Wing, J. *et al.* 2015. *Caret: Classification and Regression Training*. R package version 6.0-41, http://CRAN.R-project.org/package=caret

Leach, C.J., Wagner, T., Jones, M., Juggins, S. and Stevenson, A.C. 2008. Rapid determination of total organic carbon concentration in marine sediments using Fourier transform near-infrared spectroscopy (FT-NIRS). *Organic Geochemistry*, **39**, 910–914, https://doi.org/10.1016/j.orggeochem.2008.04.012

Li, W., Stevens, L.A. *et al.* 2021. Comparison of the impact of moisture on methane adsorption and nanoporosity for over mature shales and their kerogens. *International Journal of Coal Geology*, **237**, 103705, https://doi.org/10.1016/j.coal.2021.103705

Lis, G.P., Mastalerz, M., Schimmelmann, A., Lewan, M.D. and Stankiewicz, B.A. 2005. FTIR absorption indices for thermal maturity in comparison with vitrinite reflectance R0 in type-II kerogens from Devonian black shales. *Organic Geochemistry*, **36**, 1533–1552, https://doi.org/10.1016/j.orggeochem.2005.07.001

Marshall, C., Uguna, J. *et al*. 2015. Geochemistry and petrology of palaeocene coals from Spitzbergen – Part 2: maturity variations and implications for local and regional burial models. *International Journal of Coal Geology*, **143**, 1–10, https://doi.org/10.1016/j.coal.2015.03.013

Mas, S., De Juan, A., Tauler, R., Olivieri, A.C. and Escandar, G.M. 2010. Application of chemometric methods to environmental analysis of organic pollutants: a review. *Talanta*, **80**, 1052–1067, https://doi.org/10.1016/j.talanta.2009.09.044

Monaghan, A.A., Arsenikos, S. *et al*. 2017. Carboniferous petroleum systems around the Mid North Sea High, UK. *Marine and Petroleum Geology*, **88**, 282–302, https://doi.org/10.1016/j.marpetgeo.2017.08.019

Mostert, M.M.R., Ayoko, G.A. and Kokot, S. 2010. Application of chemometrics to analysis of soil pollutants. *TrAC - Trends in Analytical Chemistry*, **29**, 430–445, https://doi.org/10.1016/j.trac.2010.02.009

Newell, A.J., Vane, C.H., Sorensen, J.P.R., Moss-Hayes, V. and Gooddy, D.C. 2016. Long-term Holocene groundwater fluctuations in a chalk catchment: evidence from Rock-Eval pyrolysis of riparian peats. *Hydrological Processes*, **30**, 4556–4567, https://doi.org/10.1002/hyp.10903

Peters, K.E., Walters, C.C. and Moldowan, J.M. 2005. *The Biomarker Guide*. Vol. **1**. Cambridge University Press, Cambridge.

R Development Core Team. 2016. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, http://www.r-project.org/

Scotchman, I.C. 2014. Exploration for unconventional hydrocarbons: shale gas and shale oil. *Issues in Environmental Science and Technology*, **39**, 69–103, https://doi.org/10.1039/9781782620556-00069

Słowakiewicz, M., Tucker, M.E., Vane, C.H., Harding, R., Collins, A. and Pancost, R.D. 2015. Shale-gas potential of the mid-Carboniferous Bowland-Hodder unit in the Cleveland Basin (Yorkshire), central Britain. *Journal of Petroleum Geology*, **38**, 59–75, https://doi.org/10.1111/jpg.12598

Tinti, A., Tugnoli, V., Bonora, S. and Francioso, O. 2015. Recent applications of vibrational mid-infrared (IR) spectroscopy for studying soil components: a review. *Journal of Central European Agriculture*, **16**, 1–22, https://doi.org/10.5513/JCEA01/16.1.1535

Vane, C.H., Drage, T.C. and Snape, C.E. 2003. Biodegradation of Oak (*Quercus alba*) wood during growth of the shiitake mushroom (*Lentinula edodes*): a molecular approach. *Journal of Agricultural and Food Chemistry*, **51**, 947–956, https://doi.org/10.1021/jf020932h

Varma, A.K., Mishra, D.K., Samad, S.K., Prasad, A.K., Panigrahi, D.C., Mendhe, V.A. and Singh, B.D. 2018. Geochemical and organo-petrographic characterization for hydrocarbon generation from Barakar Formation in Auranga Basin, India. *International Journal of Coal Geology*, **186**, 97–114, https://doi.org/10.1016/j.coal.2017.12.002

Washburn, K.E. and Birdwell, J.E. 2013. Multivariate analysis of ATR-FTIR spectra for assessment of oil shale organic geochemical properties. *Organic Geochemistry*, **63**, 1–7, https://doi.org/10.1016/j.orggeochem.2013.07.007

Waters, C.N., Browne, M.A.E., Dean, M.T. and Powell, J.H. 2009. *Lithostratigraphical Framework for Carboniferous Successions of Great Britain (Onshore)*. British Geological Survey Research Report **RR/07/01**.

Waters, C.N., Vane, C.H., Kemp, S.J., Haslam, R.B., Hough, E. and Moss-Hayes, V.L. 2020. Lithological and chemostratigraphic discrimination of facies within the Bowland Shale Formation within the Craven and Edale basins, UK. *Petroleum Geoscience*, **26**, 325–345, https://doi.org/10.1144/petgeo2018-039

Whitelaw, P., Uguna, C.N. *et al*. 2019. Shale gas reserve evaluation by laboratory pyrolysis and gas holding capacity consistent with field data. *Nature Communications*, **10**, 3659, https://doi.org/10.1038/s41467-019-11653-4

Wilson, A.A. and Stevenson, I.P. 1973. *Institute of Geological Sciences Record of Shaft or Borehole*. British Geological Survey, Nottingham, UK.