Article

Identification and Discrimination of Petrol Sources by Nuclear Magnetic Resonance Spectroscopy and Machine Learning in Fire Debris Analysis

Yanita Yankova [1], Silvia Cirstea [2,*], Michael Cole [3] and John Warren [4]

1   Eurofins Forensic Services, 1 Dukes Green Avenue, Feltham TW14 0LR, UK;
    yanita.yankova@forensicsuk.eurofins.com
2   School of Computing and Information Sciences, Anglia Ruskin University, East Road,
    Cambridge CB1 1PT, UK
3   School of Life Sciences, Anglia Ruskin University, East Road, Cambridge CB1 1PT, UK;
    michael.cole@aru.ac.uk
4   Jazz Pharma, Unit 840 Broadoak Rd, Sittingbourne ME9 8AG, UK; john.warren@jazzpharma.com
*   Correspondence: silvia.cirstea@aru.ac.uk

Featured Application: Machine learning in forensic science and the use of MATLAB for identification and classification of petrol source.

Abstract: Petrol is considered the most common fire accelerant. However, the identification and classification of petrol sources through the years has proven to be a challenging field in the investigation of fire debris analysis. This research explored the possibility of identifying petrol sources by high-field NMR methods accompanied by ML (machine learning). The automated identification and classification of petrol brands were achieved for first time based on the ML classification model developed in this research. A hierarchical classification model was constructed using local classifiers to categorize neat or weathered petrol into its sources.

Keywords: machine learning; petrol; fire investigation; NMR; MATLAB

1. Introduction

Fire investigation is considered one of the most challenging forensic science disciplines. Current gas chromatographic and spectroscopic analytical methods in fire investigation cannot discriminate or individualize petrol sources based on class compounds within the petrol samples from the fire scene to give an indication of the country of origin, refinery (source), natural weathering/age or fire exposure. As petrol is considered one of the most common petroleum products (PPs) used as an ignitable liquids (ILs) in fire investigation, it was the primary concern of this study [1]. The characterization and identification of petrol samples is a crucial challenge in the scientific investigation of fire as the current reference data relating to petrol do not highlight the broad range of different petrol compositions [2]. Identifying individual compounds in the petrol contributes to understanding complex petrol chemical compositions and their additive and blending agents as the refineries do not reveal the exact composition of their petrol. Those compounds have not been previously identified due to their volatility and trace amounts in the petrol mixture. Gas Chromatography (GC)–Mass Spectroscopy (MS) analysis of ILs using chemometric analysis for comparisons of unevaporated, evaporated and "on substrate" petrol samples from stations across the UK displayed very similar chromatographic patterns regardless of the petrol grade or type; hence, discrimination by grade, type, or brand could be very challenging [3]. The authors used Principal Component Analysis (PCA) analysis to target C_2–C_4 alkyl benzenes; the PCA achieved a grouping of petrol brands based on their grade (premium and regular). A Hierarchical Cluster Analysis (HCA) was applied to the data

and no substantial clustering based on petrol type or brand was revealed. However, the HCA dendrogram demonstrated a linkage of the samples according to their degree of evaporation [3].

A method based on Gas Chromatography (GC)–Flame Ionization Detector (FID) analysis combined with an ANN (artificial neural network) algorithm was explored for the discrimination of petrol brands from five petrol stations in Spain based on the entire chromatogram [2]. It was concluded that despite there not being significant variations in the chromatogram, mathematically, the different petrol samples could be classified according to their brand. The authors suggested that the potential difference that contributed to the discrimination was the content of oxygenate and hydrocarbon groups such as aromatics and olefins. In that experiment, native petrol samples were only considered for identification purposes and no identification of specific compounds was made [2].

Research by Monfreda and Gregori [4] offered promising results where unevaporated samples from different petrol sources were correctly grouped based on aromatic compounds. In addition, Barrett et al. [5] used Direct Analysis Real-Time Mass Spectroscopy (DART-MS) combined with the Partial Least Squares Discriminant Analysis (PLS-DA) model to classify petrol sources on different substrates; however, the petrol samples were grouped by the already identified class rather than unknown classes.

Even though many spectroscopic and chromatography techniques have been considered, it can be concluded that the identity of the source of ILs recovered from a fire scene is still a challenging and ongoing research area. Therefore, there is a need for the individualization and classification of petrol sources to enhance ILs' evidential value.

Nuclear Magnetic Resonance (NMR) is a non-destructive spectroscopic technique that studies the nuclei of atoms within a molecule and their chemical environment. NMR spectroscopy is sufficient to completely determine the structure of an unknown molecule and to differentiate between isomers or related compounds which can be difficult using GC-MS. Various NMR pulse sequences allow complex spectra to be dissected by focusing on individual small spectral regions and extracting the spectra of the coupled spin systems that have a resonance within that region, even when their spectra severely overlap. Therefore, NMR spectroscopy has the capability to extract the sub-spectra of an individual component without prior separation from highly complex spectra [6].

A simple ^{1}H NMR method has been successful in the determination of petrol composition and some individual compounds with rapid and accurate results. Further investigation of NMR applications in the petroleum industry displayed the capabilities of ^{1}H NMR coupled with PCA, k-NN (k-Nearest Neighbors), HCA and SIMCA (Soft Independent Modeling of Class Analogy), which proved to be a useful tool for categorizing petrol samples with adulteration (solvents), fuel additives and blends, petroleum mixtures (kerosene and diesel mixtures) and petrol samples with different octane numbers; NMR is a proven analytical tool for the identification and quantification of low-molecular-weight components [7–11]. The primary application of high-field NMR spectroscopy in the petroleum industry is in the quality control of hydrocarbon classes in a sample rather than individual compounds of the overly crowded complex spectra. ^{1}H NMR methods coupled with clustering and multivariate classification techniques were used for the successful identification of adulteration in two types of samples. The potential of NMR spectroscopy for the structural elucidation of petrol components in a sample has been established.

Considering the application of NMR in various scientific fields, forensic NMR is still in the early stages of development with a particular focus on the chemical compositions of single compounds. A ^{1}H NMR method was combined with statistical analysis to identify the chemical "fingerprint" of cocaine samples and to link cocaine samples based on this information. It was concluded that the NMR method could establish a link between seized samples obtained at different locations or in possession of different individuals. The relative ratios of the minor components in coca leaves are closely associated with plant varietal, cultivar and agronomic differences that can be exploited for the assignment of geographical origin, at least when suitable authentic databases are available [12]. One of the

disadvantages of ^{1}H NMR is that it is generally used for nonselective analysis compared to the selectivity of MS. Peak overlaps from multiple detected compounds pose major challenges in the complex ^{1}H NMR spectrum of petrol. Therefore, band-selective sequences including selective (sel) TOCSY and pure shifts that use tailored pulses, which narrow the excitation bandwidth to the region of interest in a signal measurement to obtain information for a single spin system, are recommended.

Machine learning has been proven to be beneficial in forensic science in its various fields such as public safety, image and video analysis, image recognition, gunshot detection, firearms identification, 3D crime scene reconstruction, huge digital data analysis, building statistical evidence, handwriting identification, time since death estimation, dental age estimation and personal identification through dental findings [13], sex determination of skeletal remains, 3D facial reconstruction from an unidentified skull, cybercrimes and digital evidence detection [14], bloodstain pattern analysis [15] and pattern recognition, which involves pattern evidence such as bite marks, lip prints, bullet marks, tool marks, shoe prints and fingerprint comparison and identification with more accuracy and ultimately higher speeds than human experts [16,17]. NMR provides rapid and accurate data collection in high output forensic laboratories. The objective of this work, using high-field (600 Hz) NMR spectroscopy that delivers automated sample changing, was to uniquely individualize and discriminate aliquot petrol sources based on (1) source (origin of the crude oil); (2) refinery processes and procedures (blending agents); and (3) brand (additive package). Within forensic science, the identification and classification of petrol sources could help police forces in the investigation of various fuel offenses, including arson, motor vehicle incidents, environment spillage, fuel smuggling and petrol bomb-related incidents. Therefore, the objective of the study also included individualization and discrimination of evaporated and Ignitable Liquid Residue (ILR) samples (fire debris residues) to characterize the petrol sample collected at a fire scene. Petrol samples that are not hermetically sealed or are exposed to high temperatures undergo evaporation losses, which is called weathering. Therefore, evaporated samples and fire simulated debris samples are considered weathered. Several NMR pulse sequences were evaluated for this research. The one-dimensional (1D) ^{1}H selTOCSY (Selective Total Correlation Spectroscopy) method (also known as selective HOHAHA (Homonuclear Hartmann–Hahn)) was established as the most suitable pulse sequence for the evaluation of petrol sources, establishing ^{1}H connectivity through J-coupling, which allows for selective proton nuclei excitation by a shaped pulse to enable a response within the entire spin system [18]. In a mixture, an excitation of a range of similar signals and a specific chemical shift occurs depending on the shape of the selective pulse, which is calculated based on the width of the integral region. [1]H selTOCSY creates a correlation between all protons in a spin system. A ^{1}H selTOCSY experiment is an indispensable tool for unravelling the spectra of complicated molecules. This is achieved by the multistep transfer of magnetization over many spins. Increasing the isotropic mixing time causes the net magnetic polarization to spread through an increasing number of bonds [19]. For fire debris samples, a ^{1}H NMR NOESY (Nuclear Overhauser Effect) method with solvent suppression was used to improve digitization without compromising the solute as the sample contained a diluted extracted of ILs in protonated solvent which gave an intense signal; this was then followed by ^{1}H selTOCSY. The solvent suppression method is mainly used to attenuate multiplets by employing shaped pulses, which have a broader excitation profile. The merits of this technique include its easy application, easy implementation within most NMR experiments and possibility of multiple pre-saturation. However, its application may lead to the absence of 2D cross peaks due to the saturation of peaks with resonances close to the solvent frequency. This technique sometimes leads to the transfer of saturation to slowly exchanging protons, which could be detected without saturation [20].

This study developed an automated classification model to individualize and classify unknown native and fire debris petrol samples based on class characteristics of their source by using machine learning.

An automated hierarchical model for classification using local classifiers for each leaf used for the predication of the petrol source is described in this paper and the experimental results and limitations of this model are discussed. The key contributions of this study are as follows: (1) an automated classification model was developed that can successfully classify petrol sources; (2) the machine learning and statistical analysis results can support opinion-based decision making when identifying petrol samples in fire debris analysis; and (3) a new dataset of different petrol sources from UK and Ireland was created.

2. Materials and Methods

The main steps of this study's methodology included the NMR analysis of petrol, data acquisition, feature selection, design, training, optimization and evaluation of the classification model.

2.1. Materials

This study used 58 petrol samples that represented British Petroleum (mainland (M) and Scotland (S)), Jet, Esso, Texaco and Shell sources across petrol stations in the UK and Ireland. To address the issues associated with evaporation and matrix interference, the experimental protocol was followed to analyze (1) evaporated petrol samples (as per the laboratory protocol described below) and (2) simulated fire debris petrol samples burned to 50% of the original weight. For each petrol brand collected, a set of three evaporated samples was generated. In a dry bath at approximately 25 °C (room temperature), 10 mL of neat petrol samples from various petrol sources in triplicate were pipetted into 15 mL plastic tubes and placed under a nitrogen stream until the evaporation percentages were approximately 25%, 50%, 75% and 90% corresponding to volume reductions of 2.5 mL, 5.0 mL, 7.5 mL and 9.0 mL, respectively. During the evaporation process, 100 microliters of the petrol samples was collected at each stage of the volume reduction. The samples were prepared for analysis by diluting them in non-deuterated cyclohexane. Finally, petrol sources (2 mL) were burned up to 50% of their original weight on their own and on a 3 cm × 3 cm substrate (flooring material, carpets, fabrics and paper materials) and subsequently extracted by immersing the substrate with 10 mL cyclohexane. Cyclohexane is a solvent with good miscibility properties. It represents a single resonance peak in the aliphatic area, which does not interfere with the area of interest using the solvent suppression pulse sequence (NOESY). The resolution of the trace amount peaks in the olefin area and the chemical shift variance met the acceptance criteria. The cyclohexane resulted in a minimum chemical shift variance of <0.01 ppm, and showed consistency in different solute strengths (10%, 50% and 90%) with 16 scans. Based on the consistent chemical shift, peak resolution and minimal coupling overlap, it was concluded that non-deuterated cyclohexane was the most appropriate solvent. In addition, non-deuterated cyclohexane proved to be suitable for the direct solvent substrate extraction of fire debris required for this study compared with the deuterated cyclohexane with its low volume availability and high cost.

To impartially compare the NMR method for the discrimination of neat, weathered and burnt petrol samples to the current laboratory method using Automated Thermal Desorption (ATD)–Gas Chromatography–Mass Spectroscopy (GCMS) (in-house developed method used by Eurofins Forensic Services) in the analysis of ILs and their residues for the interpretation of volatile compounds and ignitable liquids, a set of neat, evaporated, burnt and fire debris samples was created. Different petrol samples were prepared by an independent laboratory examiner/analyst; the samples prepared included different brands of neat petrol samples, evaporated petrol samples, petrol samples burned on their own, and petrol samples burned on different substrates. The corresponding burned on substrate samples were collected and packed into a control nylon bag for further extraction (Table 1).

Table 1. Summary table of the blind sets of (1) neat petrol samples and (2) combinations of evaporated, burnt and fire debris petrol samples.

(1) Summary table of the neat petrol samples used for the double-blind study.		
Blind Sample Name	**Class**	
BLIND A	Jet	
BLIND B	Esso I (from regions)	different
BLIND C	Esso II	
BLIND D	Esso III	
BLIND E	Texaco I	
BLIND F	Texaco II	
BLIND G	Shell I	
BLIND H	BP M	
BLIND I	Shell II	
BLIND J	BP S	
(2) Summary table of the evaporated, burnt, and burned on substrate petrol samples used for double-blind study.		
Blind Exhibit Name	**CLASS**	**Weathered Status**
BLIND EXHIBIT A	BPM	Evaporated 50%
BLIND EXHIBIT B	BPM	Cardboard Substrate
BLIND EXHIBIT C	JET	Burnt
BLIND EXHIBIT D	JET	Evaporated 25%
BLIND EXHIBIT E	ESSO	Evaporated 25%
BLIND EXHIBIT F	SHELL	Cardboard Substrate
BLIND EXHIBIT G	SHELL	Burnt
BLIND EXHIBIT H	TEXACO	Burnt
BLIND EXHIBIT I	TEXACO	Evaporated 25%

2.2. Data Acquisition

The data in this study was acquired using a Bruker high-field 600 MHz NMR spectrometer (Bruker, London, UK) with a 5 mm broadband inverse diameter probe. The Icon NMR software (Version 3.0) was used to set the NMR experiments and control the data acquisition. The NMR experiment was a simple single-pulse sequence (zg30 from the Bruker library) for (1) neat petrol, and (2) a second dataset was acquired in cyclohexane with a solvent suppression pulse sequence (NOESY) for the evaporated (due to limited volume) and burnt petrol samples. A pulse sequence program (seldigpzs from the Bruker library) was used for the acquisition of ^{1}H sel (selective) TOCSY. Data were collected with 64 k points as the size of the free induction decay (fid), with a spectral width of 20.0 ppm, a mixing time of 0.06 s, an acquisition time of 2.7 s, a pre-scan delay of 6.5 s and a minimum of 16 scans for the neat petrol samples. The acquisition parameters were based on the default pulse sequences in the Bruker library. ^{1}H selTOCSY was performed on the following bands of chemical shifts: 4.65–4.72 ppm (olefin set 1), 4.73–4.85 ppm (olefin set 2), 4.95–5.10 ppm (olefin set 3) and 5.10–5.35 ppm (olefin set 4). The couplings were resolved and used for the assignment of the chemical species. The four discriminative sets of olefins were identified as 3-methyl-1-butene by irradiating the signal at 4.64–4.72 ppm, a mixture of 3-methyl-1-butene and 1-pentene by irradiating the signal at 4.73–4.85 ppm, 2-methyl-2-butene by irradiating the signal at 4.95–5.10 ppm and a mixture of cis- and trans-2-pentene by irradiating the signal at 5.10–5.35 ppm. For the double-blind study, the exhibits were analyzed using headspace-ATD-GC-MS using a Tenax TA sorbent sampling

tube. A 1 mL headspace was taken from within the packaging after a period of incubation at around 100 °C. The interpretation of the results was based on pattern recognition and comparing the chromatography results obtained from evidential items with the standard references. Where possible, comparison against a reference of the relevant liquid was preferable, but if not possible, the sample was compared to the laboratory reference database or the published literature.

2.3. Data Pre-Treatment and Pre-Processing

The ^{1}H NMR spectrum of petrol is a complex mixture consisting of multiple detectable and overlapping peaks. The position, intensity and width of the peaks of interest significantly impact the quality of the NMR spectrum and its subsequent interpretation. The acquired ^{1}H NMR and ^{1}H TOCSY data were processed with Mestre Nova (version 10.1.0 LITE-SE) software, where different processing parameters were applied to achieve the most efficient dataset. The processing included (1) chemical referencing, (2) phasing, (3) baseline correction, (4) sub-spectral selection and filtering, (5) normalization and (6) binning (Figure 1). The detailed methodology is included in Supplementary Materials.
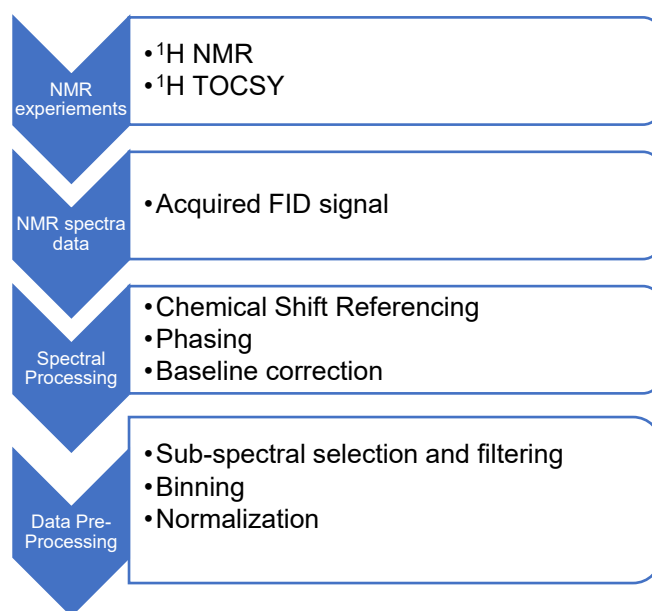

Figure 1. A summary of spectra processing and data pre-processing steps required for optimizing the NMR data of petrol for further machine learning classification.

2.4. Feature Selection

The datasets underwent unsupervised machine learning by applying PCA and a supervised analysis by applying PLS-DA in MetaboAnalyst. PCA was chosen as the explorative tool of the pre-processed data to display any natural groupings. The score plots were a visual representation of the clustering between groups. A loading plot displays how strongly each characteristic influences a principal component. Therefore, PLS-DA was then used for the classification and feature selection of the variables, using cross-validation to select an optimal number of components for classification. The bins that contained important variable information for classification were identified by PLS-DA through the Variation Importance Projection (VIP) score. The VIP score is a measure of a variable's importance in the PLS-DA model [21]. It summarizes the contribution a variable makes to the model. The VIP score of a variable is calculated as a weighted sum of the squared correlations between the PLS-DA components and the original variable. A statistical analysis was performed exploiting the real-time interactive web-based application MetaboAnalyst. Firstly, the non-targeting approach, considering all the spectral information, was explored for classification purposes, and then the targeting approach was used, where the four sets

of olefins were evaluated to achieve better clustering. The dataset from the ^{1}H selTOCSY spectral data was used, which edited out many NMR peaks by filtering out all signals that did not have a component of their spin system in the selective excitation.

2.5. Classification Model

For the first time, this study used machine learning techniques to automatically individualize and classify the petrol sources of native, evaporated, and fire debris samples. They were implemented in MATLAB (R2019b) using the Classification Learner app. The evaluation of the model was performed using selected datasets, which are essential in experimental model development.

The classification model research design is outlined as follows:

Step 1—data collection. The datasets evaluated in this research were as follows: (1) entire ^{1}H NMR spectrum of neat petrol samples; (2) ^{1}H selTOCSY spectrum of the four olefins of neat petrol samples; (3) ^{1}H selTOCSY spectrum of the four olefins of neat and evaporated petrol samples; (4) ^{1}H selTOCSY spectrum of the four olefins of neat, evaporated, and fire debris residue samples.

The datasets were divided into (i) training data, comprising non-targeting (contained all the NMR spectrum information) and targeting (^{1}H selTOCSY spectrum that consisted of selected features that were recognized as an important feature for discrimination purposes) datasets and (ii) a blind study testing dataset (for the practical validation of the model by comparing to a real-world dataset). The training and testing datasets were used to determine the best classifier model for the classification of petrol brands based on the NMR spectrum.

Step 2—reduction of data dimensionality by selecting only a subset of measured features (predictor variables) to create a cluster model through PCA or for the feature selection function (using the featured chemical bins from the PLS-DA VIP scores). The PCA function was enabled with a component reduction criterion of 80% of the explained variance as this represents a sufficient information variance; typically, the first few PCs correspond to cumulative eigenvalues accounting for 80% or above of the variation within a dataset and are sufficient to describe or explain most of the variability in the given dataset, thus reducing the dimensionality [3]. For optimal results, the study aimed to choose a classifier model with a minimum of 60% accuracy (validation).

Step 3—dataset optimization. The effect of the pre-processed parameters on the classification training model for the discrimination of petrol samples was tested. Two different parameter pre-processing methods were investigated in this study: filtering of the redundant spectral bins and the normalization parameter. Data filtering was applied to set any spectral bin value less than 1 to 0. For normalization, (i) single peaks were normalized (to the highest peak of the spectrum) and (ii) normalized with the total area sum (LOG function).

Step 4—dataset splitting. The datasets were split into training and testing datasets using the cross-validation function with K-folds. The cross-validation method with 5 and 10 folds was investigated.

Step 5—evaluating different classifier models such as Decision Trees, Discriminant Analysis (DA), Support Vector Machines (SVMs), Logistic Regression, k-Nearest Neighbor (k-NN), Naïve Bayes, Ensembles, and Artificial Neural Networks (ANNs).

Step 6—after training multiple models, their performances were compared, and then the most robust and effective classification model was chosen. The Classification Learner app displayed the results of the validated model. Performance measures, such as model accuracy, and visual representation plots, such as the confusion matrix chart, reflect the validated model results. The confusion matrix table displayed six petrol brands as true classes in the rows and predicted classes in the columns.

The goal of the classification model method was to investigate different datasets of native, evaporated and fire debris petrol samples with different pre-treatment techniques

including data filtering and normalization to identify the most desirable classifier which provides the highest classification accuracy.

3. Results

3.1. NMR Results

The typical ^{1}H NMR spectrum of the petrol samples contained signals in the chemical shift region of 0 ppm to 8 ppm (Figure 2). The ^{1}H-NMR chemical shifts can be grouped into several broadly defined regions: alkylates have a chemical shift range of 0.5–1.5 ppm; normal and branched iso-paraffins (alkanes) and oxygenates have a chemical shift range of 2.0–4.0 ppm for alcohols, chemical shift range of 2.0–2.2 ppm for (HCOO-CR) esters and for benzylic compounds (Ar-CH), chemical shift range of 3.7–4.1 ppm for esters (RCOO-CH), and chemical shift range of 3.3–4.0 ppm for ethers; olefins have a chemical shift range of 4.0–6.0 ppm; and aromatics have a chemical shift range of 6.5–7.5 ppm (benzene, toluene, ethylbenzene and o-, m- and p-xylene). The ^{1}H NMR profiling of various chemical shift (∂) regions clearly showed that each petrol source had a diagnostic "fingerprint" with specific chemical markers that could be potentially used for identification, classification and ultimately linking an unknown sample to its source. The primary region of interest in the ^{1}H NMR spectrum was the area that represented the additives and blending agents added to the base petrol like olefins and oxygenates produced during the refinery procedure with chemical shifts (∂) of 4.00–6.00 ppm (Figure 3).
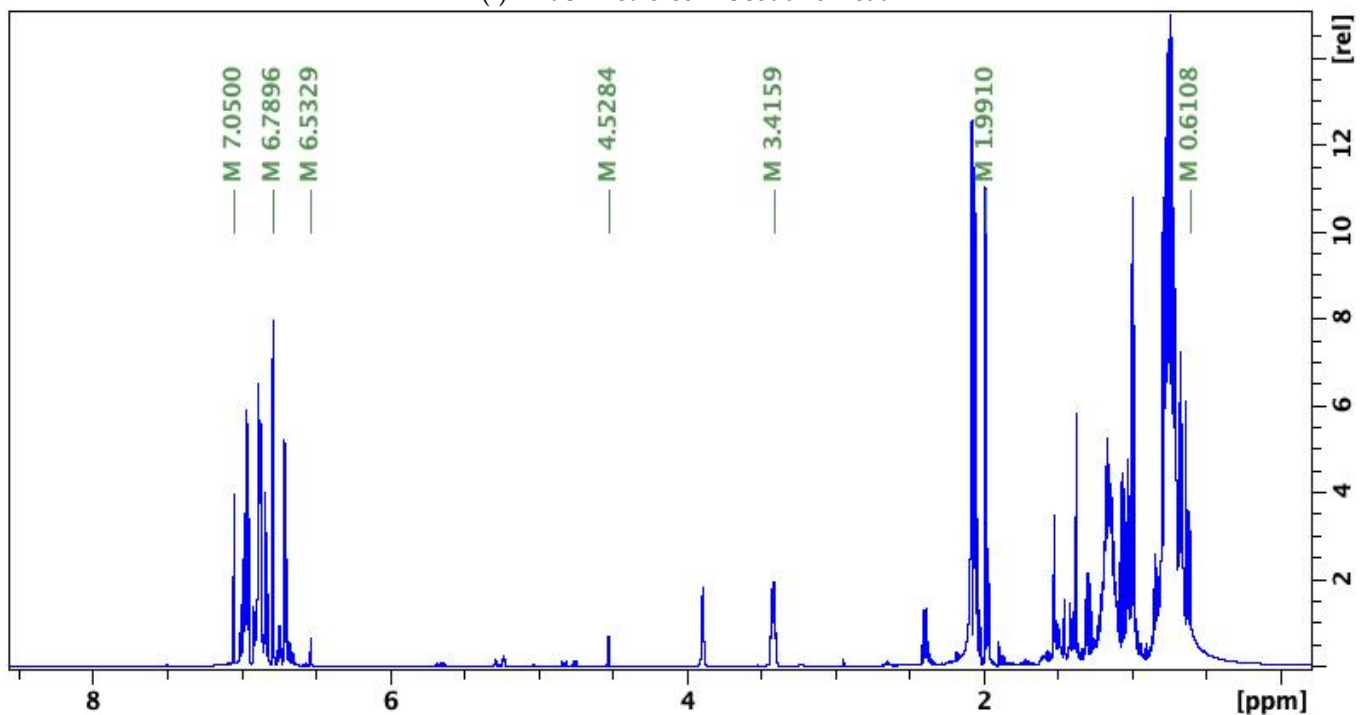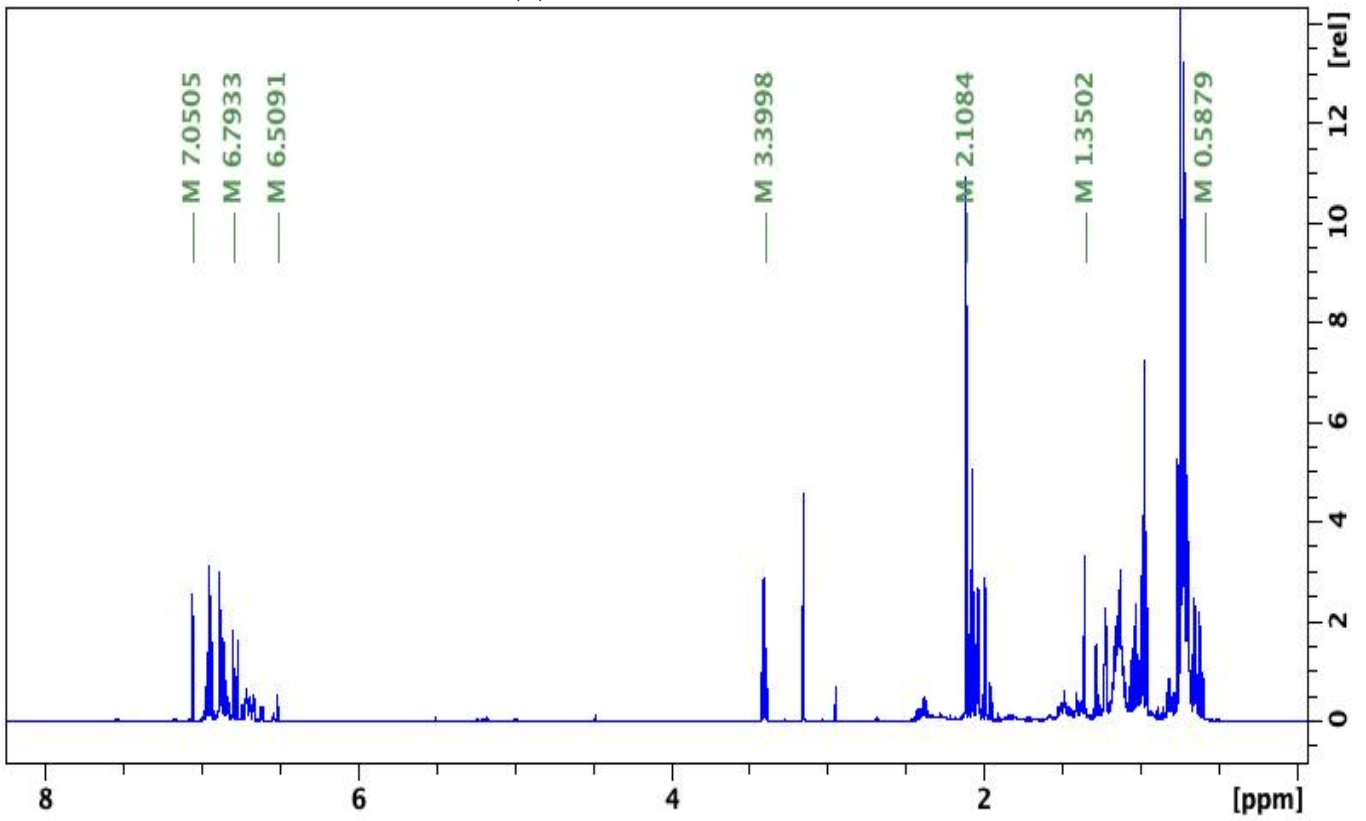
(**i**) British Petroleum Scotland neat

Figure 2. *Cont.*

(**ii**) British Petroleum M neat
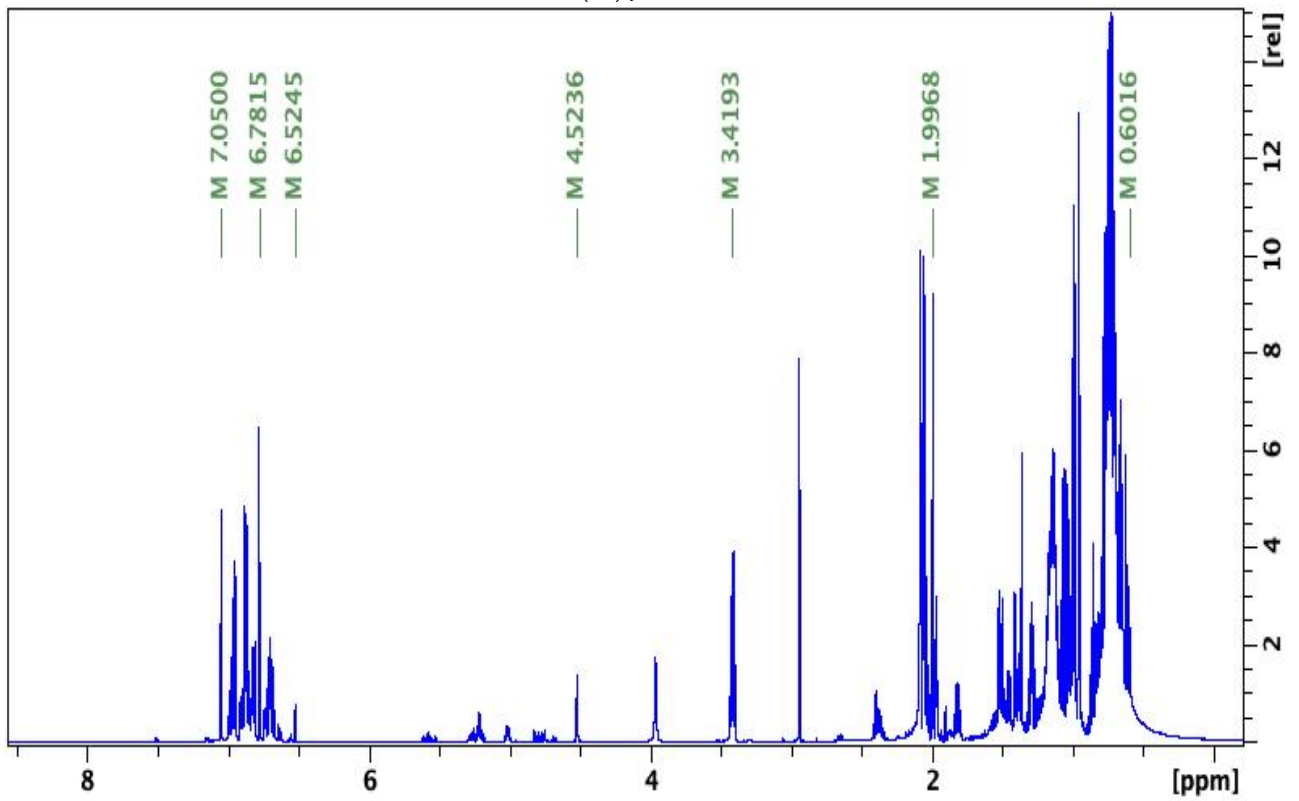

(**iii**) Jet neat
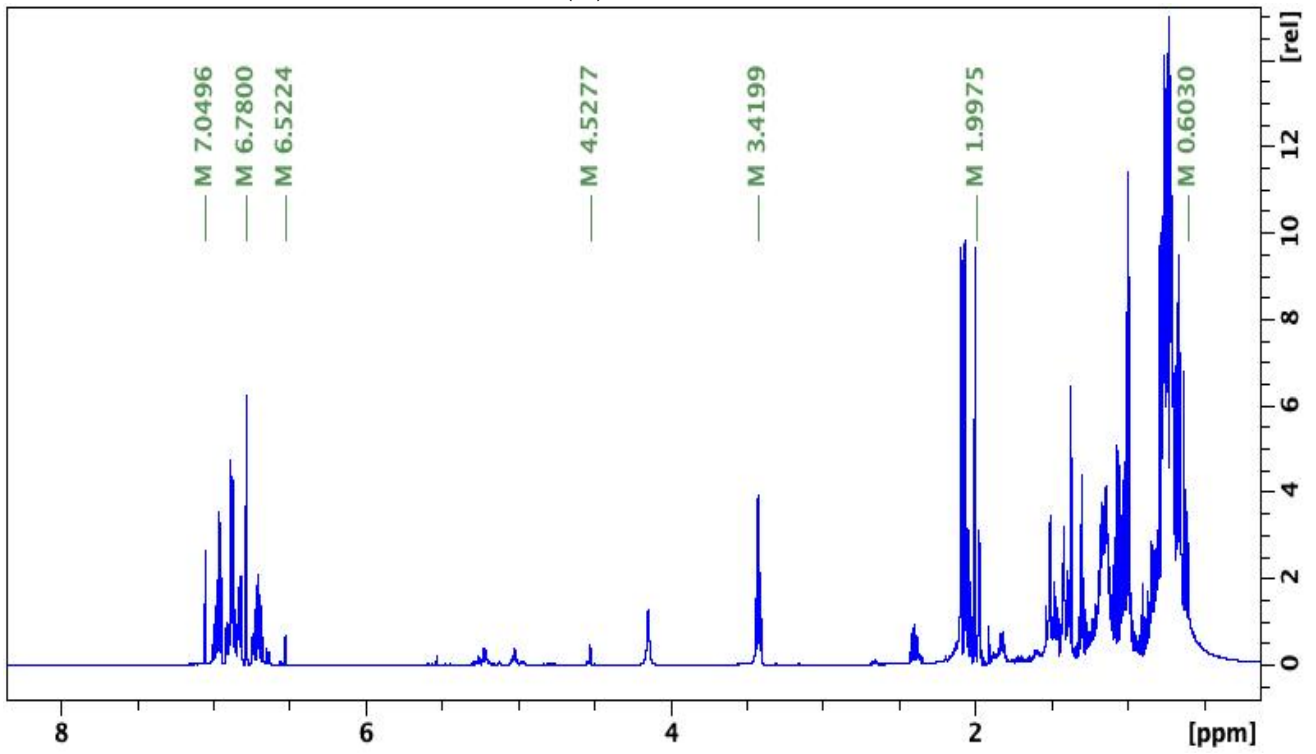

Figure 2. *Cont.*

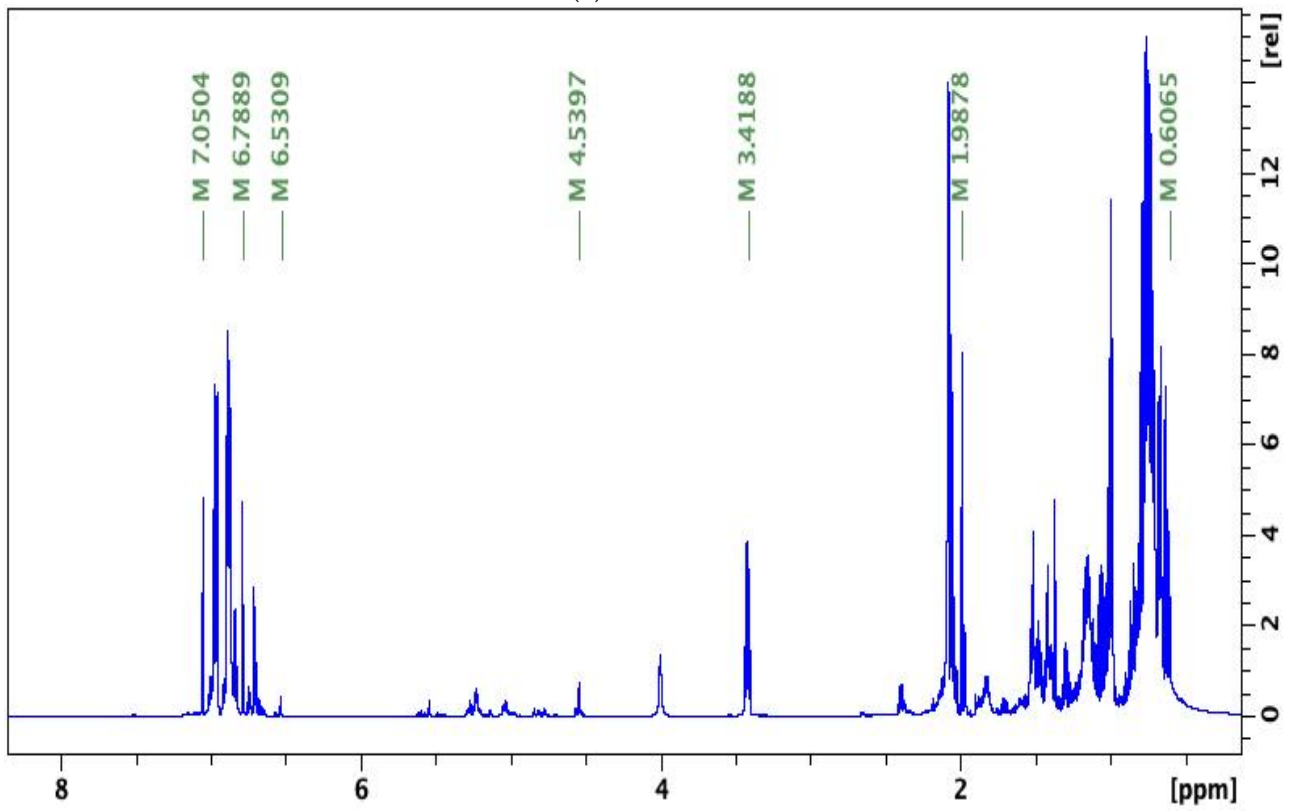(**iv**) Texaco neat

(**v**) Shell neat

Figure 2. *Cont.*
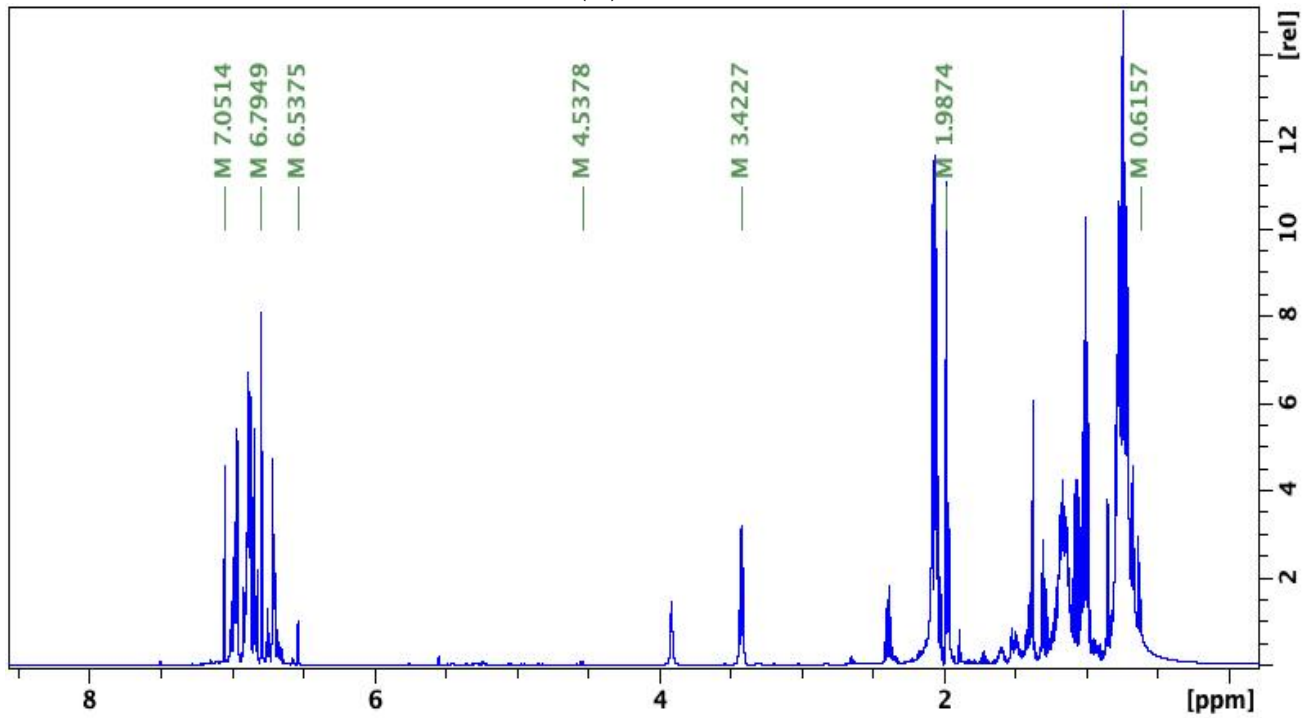
(**vi**) Esso neat

Figure 2. ^{1}H NMR spectra variation between neat petrol from different sources: (**i**) British Petroleum Scotland, (**ii**) British Petroleum UK, (**iii**) Jet (Concord), (**iv**) Texaco, (**v**) Shell and (**vi**) Esso.

(**i**) British Petroleum (UK) neat
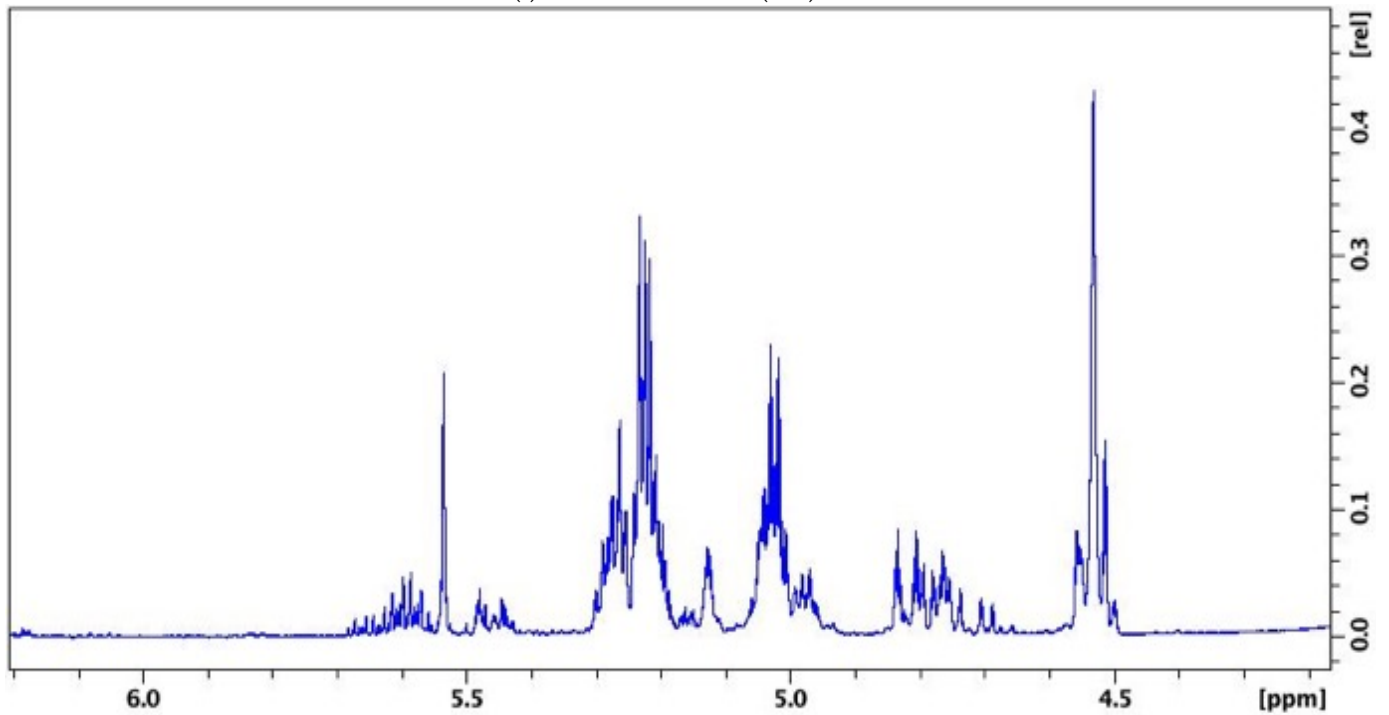
Figure 3. *Cont*.

(**ii**) British Petroleum (Scotland) neat
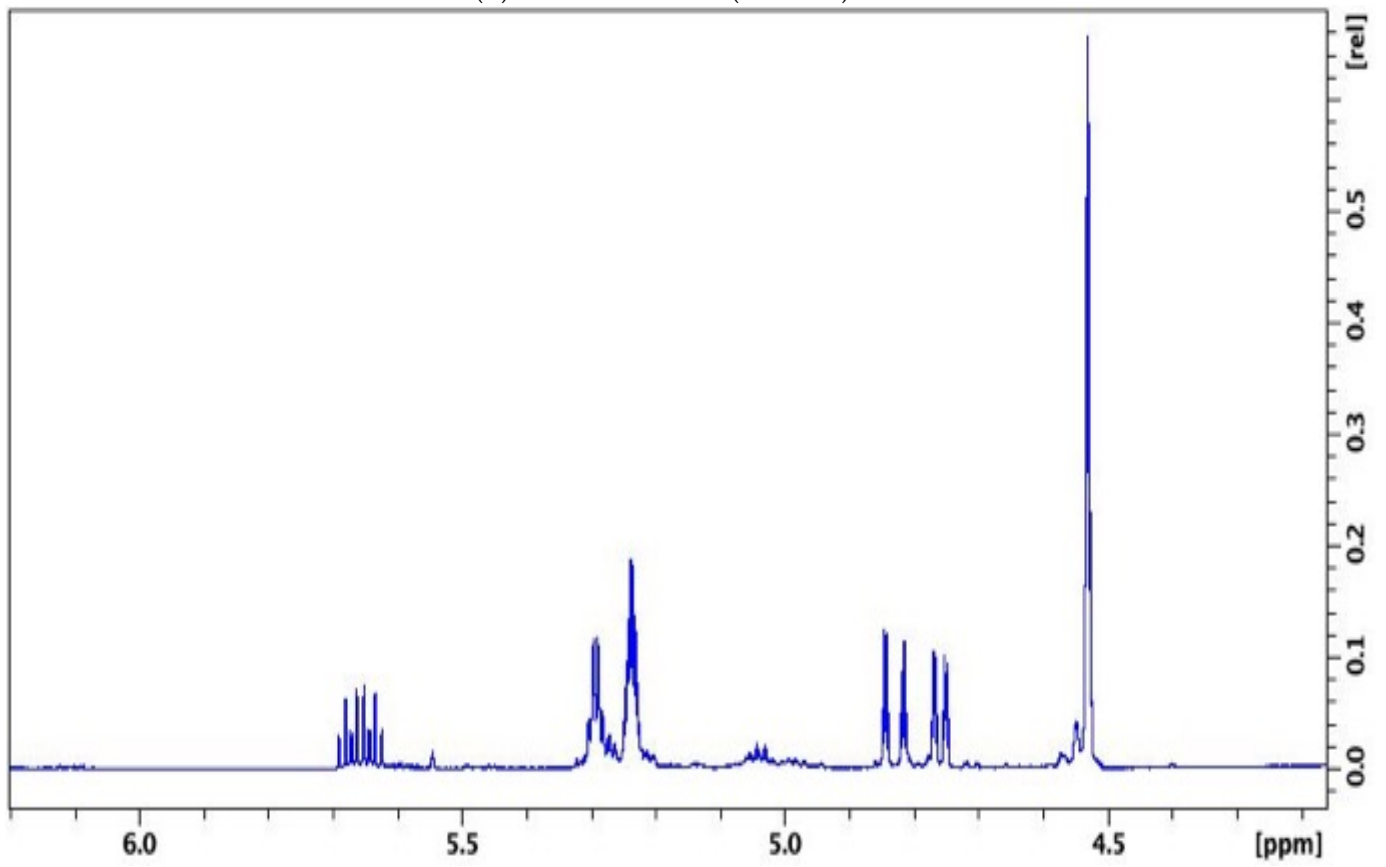
(**iii**) Jet neat
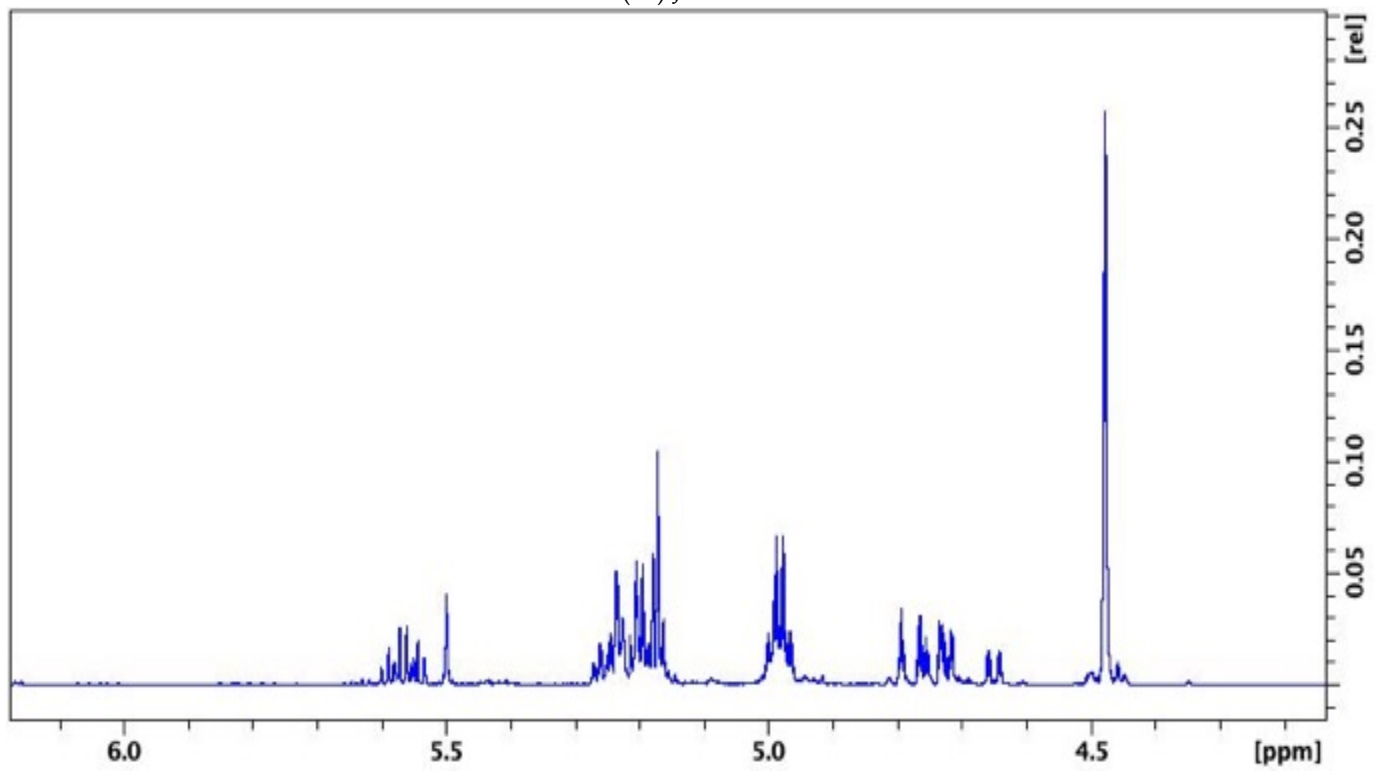
Figure 3. *Cont*.

(**iv**) Texaco neat
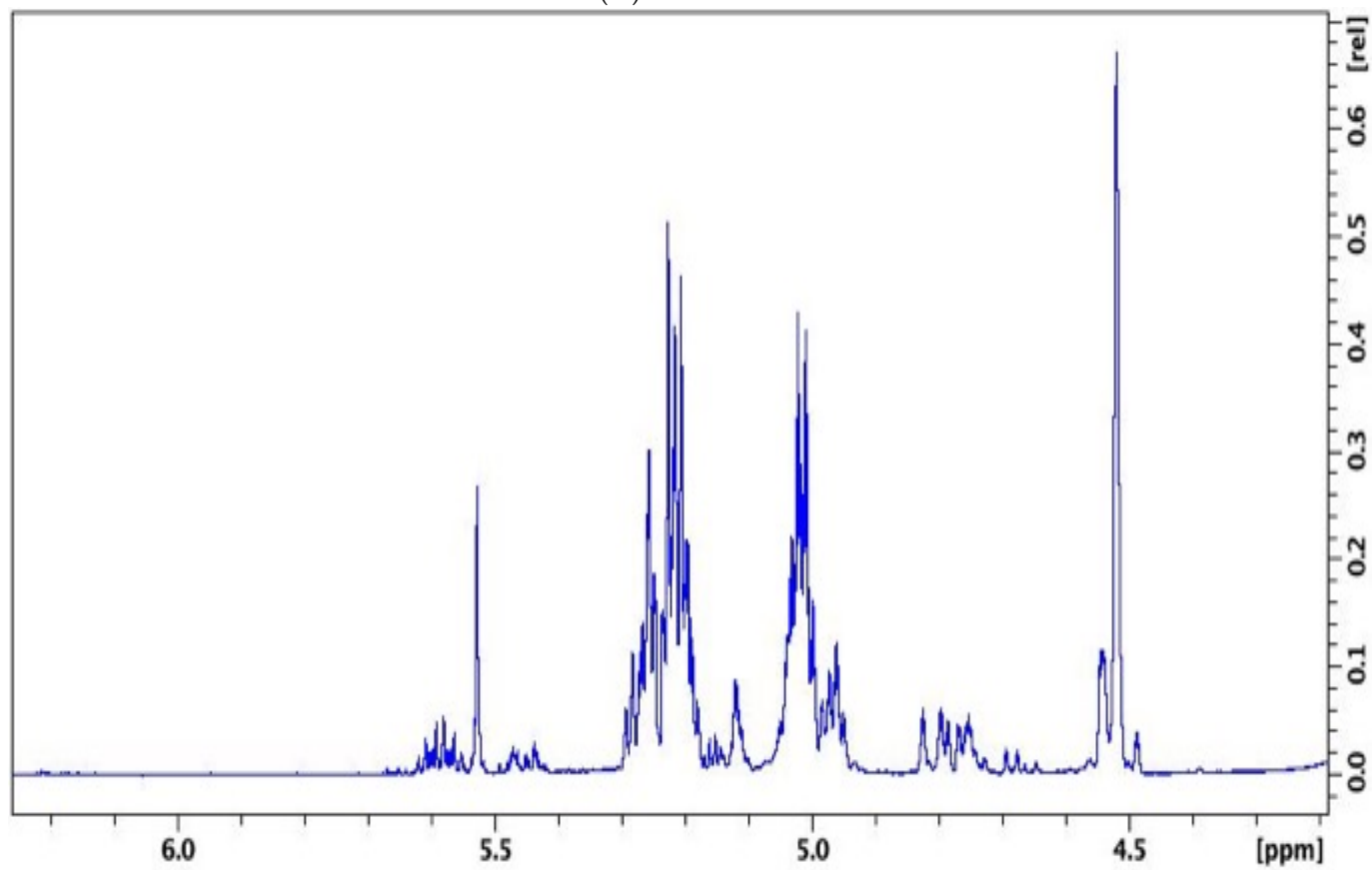
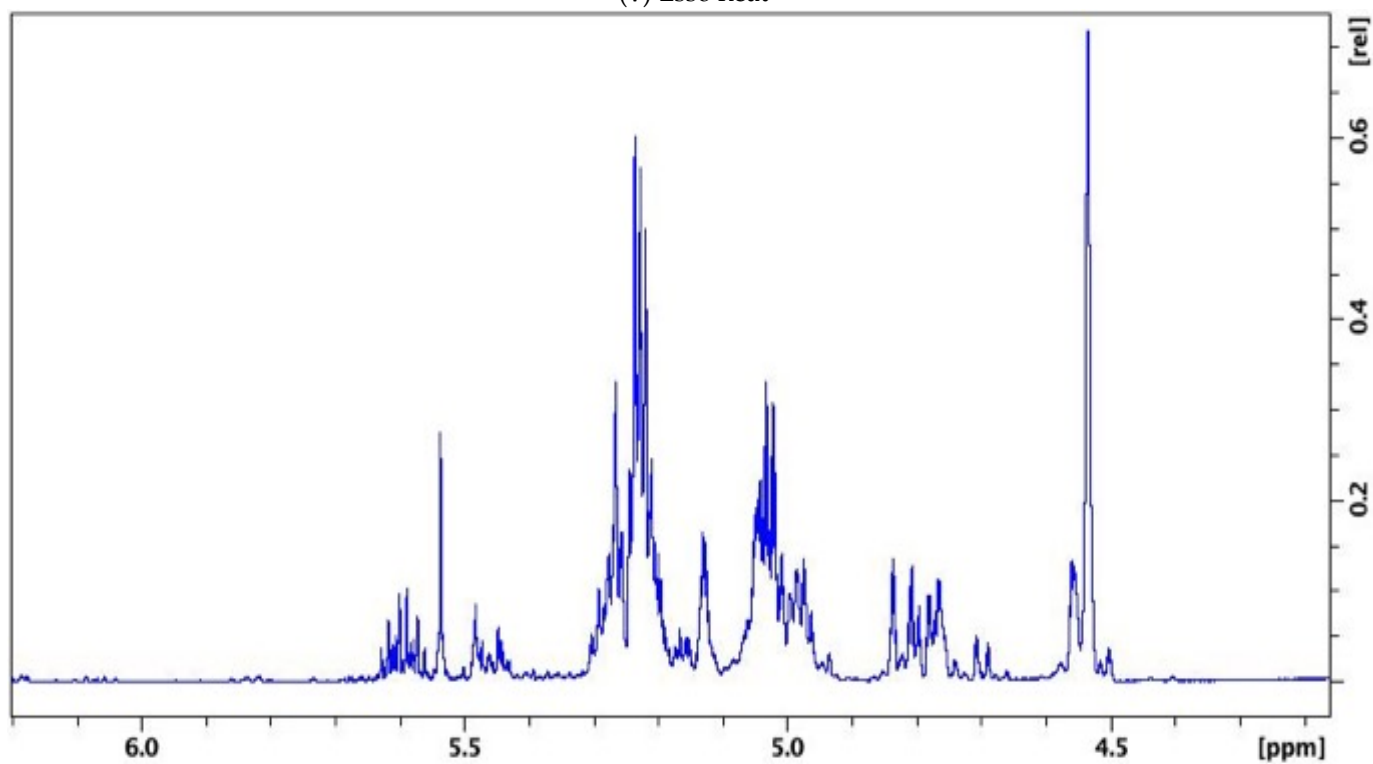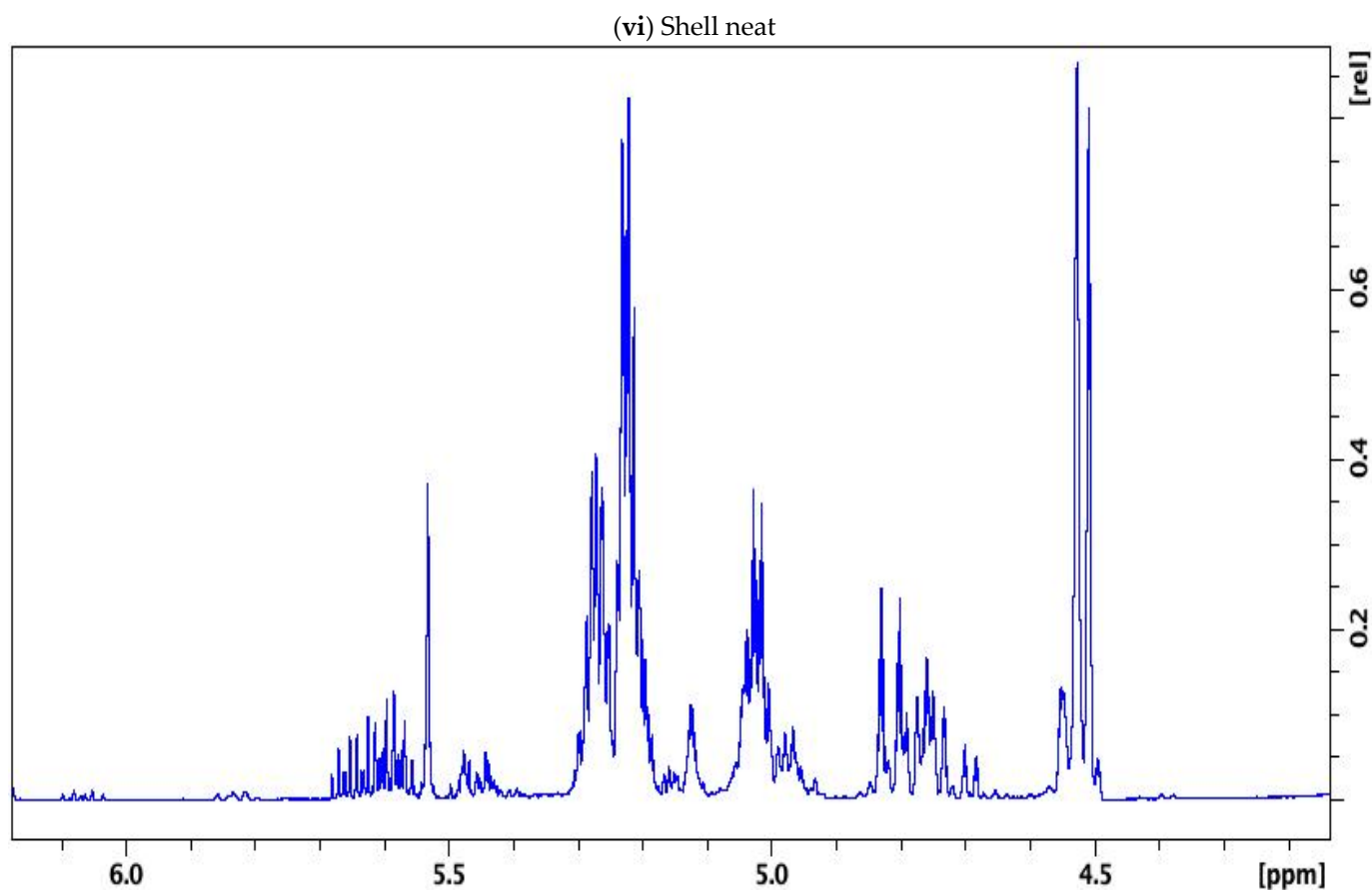(**v**) Esso neat

Figure 3. *Cont.*

(**vi**) Shell neat

Figure 3. Representation of the olefin region of the ^{1}H NMR spectra of neat petrol from different sources: (**i**) British Petroleum Scotland, (**ii**) British Petroleum UK, (**iii**) Jet (Concord), (**iv**) Texaco, (**v**) Esso and (**vi**) Shell.

3.2. [1]H selTOSCY Results for Native Petrol Samples

The ^{1}H selTOCSY NMR experiment allowed for more detailed structural and chemical identification of the 'source-related' compounds in the olefin region of the spectra. The ^{1}H selTOCSY NMR experiments on these alkene signals removed most of the non-alkene-related signals from the spectra, providing additional clarity in the previously heavily overlapped regions of the ^{1}H NMR spectrum. Therefore, selective ^{1}H TOCSY was used to elucidate the 'target' compounds of interest not only to identify them but also to use the obtained ^{1}H selTOCSY data for building a successful classification model. Figure 4 illustrates spectral examples of the great potential of the [1]D selTOCSY experiments in revealing the whole spin system of a band-selective chemical shift. The couplings were resolved and can be used to assign the chemical species. Four discriminative sets of olefins were identified: (a) 3-methyl-1-butene by irradiating the signal at 4.64–4.72 ppm, (b) a mixture of 3-methyl-1-butene and 1-pentene by irradiating the signal at 4.73–4.85 ppm, (c) 2-methyl-2-butene by irradiating the signal at 4.95–5.10 ppm and (d) a mixture of cis and trans-2-pentene by irradiating the signal at 5.10–5.35 ppm.

(**a**)
(i) Aliphatic area of the spectra
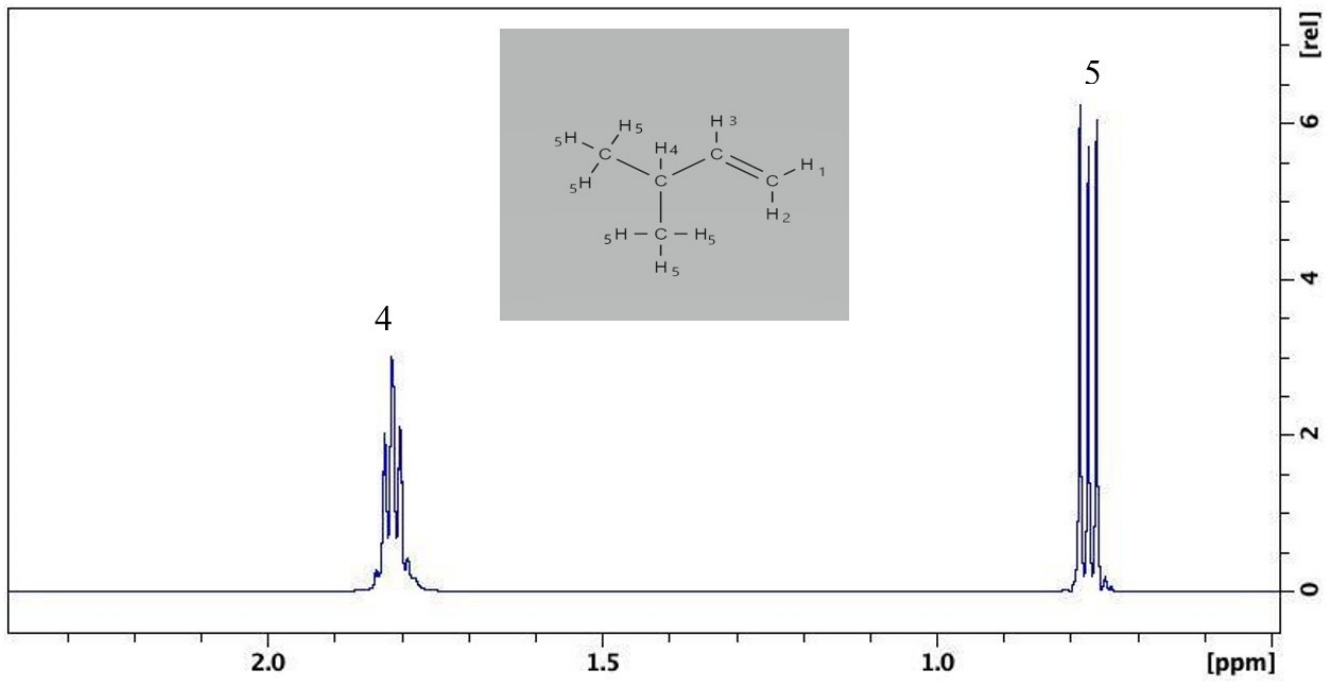
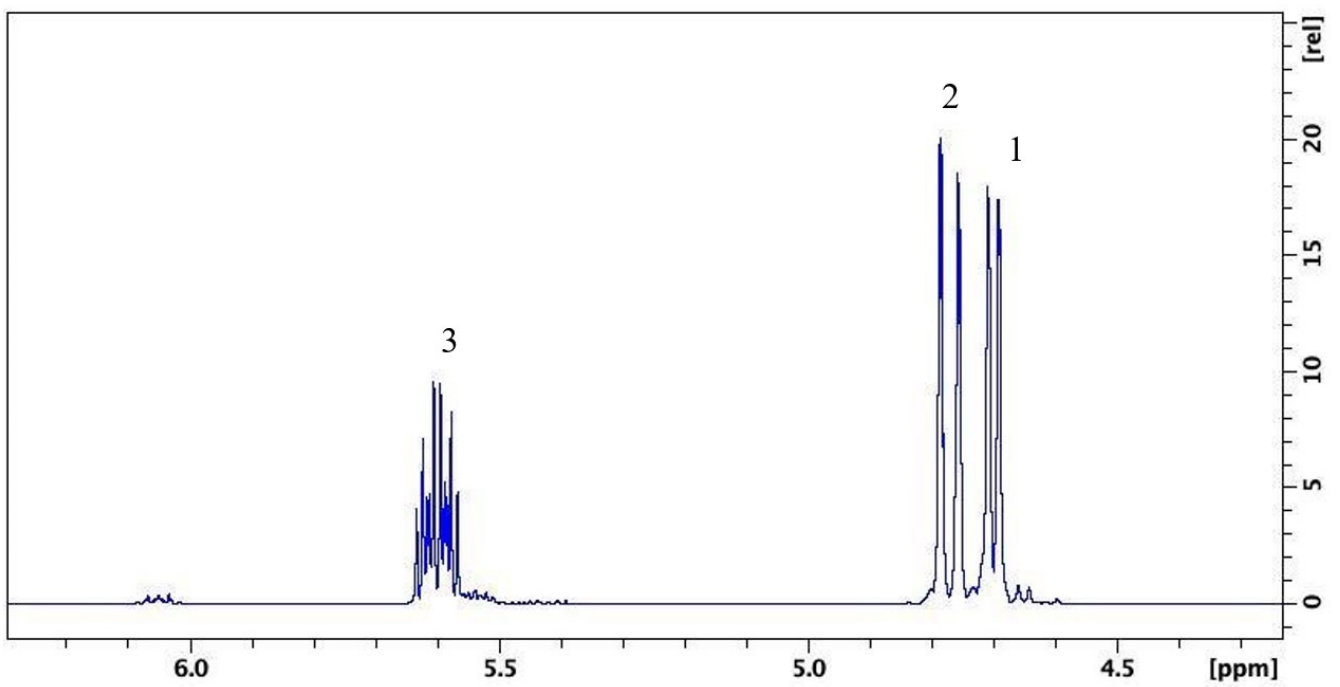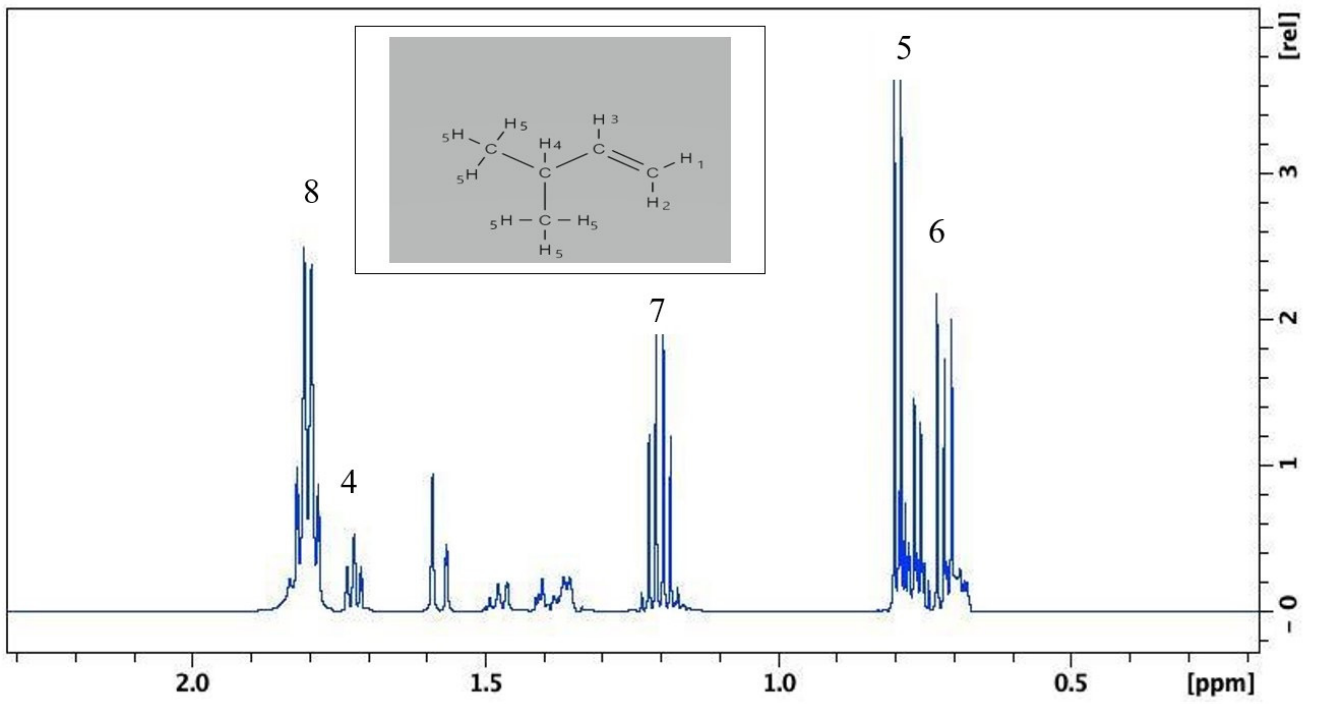(ii) Olefin area of the spectra

Figure 4. *Cont*.

(**b**)

(i) Aliphatic area of the spectra

(ii) Olefin area of the spectra
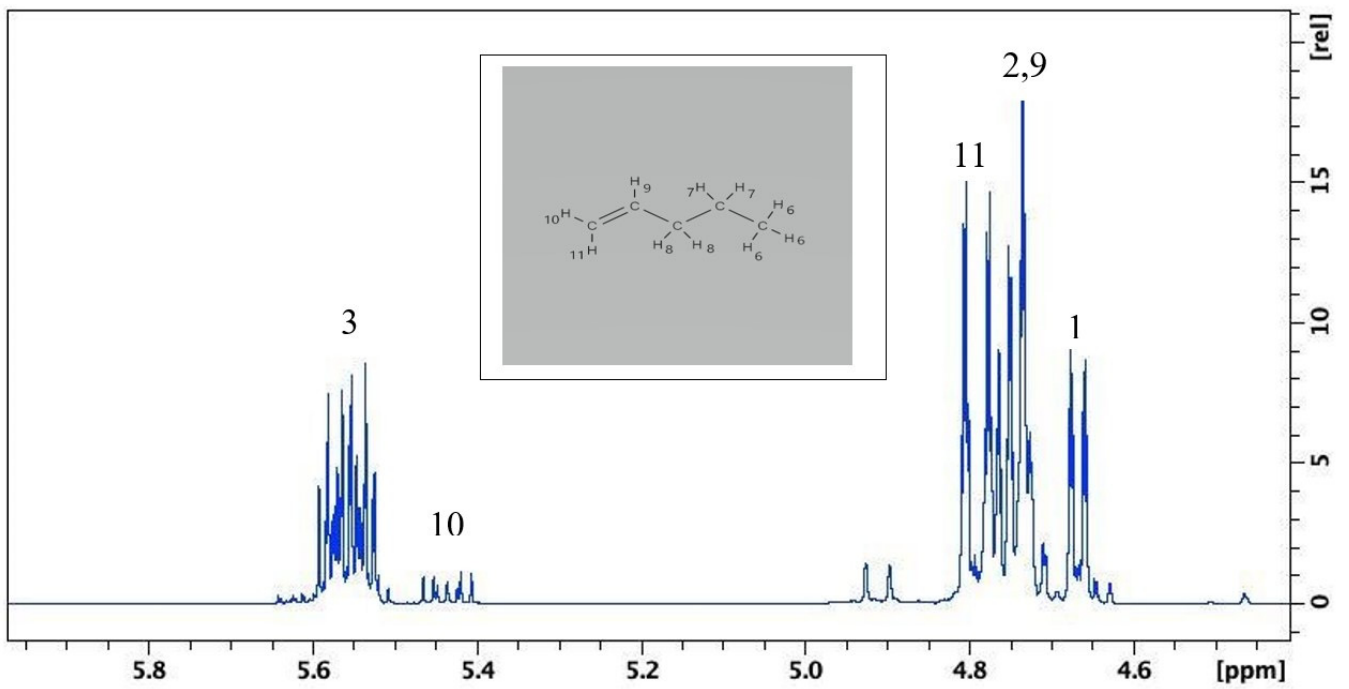
Figure 4. *Cont.*

(**c**)
(i) Aliphatic area of the spectra

(ii) Olefin area of the spectra

Figure 4. *Cont.*

(d)
(i) Aliphatic area of spectra
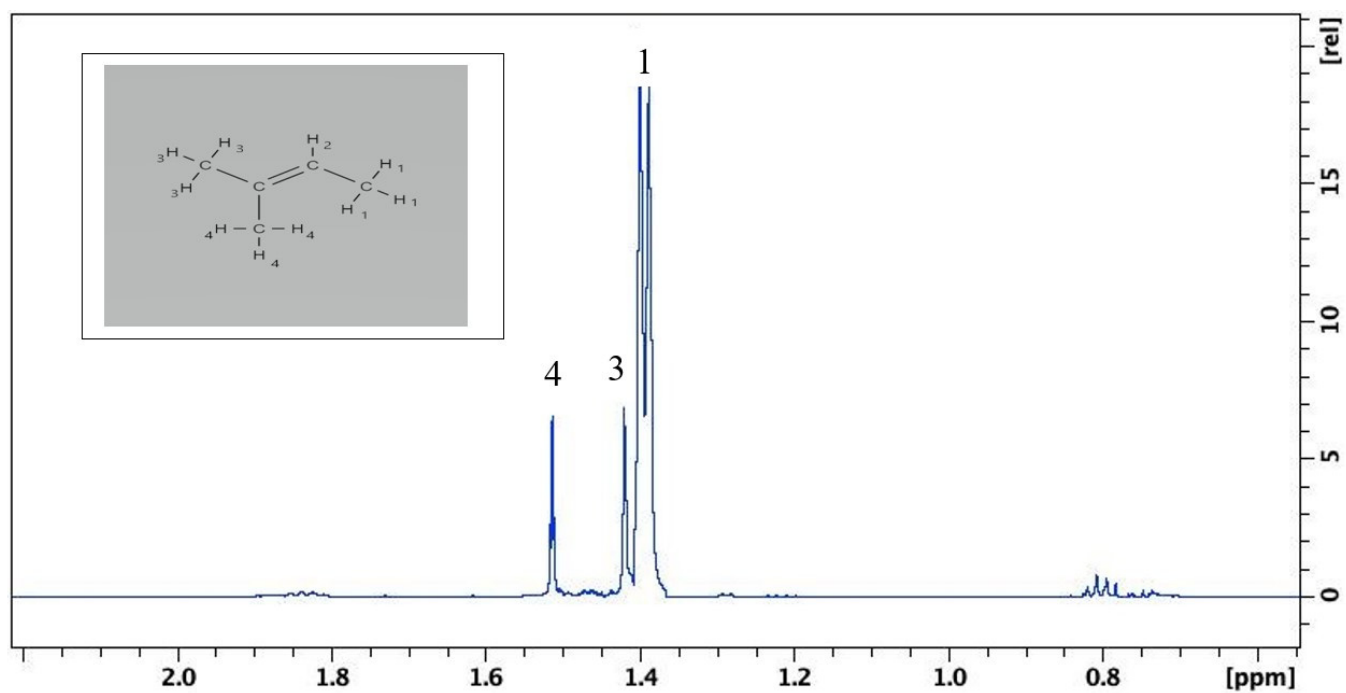
(ii) Olefin area of the spectra
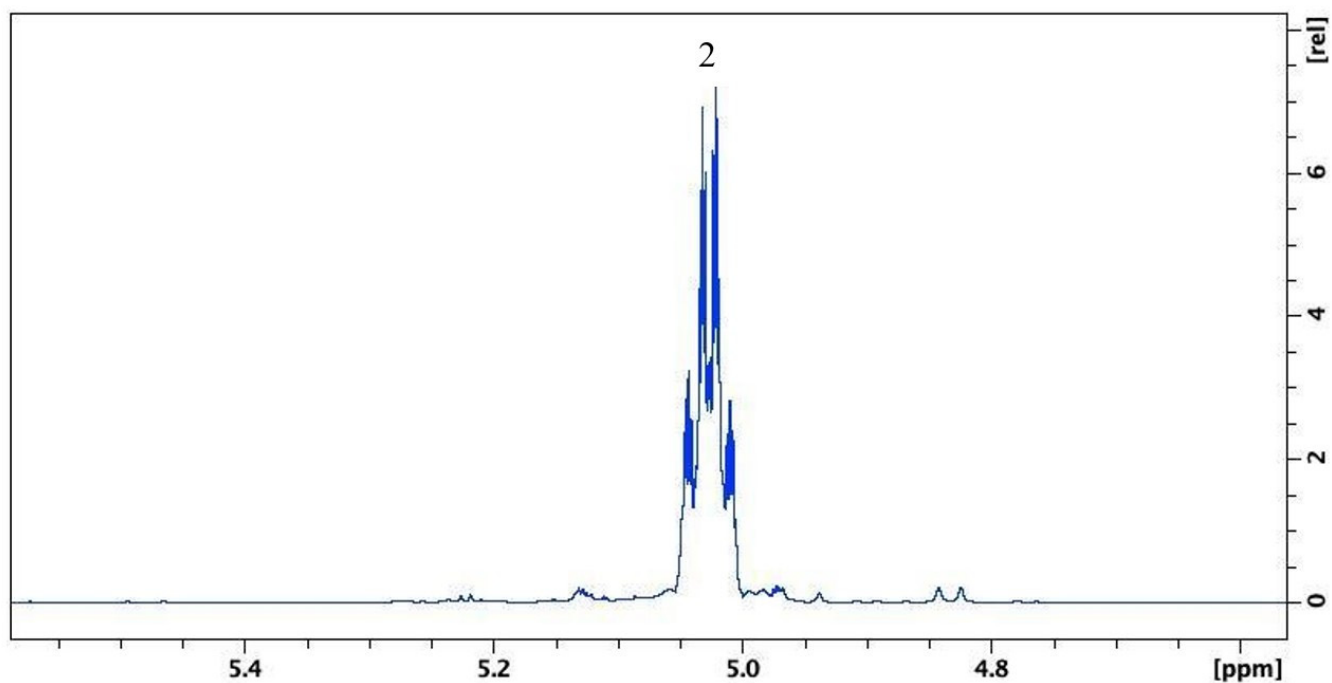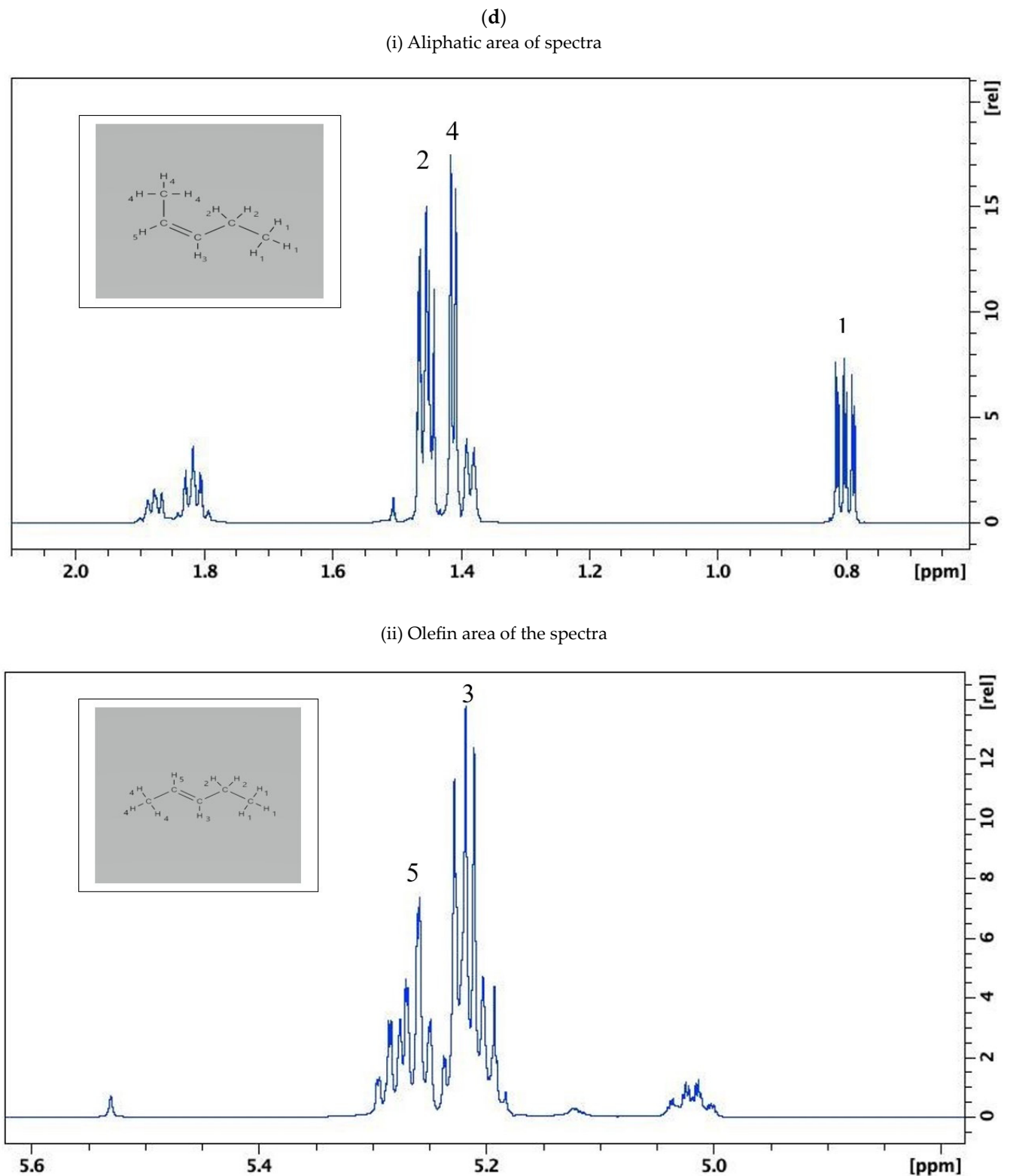
Figure 4. An illustration of the ^{1}H selTOCSY spectra of the distinctive olefins found in petrol source. (**a**) ^{1}H selTOCSY spectra displaying the coupling assignment of the irradiated signal at 4.65–4.72 ppm of 3-methyl-1-butene in Jet petrol source. (**b**) ^{1}H selTOCSY spectra displaying the couplings assignment of the irradiated signal at 4.73–4.85 ppm of the mixture of 3-methyl-1-butene and 1-pentene in Jet petrol source. (**c**) ^{1}H selTOCSY spectra displaying the couplings assignment of the irradiated signal at 4.95–5.10 ppm of 2-methyl-2-butene in Jet petrol source. (**d**) ^{1}H selTOCSY spectra displaying the coupling assignment of the irradiated signal at 5.10–5.35 ppm of the mixture of cis- and trans-2-pentene in Jet petrol source.

It was concluded that petrol sources from BP in Scotland have a unique combination of olefinic compounds that clearly distinguish them from the rest of the petrol sources. The BP M and Jet exhibited similar patterns of the sets of olefins and could not be discriminated on these alone. However, BP M and Jet have the potential to be discriminated from the other three brands based on the couplings of 2-methyl-2-butene. Texaco, Esso and Shell could be potentially discriminated from the other petrol brands based on 2-methyl-2-butene and the mixture of cis- and trans-2-pentene.

The ^{1}H selTOCSY method identified the presence/absence of four important sets of olefins representing the minor differences in the signals (Table 2). The summary table shows the combination of different sets of olefins in the petrol brands.

Table 2. Summary of the alkene compounds present ($\sqrt{}$) and absent (X) in the petrol brands.

	Set 1	Set 2	Set 3	Set 4
Petrol Brand	3-methyl-1-butene	Mixture of 1-pentene and 3-methyl-1-butene	2-methyl-2-butene	Mixture of cis- and trans-2-pentene
BP S	$\sqrt{}$	X	X	$\sqrt{}$
BP M	$\sqrt{}$	$\sqrt{}$	$\sqrt{}$	$\sqrt{}$
Jet	$\sqrt{}$	$\sqrt{}$	$\sqrt{}$	$\sqrt{}$
Texaco	$\sqrt{}$	$\sqrt{}$	X	X
Esso	$\sqrt{}$	$\sqrt{}$	X	X
Shell	$\sqrt{}$	$\sqrt{}$	X	X

3.3. selTOCSY Results for Evaporated Petrol Samples

The investigation of different petrol sources evaporated to 25%, 50%, 75% and 90% of their original weight found distinctive couplings to discriminate between the petrol sources based on 2-methyl-2-butene and the mixture of cis- and trans-2-pentene up to 50% evaporation, except for the BP S petrol source which displayed poor recovery for all relevant olefins. It should be noted that 2-methyl-2-butene and the mixture of cis- and trans-2-pentene were the main olefins contributing to the differentiation between petrol sources. The 75% evaporation samples displayed poor peak resolution, changes in peak intensity and low signal-to-noise ratios resulting in spectra which could not be used for individualization and classification of the petrol sources. The 90% evaporated petrol samples exhibited a complete loss of all the alkene signals that could be used to discriminate between the petrol sources. The petrol samples evaporated to 25% and 50% of their original weight were suitable for individualization and discrimination of petrol sources. The use of ^{1}H selTOCSY dramatically improved the ability to identify the alkenes which subsequently has the potential to link a petrol sample to its source (Table 3). The spectra are shown in Supplementary Materials.

Table 3. A summary table of the aliphatic and olefinic couplings of the individual sets of olefins at 25%, 50% and 75% level of evaporation and their potential to be used for discrimination.

Distinctive Set of Olefins	Aliphatic Couplings	Olefinic Couplings	Potential to Discriminate
1. 3-methyl-1-butene	loss of -CH$_3$ methyl groups and CH couplings at 25% evaporation rate	decrease in relative intensity of CH=CH$_2$ couplings up to complete loss for 50% evaporation rate	X
2. mixture of 1-pentene and 3-methyl-1-butene	loss of all aliphatic couplings at 50% evaporation rate	decrease in relative intensity for all couplings up to 50% evaporation rate, poor resolution was observed for 75% evaporation rate	X

Table 3. *Cont.*

Distinctive Set of Olefins	Aliphatic Couplings	Olefinic Couplings	Potential to Discriminate
3. 2-methyl-2-butene	Preserved of all aliphatic couplings up to 75% evaporation with loss of resolution	decrease in relative intensity for all couplings up to 50% evaporation rate, poor resolution was observed for 75% evaporation rate	√
4. Mixture of cis- and trans- 2-pentene	Preserved of all aliphatic couplings up to 75% evaporation with loss of resolution	decrease in relative intensity for all couplings up to 50% evaporation rate, poor resolution was observed for 75% evaporation rate	√

3.4. Identification of Target Olefin Sets in Petrol Samples Burned on Their Own and on Substrates

The next stage of this study was analyzing the 50% burnt petrol samples from different sources. The aim was to identify the recovered distinctive olefins. The resulting ^{1}H NMR spectra of the burnt petrol samples displayed a preservation of the discriminative olefins with changes in the peak intensity and resolution for all petrol sources except for the BP S petrol source, where no olefins were recovered in the olefin area. The full and olefin area of the ^{1}H NMR spectra of the different petrol samples are shown in Supplementary Materials. However, challenges were met when different petrol sources were spiked on different substrates; Table 4 summarizes the ability to discriminate the petrol source from a variety of different simulated fire debris.

Table 4. Summary table of petrol samples burned on substrates and their significance for forensic investigations.

Type of Substrate	Background Interference	Application	Potential to Discriminate
Wood (flooring): Oak Ash White pine Yellow pine Hickory	Substrate background interference from the pyrolysis of the wood: cellulose and levoglucosan from oak; unidentified peaks in ash, hickory and yellow pine; 2-furalaldehyde and m-xylene from white pine	Household fires	X
Carpets 100% polyester with Acton/Hessian (14 mm thickness) 100% polypropylene with felt backing (12 mm thickness) 50% wool with Action/Hessain (5 mm thickness)	Substrate background interference from the backing of substrate itself due to the polymer styrene	Household/motor vehicle fires	X
Fabrics 100% cotton 100% linen 100% polyester fabric Cotton and linen Cotton and polyester Viscose and linen	Partially recovery of 2-methyl-2-butene and a mixture of cis- and trans-2-pentene identified by ^{1}H selTOCSY method but inconsistent among petrol sources	Household fires, petrol bombs, fires set by humans	X

Table 4. *Cont.*

Type of Substrate	Background Interference	Application	Potential to Discriminate
Paper materials	All sets of olefins were lost due to absorbance and retention capabilities of paper materials	Household and office fires, destruction of evidence	X
Cardboard	Fully recovered 2-methyl-2-butene and a mixture of cis- and trans-2-pentene, which were identified by ^{1}H selTOCSY method	Household and office fires, destruction of evidence	√

3.5. Classification of Neat Petrol Classes

The first stage of the classification experiment aimed to discriminate neat petrol samples by their brand (source/origin). Different classifier models such as Decision Trees, Discriminant Analysis (DA), Support Vector Machines (SVMs), Logistic Regression, k-Nearest Neighbor (k-NN), Naïve Bayes and Ensemble techniques were used. The accuracy of the best performing models is summarized in Table 1, where the training and testing were conducted with 5- and 10-fold cross-validation. The classification model of the entire ^{1}H NMR spectral dataset of neat petrol samples that were filtered and normalized by the sum of the total area displayed an advantage over the non-filtered single peak normalized dataset, successfully classifying BP S and Texaco. The lower accuracy for BP M, Jet, Esso and Shell could be explained by the overall chemical similarities in the whole spectra; as previously stated, the potential discriminative features were the four olefinic compounds/mixtures. Therefore, when feature selection was applied using the spectral bins representing specific couplings, BP S, BP M, Texaco, and Jet were successfully classified previously. In addition, the high intensity of the aliphatic and aromatic region of the ^{1}H NMR spectra could suppress the low-intensity olefin components, which contributed to the model's low accuracy when no feature selection was used. No significant difference was observed when evaluating the dataset with 10 folds.

The classification model of the olefin datasets did not achieve a sufficient accuracy on the classification of all the petrol brands when classifying using the feature selection technique. This could be simply explained by the alkene couplings being spread across the chemical shift range of 4.0–6.0 ppm (allylic carbons) and 1.6–2.6 ppm (aliphatic carbons) rather than couplings only in the 4.0–6.0 ppm olefin region. Comparing the individual olefins and the combination of all four sets of olefins for the training of the classification model revealed that the classification model with the highest accuracy for discriminating between the native petrol samples based on their brand was the one that used the combination of the four featured olefins. The Ensemble Classifier applied to the filtered and normalized by sum of the total area data with 10-fold cross-validation was the most satisfactory classification model that successfully classified all the petrol brands with an overall accuracy of 81% (Table 5).

Table 5. Summary table of the classification models for neat petrol samples with the most successful classifiers of different petrol sources with prediction rates >60% and <60%.

Dataset	Classifier	PCA	Feature Selection	k-Folds	BP S	BP M	Jet	Esso	Shell	Texaco
Entire ^{1}H NMR spectra	Ensemble	√		5	92.3%					91.7%
Entire ^{1}H NMR spectra	SVM		√	10	100%		71.4%			83.3%
Olefin Region	NN	√		5	76.9%					66.7%
Olefin Region	NN		√	10	92.3%					75%
3-methyl-2-butene	NN	√		10	100%		83.3%		88.9%	100%
3-methyl-2-butene	Ensemble		√	5	91.7%		66.7%	62.5%	77.8%	90%

Table 5. *Cont.*

Dataset	Classifier	PCA	Feature Selection	k-Folds	BP S	BP M	Jet	Esso	Shell	Texaco
Mixture of 3-methyl-2-butene and 1-pentene	Ensemble	√		5	n/a	100%				100%
Mixture of 3-methyl-2-butene and 1-pentene	kNN		√	10	n/a				66.7%	76%
2-methyl-2-butene	SVM	√		10	n/a	66.7%	85.7%			76%
2-methyl-2-butene	Ensemble		√	5	n/a		71.4%	71.4%	66.7%	69.2%
Cis- and trans-2-pentene	Linear Discriminant	√		10	85.7%	71.4%	60%		60%	83.3%
Cis- and trans-2-pentene	Ensemble		√	10	100%	71.4%	60%		60%	83.3%
Combined Olefins	Ensemble	√		10	100%	77.8%	71.4%	71.4%	77.8%	76.9%
Combined Olefins	Ensemble		√	10	100%	66.7%	71.4%	71.4%	88.9%	76.9%

3.6. Classification of Native, Evaporated, Burned on Its Own and Burned on Substrate Petrol Classes

The second stage of this research aimed to build a classification model suitable for classifying evaporated, burnt, and burned on substrates petrol samples according to the brand. The ability to link a sample that has undergone compositional changes (through weathering) back to its unevaporated source was investigated. To build the most robust classification model, different classification models were evaluated: (1) the dataset of 2-methyl-2-butene and the mixture of cis- and trans-2-pentene in the neat petrol samples, (2) the evaporated petrol sample dataset, and (3) the combined dataset of neat and evaporated petrol samples.

It was concluded that based on the above evaluated classification models, it was not possible to compare the native and the weathered petrol samples directly. Table 6 summarizes the classification models for the neat, the evaporated, the combined neat and evaporated and the combined evaporated, burnt and substrate sample datasets.

Table 6. A summary table of the classification models for neat, evaporated, burned on its own and on variety of substrates samples petrol samples with the most successful prediction of different petrol source with prediction rate >60% and <60% .

Dataset	Classifier	k-Folds	BP S	BP M	Jet	Esso	Shell	Texaco
Neat Combined	Linear Discriminant	5	85.7%	88.9%				76.9%
Evaporated petrol samples	NN	10				75%		60%
Neat and evaporated petrol samples	NN	10		60%			69.2%	70.6%
Neat, evaporated, burnt and burned on substrate petrol samples	NN	5	100%	62.5%				

3.7. Hierarchical Classification Model for Individualization of Petrol Source

Firstly, the petrol samples status native (unevaporated) or weathered (evaporated and simulated fire debris) needed to be determined. The evaluation of the blind dataset of native petrol brands showed that the classification model with the combination of all four olefins correctly predicted 80% of the petrol brands. On the other hand, it was found that classification model with the combination of 2-methyl-2-butene and the mixture of cis- and trans-2-pentene was not successful in discriminating between petrol brands of the native samples. The combined native and evaporated model displayed an accuracy of 44.4% on the blind dataset. The evaporated and simulated fire debris classification model displayed an accuracy rate of 33%. The multi-classifier classification models for the evaporated and simulated fire debris samples were not satisfactory when evaluated on the blind sets of samples. Therefore, a binary classifier for single sources of petrol vs. the rest of the petrol brands was needed. The binary models were created based on the discriminative potential of the minor compounds using ^{1}H selTOCSY; the four sets of olefins displayed minor differences in the NMR spectra, which could contribute to the distinction between the petrol samples from different sources. The BP M and Jet samples displayed similar NMR spectral fingerprints compared to the other petrol brands based on 2-methyl-2-butene

and the mixture of cis- and trans-2-pentene. In addition, Texaco and Shell shared similar NMR profiles; however, the ^{1}H selTOCSY results for 2-methyl-2-butene and the mixture of cis- and trans-2-pentene could potentially assist with distinguishing these two brands from the rest of the petrol sources. For the weathered dataset, the combination of evaporated petrol samples (25% and 50% of sample weight), petrol samples burned on their own and petrol samples burned on a substrate (cardboard) was used for the training. Due to the limitation of the results of the multi-class classification models, a hierarchical classification model was created to build a more robust classification model.

For this study, a hierarchically structured local classifier model was constructed. Local classifiers were used for each parent node. The first parent node was a binary classification model between the types of petrol samples, native vs. weathered petrol, which produced two child nodes. The binary classifier predicted the source of the petrol sample based on its weathered status using a Linear Discriminant Classifier. Each child node had its local classifier. If the sample was predicted to be native petrol, a multi-class classification model was applied to individualize the native petrol sample based on its source. The Ensemble Classifier used the combination of four olefins compounds to create a multi-class classification model. The classification model for the weathered samples was more complex than the classification model for the native petrol samples due to the complications in the recovered spectra and background interference from the substrate. The weathered samples contained recovered 2-methyl-2-butene and the mixture of cis- and trans-2-pentene. Therefore, for the child node containing weathered samples, a local binary classifier was used for each level. For each level of the second child node, a local binary classifier was used to identify if a sample belonged to a petrol source; if it did not belong to the first petrol source, the sample was input through to the next level and classified. The first leaf level of the binary classification model used the k-NN classifier to determine if a petrol sample was BP M vs. other petrol sources. If BP M was not identified as the source of the sample, the second leaf level used a Logistic Regression Classifier to classify the sample as Jet vs. other petrol sources. If a sample was not identified as BP M and Jet, the third leaf level used a Neural Network Classifier to identify the petrol sample as Esso vs. other petrol sources. The last leaf level was a binary classifier between Shell and Texaco using a Neural Network Classifier (Figure 5).
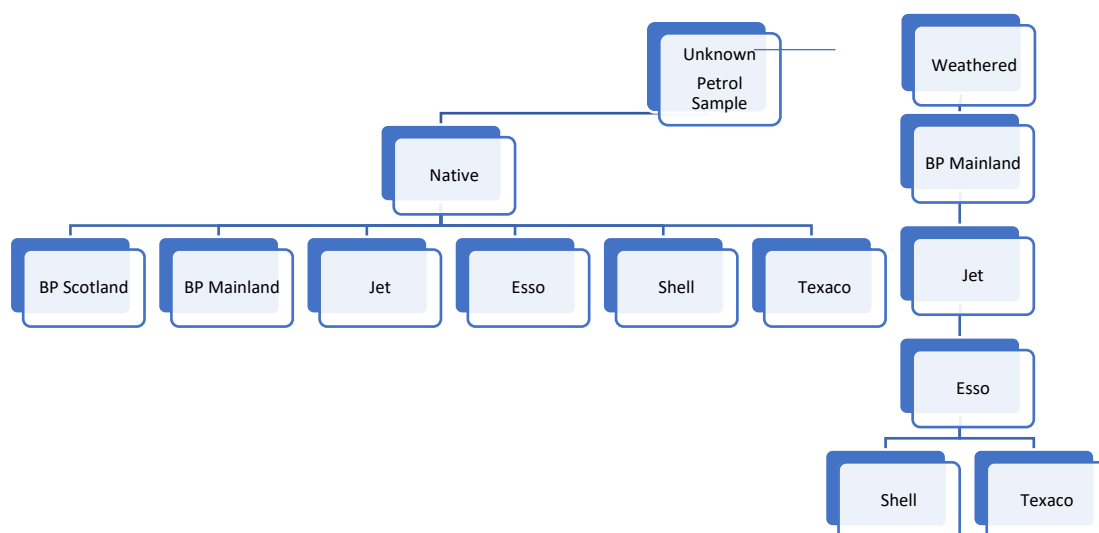

Figure 5. A representation of the hierarchy classification model for native, evaporated, and burnt petrol, and petrol burned on a substrate.

3.8. Blind Study

The first blind dataset that represents native petrol samples was input into the classification model. All the samples were correctly identified as native petrol samples. Table 4 represents the multi-class classification model's output for the prediction of the

petrol sources. The Jet sample was misclassified as Esso; however, this can be explained by the low number of Jet petrol samples that were available for building the classification model. Esso and Shell displayed strong similarities in the ^{1}H selTOCSY results for the four sets of olefins, which were minor compounds potentially contributing to the incorrect classification. Overall, the success rate of classifying the blind dataset of the native petrol samples was 80% compared to the currently used ATD-GC-MS method (30% classification) which uses target compounds and visual interpretation for the comparison of ignitable liquids (Table 7). The second blind dataset, which contained a combination of evaporated petrol samples (25% and 50%), petrol samples burned on their own and petrol samples burned on substrates (petrol samples extracted from cardboard substrate), was input into the model. The goal was to identify and link the petrol samples despite their evaporation status to their source (brand). Firstly, all the blind samples were correctly identified as weathered through the binary model (native vs. weathered). Then, the blind weathered dataset was input through the binary classifiers for each leaf. The NMR classification model displayed an overall accuracy of 78%; one of the Shell and one of the Texaco petrol samples was misclassified compared to the ATD-GC-MS method, which was not suitable for the identification and differentiation of petrol source (Table 8). The local classifier between Shell and Texaco had a 60% success rate, which could have contributed to the misclassification of the Shell petrol sample as a Texaco petrol sample. In addition, both petrol brands displayed similar ^{1}H selTOCSY profiles for the four sets of olefinic compounds. The Esso petrol samples displayed similarities in their ^{1}H selTOCSY profiles for the four sets of olefin compounds to the Shell and Texaco samples investigated.

Table 7. Summary table of the results of the blind dataset of native petrol sources with >60% correct classifications and <60% incorrect classifications.

Sample N	Class	Native vs. Evaporated	Predicted Class by NMR Hierarchical Classifier	ATD-GC-MS
BLIND A	JET	native	ESSO	Identified as unique petrol source
BLIND B	ESSO	native	SHELL	Identified as unique petrol source or similar to J, E, F and H
BLIND C	ESSO	native	ESSO	Sample G identified as similar to Sample C
BLIND D	ESSO	native	ESSO	Sample D identified to be similar to Sample I
BLIND E	TEXACO	native	TEXACO	Sample E and F identified as same petrol source
BLIND F	TEXACO	native	TEXACO	Sample E and F identified as same petrol source
BLIND G	SHELL	native	SHELL	Sample G identified as similar to Sample C
BLIND H	BP M	native	BP M	Sample H and J were grouped with Texaco petrol source
BLIND I	SHELL	native	SHELL	Sample I identified as similar to Sample D
BLIND J	BP S	native	BP S	Sample H and J were grouped with Texaco petrol source

Table 8. Summary table of the results of the blind dataset of weathered petrol sources with >60% correct classifications and <60% incorrect classifications.

Sample N	Class	Native vs. Weathered	BP M Classifier	Jet Classifier	Esso Classifier	Shell/Texaco Classifier	Predicted Class by NMR Hierarchical Classifier	ATD-GC-MS
BLIND A	BP M 50% evaporated	Weathered	BP M				BP M	No differentiation achieved
BLIND B	BP M on cardboard	Weathered	BP M				BP M	Differentiated as different petrol source
BLIND C	JET burnt	Weathered	Other	JET			JET	No differentiation achieved

Table 8. *Cont.*

Sample N	Class	Native vs. Weathered	BP M Classifier	Jet Classifier	Esso Classifier	Shell/Texaco Classifier	Predicted Class by NMR Hierarchical Classifier	ATD-GC-MS
BLIND D	JET 25% evaporated	Weathered	Other	JET			JET	No differentiation achieved
BLIND E	ESSO 25% evaporated	Weathered	Other	Other	ESSO		ESSO	No differentiation achieved
BLIND F	SHELL on cardboard	Weathered	Other	Other	Other	SHELL	SHELL	No differentiation achieved
BLIND G	SHELL burnt	Weathered	Other	Other	Other	TEXACO	TEXACO	No differentiation achieved
BLIND H	TEXACO burnt	Weathered	Other	Other	Other	TEXACO	TEXACO	No differentiation achieved
BLIND I	TEXACO 25% evaporated	Weathered	Other	Other	ESSO		ESSO	No differentiate achieved

In all cases, the native and weathered petrol samples were correctly classified into their respective class regardless of their evaporation and substrate interference status with more than 60% accuracy (Table 9). These results demonstrated that the development of the hierarchical classification model objectively individualized and correctly discriminated petrol brands under one classification model despite their evaporated products or in the presence of interfering products from cardboard substrates.

Table 9. Summary of machine learning model performance output using different classifiers and their accuracy in the classification of different petrol sources.

Classifier	Overall Accuracy (%)	Classification
Linear Discriminant	98.5	Native vs. weathered
Ensemble	80	BP S vs. BP M vs. Jet vs. Texaco vs. Shell vs. Esso
k-NN	84.4	BP M vs. other petrol brands
Logistic Regression	82.4	Jet vs. other petrol brands
ANN	82.1	Esso vs. other petrol brands
ANN	60	Texaco vs. Shell

The model performance was affected by the availability of the data (petrol samples used for training the classification model were limited in number). A higher number of samples would strengthen the model and result in greater accuracy. To compensate for the lower number of data points, the cross-fold validation method was applied to the data to avoid overfitting. In addition, we explored different types of classifiers such as the Ensemble Classifier which combines weaker classifiers to build a more robust classification. Moreover, the limitations that come from the properties of the additives and their chemical alternation (the loss of 3-methyl-1-butene and the mixture of 3-methyl-1-butene and 1-pentene during weathering (evaporation and burning)) impacted the classification model by complicating the like-for-like comparison of native and evaporated/weathered petrol samples.

4. Discussion and Conclusions

The NMR method combined with ML was successfully applied for the individualization and classification of petrol samples from different sources. The NMR spectroscopy method had not been previously evaluated for fire debris analysis. The use of ^{1}H sel TOCSY NMR spectroscopy is also a new approach for fire debris analysis to identify distinctive compounds, background interference and its sources, and pyrolysis products in petrol sources. In addition, the ^{1}H selTOCSY method is innovative for the structural elucidation of petrol samples. Our study proves it can be used for the identification of trace amounts of

specific compounds from the complex spectra of petrol and can be combined with ML for classification purposes.

A hierarchal classification model based on a multi-class classifier for native petrol samples and a combination of binary classifiers for weathered petrol samples was constructed. The overall accuracy of the classification model was 80% for native petrol samples and 78% for weathered petrol samples, significantly outperforming the alternative method (ATD-GC-MS) that is currently used. The combination of ^{1}H combined with the NOESY NMR method and ^{1}H selTOCSY displayed potential in individual identification of fire debris samples and linking them to a source or suspect. The model has the potential to identify an unknown petrol sample and linking it to its source regardless of its evaporation rate (based on to 25% and 50% evaporated samples) and whether it was burned on a cardboard substrate. In conclusion, an automated hierarchical classification model was created for the successful discrimination and individualization of petrol samples based on their source using a machine learning classifier. This paper describes the first ML model that has the potential to be used for the classification of petrol sources of fire debris.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/app14125177/s1, Figure S1: A comparison of ^{1}H selTOCSY of the mixture of 3-methyl-1-butene of (i) BP S (ii) Jet (iii) Esso (iv) Shell and (v) Texaco (vi) BP M native petrol sources with 25%, 50% and 75% evaporation rate; Figure S2: A comparison of ^{1}H selTOCSY of the mixture of 3-methyl-1-butene and 1-pentene of (i) Jet (ii) Esso (iii) Shell and (iv) Texaco (v) BP M native petrol sources with 25%, 50% and 75% evaporation rate; Figure S3: A comparison of ^{1}H selTOCSY of the 2-methyl-2-butene of (i) Jet (ii) Esso (iii) Shell and (iv) Texaco (v) BP M native petrol sources with 25%, 50% and 75% evaporation rate; Figure S4: A comparison of ^{1}H selTOCSY of the mixture of cis and trans-2-pentene of (i) BP S (ii) BP M (iii) Jet (iv) Esso (v) Shell and (vi) Texaco native petrol sources with 25%, 50% and 75% evaporation rate; Figure S5: The illustration of full and olefins area of ^{1}H NMR spectra of burnt on its own petrol in (i) Jet (ii) Esso (iii) Shell and (iv) Texaco (v) BP M; Figure S6: A representation comparison of ^{1}H selTOCSY spectra of (a) 2-methyl-2-butene and (b) the mixture of cis and trans-2-pentene in neat petrol vs in burnt cardboard in (i) Jet, (ii) Esso, (iii) Shell and (iv) Texaco, (v) BP M.

References

1. Bumbrah, G.S.; Sarin, R.K.; Sharma, R.M. Analysis of Petroleum Products in Fire Debris Residues by Gas Chromatography: A Literature review. *Arab. J. Forensic Sci. Forensic Med.* **2017**, *1*, 512–534.
2. Ugena, L.; Moncayo, S.; Manzoor, S.; Rosales, D.; Cáceres, J. Identification and Discrimination of Brands of Fuels by Gas Chromatography and Neural Networks Algorithm in Forensic Research. *J. Anal. Methods Chem.* **2016**, *2016*, 6758281. [CrossRef] [PubMed]
3. Desa, W. The Discrimination of Ignitable Liquids and Ignitable Liquid Residues Using Chemometric Analysis. Ph.D. Thesis, University of Strathclyde, Glasgow, UK, 2012.

4. Monfreda, M.; Gregori, A. Differentiation of unevaporated gasoline samples according to their brands, by SPME-GC-MS and multivariate statistical analysis. *J. Forensic Sci.* **2011**, *56*, 372–380. [CrossRef] [PubMed]

5. Barnett, I.; Bailey, F.; Zhang, M. Detection and Classification of Ignitable Liquid Residues in the Presence of Matrix Interferences by Using Direct Analysis in Real Time Mass Spectrometry. *J. Forensic Sci.* **2019**, *64*, 1486–1494. [CrossRef] [PubMed]

6. Novoa-Carballal, R.; Fernandez-Megia, E.; Jimenez, C.; Riguera, R. NMR methods for unravelling the spectra of complex mixtures. *Nat. Prod. Rep.* **2011**, *28*, 78–98. [CrossRef] [PubMed]

7. Flumignan, D.L.; Boralle, N.; De Oliveira, J.E. Screening Brazilian commercial gasoline quality by hydrogen nuclear magnetic resonance spectroscopic fingerprinting and pattern-recognition multivariate chemometric analysis. *Talanta* **2010**, *82*, 99–105. [CrossRef] [PubMed]

8. Monteiro, M.; Ambrozin, A.; Lião, L.; Boffo, E.; Tavares, L.; Ferreira, M.; Ferreira, A. Study of Brazilian Gasoline Quality Using Hydrogen Nuclear Magnetic Resonance (^{1}H NMR) Spectroscopy and Chemometrics. *Energy Fuels* **2009**, *23*, 272–279. [CrossRef]

9. Obeidat, S.M. The Use of ^{1}H NMR and PCA for Quality Assessment of Gasoline of Different Octane Number. *Appl. Magn. Reason.* **2015**, *46*, 875–883. [CrossRef]

10. Obeidat, S.; Alomary, A. Multivariate Calibration and ^{1}H NMR Spectroscopy for Uncovering Fuel Adulteration. *Appl. Magn. Reson.* **2006**, *47*, 1273–1282. [CrossRef]

11. Sun, C.; Wang, Z. ^{1}H NMR application in characterizing the refinery products of gasoline. *Concepts Magn. Reson. Part A* **2016**, *45A*, e21393. [CrossRef]

12. Pagano, B.; Lauri, I.; De Tito, S.; Persico, G.; Chini, M.G.; Malmendal, A.; Novellino, E.; Randazzo, A. Use of NMR in profiling of cocaine seizures. *Forensic Sci. Int.* **2013**, *231*, 120–124. [CrossRef] [PubMed]

13. Takano, H.; Momota, Y.; Ozaki, T.; Terada, K. Personal Identification from Dental Findings Using AI and Image Analysis against Great Disaster in Japan. *J. Forensic Leg. Investig. Sci.* **2019**, *5*, 041. [CrossRef] [PubMed]

14. Mitchell, F. The use of Artificial Intelligence in digital forensics: An introduction. *Digit. Evid. Electron. Signat. Law Rev.* **2010**, *7*, 35–41. [CrossRef]

15. Vitiello, A.; Di Nunzio, C.; Garofano, L.; Saliva, M.; Ricci, P.; Acampora, G. Bloodstain pattern analysis as optimization problem. *Forensic Sci. Int.* **2016**, *266*, e79–e85. [CrossRef] [PubMed]

16. Chinnikatti, S. Artificial Intelligence in Forensic Science. *Forensic Sci. Addict. Res.* **2018**, *2*. [CrossRef]

17. Rigano, C. Using Artificial Intelligence to Address Criminal Justice Needs. *Natl. Inst. Justice J.* **2019**, *17*, 1–10.

18. Braun, S.; Kalinowski, H.O.; Berger, S. *100 and More Basic NMR Experiments: A Practical Course*; VCH: Weinheim, Germany, 1996.

19. Pavia, D.; Lampman, G.; Kriz, G. *Introduction to Spectroscopy*, 3rd ed.; Thomson Learning: Washington, DC, USA, 2001.

20. Beyer, T.; Schollmayer, C.; Holzgrabe, U. The role of solvents in the signal separation for quantitative 1H NMR spectroscopy. *J. Pharm. Biomed. Anal.* **2010**, *52*, 51–58. [CrossRef] [PubMed]

21. Nawaiseh, A. Audit Opinion Decision Using Artificial Intelligence Techniques: Empirical Study of UK and Ireland. Ph.D. Thesis, Brunel University, London, UK, 2021.