Trends in Analytical Chemistry 57 (2014) 147–158

Contents lists available at ScienceDirect

Trends in Analytical Chemistry

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

Review

# 60 MHz $^1$H NMR spectroscopy for the analysis of edible oils ☆

T. Parker [a], E. Limer [b], A.D. Watson [c], M. Defernez [c], D. Williamson [d], E. Kate Kemsley [c,*]

[a] School of Chemistry, University of East Anglia, Norwich Research Park, Norwich NR4 7TJ, UK
[b] Oriel College, University of Oxford, Oxford OX1 4EW, UK
[c] Analytical Sciences Unit, Institute of Food Research, Norwich Research Park, Norwich NR4 7UA, UK
[d] Oxford Instruments Industrial Analysis, Tubney Woods, Abingdon, Oxford, UK

ARTICLE INFO

*Keywords:*
60 MHz $^1$H NMR
Authenticity
Bench-top NMR
Chemometrics
Double-bond vibration
Edible oil
FTIR
Hazelnut oil
Olive oil
Screening

ABSTRACT

We report the first results from a new 60 MHz $^1$H nuclear magnetic resonance (NMR) bench-top spectrometer, Pulsar, in a study simulating the adulteration of olive oil with hazelnut oil. There were qualitative differences between spectra from the two oil types. A single internal ratio of two isolated groups of peaks could detect hazelnut oil in olive oil at the level of ~13%w/w, whereas a whole-spectrum chemometric approach brought the limit of detection down to 11.2%w/w for a set of independent test samples. The Pulsar's performance was compared to that of Fourier transform infrared (FTIR) spectroscopy. The Pulsar delivered comparable sensitivity and improved specificity, making it a superior screening tool. We also mapped NMR onto FTIR spectra using a correlation-matrix approach. Interpretation of this heat-map combined with the established annotations of the NMR spectra suggested a hitherto undocumented feature in the IR spectrum at ~1130 cm$^{-1}$, attributable to a double-bond vibration.
© 2014 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/3.0/).

**Contents**

1. Introduction . . . 148
2. Materials and methods . . . 150
 2.1. Samples . . . 150
 2.2. Spectral acquisition . . . 150
 2.3. Chemometric analysis . . . 152
  2.3.1. NMR . . . 152
  2.3.2. MIR . . . 152
3. Results and discussion . . . 152
 3.1. NMR spectra: assignment and visual inspection . . . 152
 3.2. NMR spectra: quantitative data analysis . . . 152
 3.3. Chemometric analysis of NMR data . . . 153
 3.4. Comparison with IR data . . . 155
  3.4.1. Band 1 (~1.3 ppm) . . . 156
  3.4.2. Band 2 (1.8–2.1 ppm) . . . 156
  3.4.3. Band 3 (2.6–2.9 ppm) . . . 157
  3.4.4. Band 4 (5.2–5.6 ppm) . . . 157
 3.5. Chemometric analysis of MIR data . . . 157
4. Conclusions . . . 157
 Acknowledgements . . . 158
 References . . . 158

☆ This paper was originally submitted to the recently published Special Issue: 'Modern Food Analysis and Foodomics' (http://www.sciencedirect.com/science/journal/01659936/52).

* Corresponding author. Tel.: +44 (0)1603 255014.
*E-mail address:* kate.kemsley@ifr.ac.uk (E.K. Kemsley).

http://dx.doi.org/10.1016/j.trac.2014.02.006
0165-9936/© 2014 The Authors. Published by Elsevier Ltd.
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/3.0/).

## 1. Introduction

Nuclear magnetic resonance (NMR) spectroscopy has a long, distinguished history as an academic research tool, and is a cornerstone of modern analytical chemistry in laboratories across the world. However, virtually all current instruments utilize expensive high-field super-conducting magnets and probes that require cryogenic cooling, as well as specialist staff to operate the spectrometer. Hence, NMR is not generally regarded as a candidate technique for routine, low-cost analysis in industrial settings.

In the pursuit of spectral resolution and sensitivity, the combination of modern electronics, computing power and a low-field, permanent magnet for conducting NMR spectroscopy has been somewhat overlooked. Recently, innovative hardware design resulted in the production of a low-field, cryogen-free $^1$H NMR spectrometer having a bench-top footprint (Pulsar, Oxford Instruments, Tubney Woods, Abingdon, Oxford, UK) [1]. The strength of the permanent magnet is 1.4 T (60 MHz), substantially less than the 7.1 T (300 MHz) found in current, entry-level high-field instruments.

Traditionally, NMR spectra are displayed as signal amplitude *versus* parts per million (ppm) of the spectrometer base frequency because the resonance frequency of nuclei depends on the strength of the static magnetic field. As a result, spectral lines of the same width in Hz will appear broader relative to the spectral window in ppm as the static field strength decreases. In turn, this results in increased overlap between peaks, as illustrated in Fig. 1 which compares the simulated spectrum of glyceryl trioleate at 1.4 T (bottom) and 7.1 T (top) created using NMRPredict (Modgraph Consultants Ltd, Hertfordshire, UK) as implemented in MNova (version 8.1, Mestrelab Research, Santiago de Compostela, Spain). In addition to the overlap issue, some peaks can appear further apart at a lower field on a ppm scale because scalar couplings are independent of field strength. For example, in Fig. 1, the single cluster of peaks at ~2.4 ppm in the high-field regime becomes a pair of separated peaks at ~2.4 ppm and ~2.3 ppm in the low-field regime. The existing reservoir of high-field $^1$H NMR data is helpful in interpreting 60 MHz results but comparisons must be made with care.

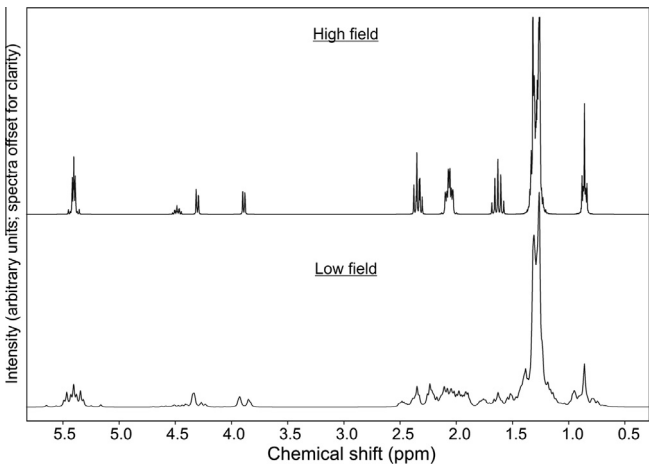



**Fig. 1.** Comparison of the simulated $^1$H NMR spectrum of glyceryl trioleate at 1.4 T (bottom, low-field, 60 MHz) and 7.1 T (top, high-field, 300 MHz) to highlight the impact of the different field strengths on an otherwise identical system. The model spectra were created using NMRPredict implemented in MNova. Note the broadening of peaks and, in some cases, the separation of peaks [e.g., the single cluster of peaks at ~2.4 ppm (high-field), which becomes a pair of peaks at ~2.4 ppm and ~2.3 ppm (low-field)]. The feature showing at ~3.9 ppm in the high-field panel is misplaced from ~4.1 ppm, the result of a modelling imperfection. Similarly, the two low-field features at ~3.9 ppm and ~4.4 ppm should be a single feature at ~4.2 ppm.

High-resolution NMR spectra of any kind have only been routinely acquired and stored in digital form since the 1990s, when field strengths of 7.1 T upwards were the norm [2]. It was around the same time that chemometrics began to emerge as a distinct discipline, concerned with the statistical analysis of large, digitally-stored datasets acquired by modern analytical instrumentation [3]. Since then, chemometric methods have been extensively applied to high-field NMR spectra in many diverse applications [4], particularly metabolomics [5,6], but, due to the chronology, there has been almost no chemometric analysis of 60 MHz NMR spectra.

In this article, we focus on the potential of 60 MHz NMR spectroscopy combined with chemometric analysis to address a simulated screening application of long-standing interest: detecting the adulteration of high-value olive oils with cheaper substitutes [7], a food authentication issue dramatized in a popular book by Mueller [8]. In the present work, hazelnut oil is used as a model adulterant, because its fatty acid composition is very similar to that of olive oil [9]. Preliminary analysis of a wide range of edible oils using 60 MHz NMR (unpublished data) confirmed that, of the oils surveyed, mixtures of olive and hazelnut oils would present the most challenging authentication problem.

Hazelnut oil adulteration of olive oil is also a genuine authenticity issue, as evidenced by a carefully documented episode in which over 20 000 tonnes of Turkish hazelnut oil was shipped to Northern Europe as sunflower oil, with much of it then being transported in French trucks to Spain, where it was mixed in proportions of 15–50% with olive oil, before being sold on to Spanish production and bottling facilities as olive oil [10,11].

Numerous analytical techniques have been used to study adulteration of olive oil, including with hazelnut oil [12]. In particular, the detection of hazelnut oil in olive oil using high-resolution $^1$H NMR spectroscopy has been studied previously [13–17] and reviewed [18]. NMR applied to quality assessment and adulteration more generally was reviewed by Dais and Hatzakis [19].

In terms of operation, 60 MHz bench-top NMR has more in common with another molecular spectroscopy technique, Fourier transform infrared spectroscopy (FTIR), than with high-field NMR. Although lacking in sensitivity and specificity in comparison to, say, chromatography or mass spectrometry techniques, the ease of use and the modest cost of FTIR make it an attractive approach for screening applications in many sectors. Edible oils are particularly amenable to FTIR analysis, and many studies have evaluated FTIR as a screening tool [20–25], including the detection of hazelnut oil in olive oil [26–29]. Consequently, high-resolution FTIR spectroscopy operating in the mid-IR (MIR) provides a natural comparison with bench-top 60 MHz NMR.

However, bench-top NMR has one immediate advantage over FTIR: the chemical specificity of the NMR spectrum. The high-field (9.4 T) $^1$H spectra of fatty acids and their derivatives have been fully assigned [30–32] and this information can be used to identify peaks in the lower field spectrum. In contrast, even in the 'fingerprint' region of MIR, there is widespread overlap of the peaks arising from the various vibrations of different chemical bonds in samples as complex as edible oils, making comprehensive peak annotation difficult.

The primary purpose of the present work is to establish whether, despite inherent overlap and peak broadening, spectra from a 60 MHz spectrometer can be combined with modern data analysis to provide useful information for applications of a certain class. If so, bench-top NMR spectroscopy will have considerable potential for use as a rapid screening tool. We also compare the results obtained using a bench-top Pulsar system with comparable analysis carried out using high-quality FTIR, on the grounds that the two analytical platforms are of similar cost and complexity, and are therefore to some extent competitors in the world of rapid screening.

**Table 1**
Details of olive and hazelnut oils, and prepared mixtures. 'EV' indicates extra virgin olive oil; 'HN' is hazelnut oil. EV01 to EV18 inclusive (note: there are no EV03, EV07 and EV17 codes), plus HN03, HN07, HN17 and HN19, are collectively termed Batch 1. EV26 to EV30 inclusive, and HN20 to HN25 inclusive, are collectively termed Batch 2. Batch 1 contains 69 sample mixtures, Batch 2 contains 75. Batch 1 was purchased and analyzed by IR in August and September 2012, and by NMR in March and April 2013. Batch 2 was purchased and analyzed by IR in October and November 2012, and by NMR in April and May 2013

| Batch 1 | | | |
|---|---|---|---|
| Mixture number | Identity of the "base" olive oil | Identity of the "adulterant" hazelnut oil | %w/w of hazelnut oil |
| 1 | EV01 | HN03 | 19.14 |
| 2 | | | 13.98 |
| 3 | | | 9.07 |
| 4 | | | 4.19 |
| 5 | EV02 | | 18.99 |
| 6 | | | 14.01 |
| 7 | | | 8.97 |
| 8 | | | 4.13 |
| 9 | EV04 | | 19.03 |
| 10 | | | 13.95 |
| 11 | | | 8.93 |
| 12 | | | 4.06 |
| 13 | EV05 | HN07 | 18.99 |
| 14 | | | 13.99 |
| 15 | | | 8.91 |
| 16 | | | 4.08 |
| 17 | EV06 | | 18.73 |
| 18 | | | 14.23 |
| 19 | | | 9.16 |
| 20 | | | 4.06 |
| 21 | EV08 | | 18.18 |
| 22 | | | 13.98 |
| 23 | | | 9.21 |
| 24 | | | 4.05 |
| 25 | EV09 | HN17 | 25.02 |
| 26 | | | 20.13 |
| 27 | | | 15.24 |
| 28 | | | 10.03 |
| 29 | | | 5.05 |
| 30 | EV10 | | 25.05 |
| 31 | | | 20.54 |
| 32 | | | 14.98 |
| 33 | | | 10.05 |
| 34 | | | 5.12 |
| 35 | EV11 | | 24.93 |
| 36 | | | 20.37 |
| 37 | | | 14.98 |
| 38 | | | 9.85 |
| 39 | | | 5.06 |
| 40 | EV12 | | 24.73 |
| 41 | | | 19.86 |

| Batch 2 | | | |
|---|---|---|---|
| Mixture number | Identity of the "base" olive oil | Identity of the "adulterant" hazelnut oil | %w/w of hazelnut oil |
| 1 | EV26 | HN20 | 5.30 |
| 2 | | | 10.05 |
| 3 | | | 14.78 |
| 4 | | | 20.16 |
| 5 | | | 24.98 |
| 6 | | HN23 | 5.17 |
| 7 | | | 10.38 |
| 8 | | | 15.78 |
| 9 | | | 19.65 |
| 10 | | | 24.58 |
| 11 | | HN24 | 5.24 |
| 12 | | | 9.83 |
| 13 | | | 15.58 |
| 14 | | | 20.18 |
| 15 | | | 25.38 |
| 16 | EV27 | HN21 | 4.96 |
| 17 | | | 9.86 |
| 18 | | | 15.18 |
| 19 | | | 19.90 |
| 20 | | | 25.42 |
| 21 | | HN22 | 4.93 |
| 22 | | | 10.08 |
| 23 | | | 15.03 |
| 24 | | | 20.28 |
| 25 | | | 25.05 |
| 26 | | HN25 | 4.93 |
| 27 | | | 10.22 |
| 28 | | | 15.04 |
| 29 | | | 19.89 |
| 30 | | | 25.08 |
| 31 | EV28 | HN22 | 5.15 |
| 32 | | | 9.84 |
| 33 | | | 15.09 |
| 34 | | | 20.11 |
| 35 | | | 25.23 |
| 36 | | HN24 | 5.27 |
| 37 | | | 9.65 |
| 38 | | | 15.04 |
| 39 | | | 20.45 |
| 40 | | | 24.60 |
| 41 | | HN25 | 5.06 |

(*continued on next page*)

**Table 1** (continued)

| | | | |
|---|---|---|---|
| 42 | | | 14.28 |
| 43 | | | 10.01 |
| 44 | | | 5.22 |
| 45 | EV13 | | 25.29 |
| 46 | | | 20.32 |
| 47 | | | 15.09 |
| 48 | | | 10.21 |
| 49 | | | 4.83 |
| 50 | EV14 | | 25.60 |
| 51 | | | 20.25 |
| 52 | | | 13.95 |
| 53 | | | 10.04 |
| 54 | | | 5.40 |
| 55 | EV15 | | 25.82 |
| 56 | | | 20.65 |
| 57 | | | 14.95 |
| 58 | | | 9.91 |
| 59 | | | 4.39 |
| 60 | EV16 | | 25.00 |
| 61 | | | 19.70 |
| 62 | | | 15.19 |
| 63 | | | 10.12 |
| 64 | | | 5.20 |
| 65 | EV18 | HN19 | 24.91 |
| 66 | | | 19.99 |
| 67 | | | 15.34 |
| 68 | | | 9.95 |
| 69 | | | 5.06 |

| | | | |
|---|---|---|---|
| 42 | | | 10.11 |
| 43 | | | 14.76 |
| 44 | | | 19.74 |
| 45 | | | 25.04 |
| 46 | EV29 | HN20 | 5.38 |
| 47 | | | 9.97 |
| 48 | | | 15.00 |
| 49 | | | 21.31 |
| 50 | | | 25.14 |
| 51 | | HN22 | 5.44 |
| 52 | | | 10.36 |
| 53 | | | 15.19 |
| 54 | | | 19.79 |
| 55 | | | 25.73 |
| 56 | | HN23 | 5.43 |
| 57 | | | 10.24 |
| 58 | | | 14.68 |
| 59 | | | 20.19 |
| 60 | | | 24.88 |
| 61 | EV30 | HN21 | 5.13 |
| 62 | | | 10.09 |
| 63 | | | 15.04 |
| 64 | | | 19.86 |
| 65 | | | 25.01 |
| 66 | | HN22 | 4.97 |
| 67 | | | 9.90 |
| 68 | | | 14.85 |
| 69 | | | 20.42 |
| 70 | | | 24.98 |
| 71 | | HN23 | 4.81 |
| 72 | | | 10.21 |
| 73 | | | 15.11 |
| 74 | | | 19.87 |
| 75 | | | 25.11 |

## 2. Materials and methods

### 2.1. Samples

20 extra virgin olive oils and 10 hazelnut oils were purchased from various UK retailers in two chronologically separated Batches (“1”, purchased in August 2012, and “2”, purchased in October 2012; see Table 1). The numbers of hazelnut and olive oils used in this work reflect the relative local availability. Mixtures were prepared by combining varying amounts of hazelnut and olive oils, in which the hazelnut oil component simulates an adulterant in the olive oil. The quantities of hazelnut oil used were in the range 4–26%w/w. The range was chosen by examining typical detection rates by IR and high-resolution NMR spectroscopy reported in the literature.

Batch 1 samples were used to prepare 69 mixtures in total, and Batch 2 to prepare 75, giving a total of 144 samples of oil mixtures. In the subsequent chemometric analyses, calibration for the hazelnut oil content was carried out using the data from Batch 1 samples (mixtures and olive oils) only. Batch 2 samples, prepared from a different set of olive oil and hazelnut oil samples and analyzed by each technique some weeks later, were used to challenge the calibration models. This protocol is intended to mimic a real-world scenario, in which the original spectral collection and model development precedes the application of the screening tool to incoming test samples. It not only tests the ability of the statistical model to generalize beyond data from samples used in its development, but also the ability of the instrumentation to produce consistent, compatible spectra over an extended period of time.

### 2.2. Spectral acquisition

NMR and MIR spectra were collected for each of the purchased oils and the prepared mixtures. The order in which samples were presented to the spectrometer was randomized within each batch. Before and between spectral acquisitions, all samples were stored in the dark at laboratory ambient temperature (air-conditioned, 21°C). All spectral analysis was completed in advance of the “Best

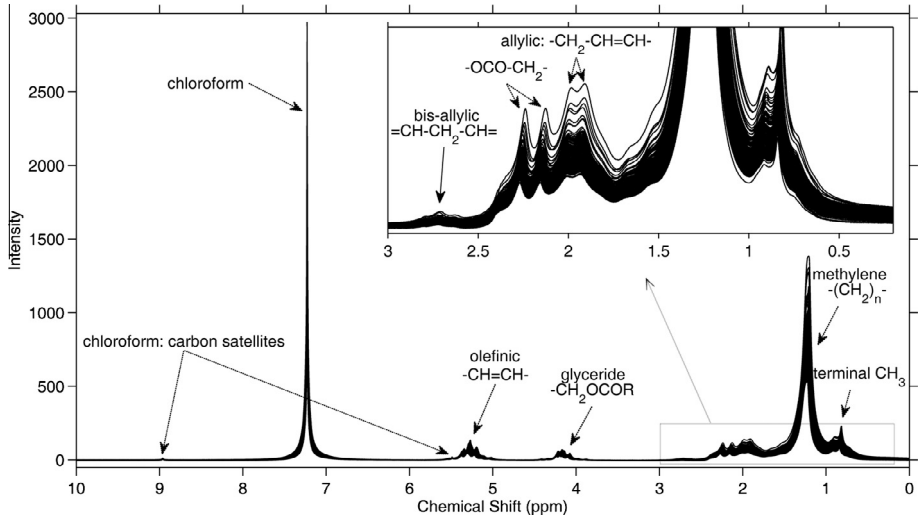



**Fig. 2.** Annotated 60 MHz $^1$H NMR spectra for 20 extra virgin olive oils, 10 hazelnut oils and 144 olive oil–hazel nut oil mixtures, as set out in Table 1. The inset shows an expansion of the chemical-shift region 0.2–3 ppm.

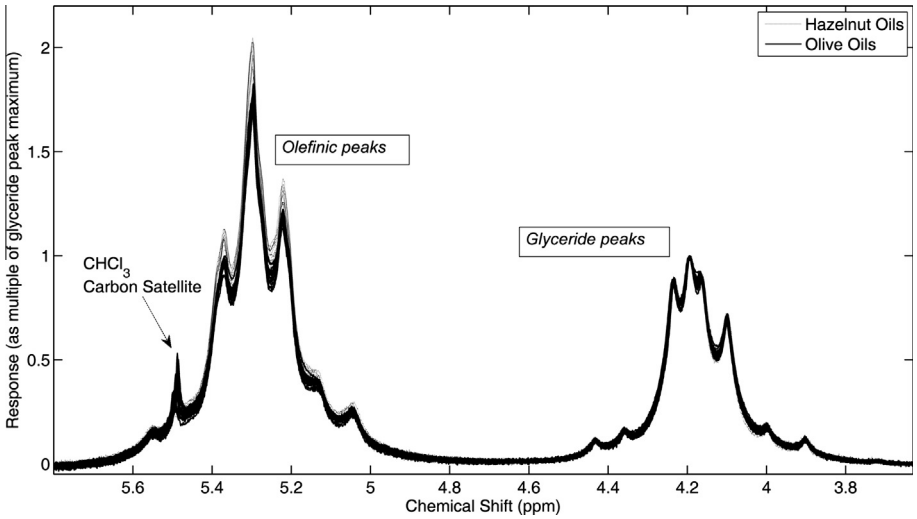



**Fig. 3.** Expansion of the olefinic and glyceride region (3.6–5.8 ppm) for 60 MHz $^1$H NMR spectra of the pure olive and pure hazelnut oils only. The spectra have been aligned and the spectral intensity scaled to the glyceride peak maximum. Hazelnut oil gives increased olefinic peak intensities.

Before" dates given on the labels. Instrumental parameters were selected such that the acquisition time per sample was similar, at around 10 min, by both techniques. Note that, for the MIR analysis, this includes acquisition of a background in advance of each sample spectrum.

For NMR analysis, approximately 2 mL of each sample was introduced into a standard 5 mm disposable NMR tube, along with approximately 2 mL of non-deuterated chloroform (Sigma Aldrich 288306-100ML). The addition of chloroform serves two purposes: first, it reduces the viscosity of the oils, leading to sharper spectral peaks; and, second, it provides a reference ppm value, allowing a chemical-shift scale to be properly attached to each spectrum. Due to its high volatility, some evaporation of the chloroform fraction in advance of spectral analysis was unavoidable, despite the use of stoppered NMR tubes. The resultant concentration variance impacts on the relative magnitude of the oil and chloroform signals, but this can be readily dealt with using standard data normalization techniques.

NMR $^1$H spectra were acquired on a Pulsar low-field spectrometer (Oxford Instruments, Tubney Woods, Abingdon, Oxford, UK) operating at 60 MHz. The sample temperature inside the spectrometer was 37°C. In all cases, 16 free induced decays (FIDs) were collected from each sample. No resolution enhancement methods were applied. The FIDs were Fourier-transformed, linearly aligned using the chloroform peak as a reference, and co-added and manually phase-corrected to present a single frequency-domain spectrum from each sample.

MIR spectra were acquired on a Nicolet MagnaIR 860 FTIR spectrometer equipped with a mercury-cadmium-telluride detector (Thermo Scientific, Hemel Hempstead, Hertfordshire, UK) and a GoldenGate single-reflection diamond attenuated-total-reflectance (ATR) sampling unit (Specac Ltd, Orpington, Kent, UK). To obtain each absorbance spectrum, a single-beam background of the clean, dry ATR crystal was first collected. Next, approximately 1 mL of sample was applied to the ATR crystal using a transfer pipette, and a single-beam spectrum collected of the sample. This was then converted to an absorbance spectrum using the background. All single-beam spectra were acquired using the same conditions: the spectral range was 600–4000 $cm^{-1}$, the nominal resolution 4 $cm^{-1}$, and 512 interferograms were co-added before Fourier transformation. After collecting each sample spectrum, the ATR plate was cleaned *in situ* by scrubbing with tissue and ethanol.

Cleanliness was verified by examining in real-time the single beam of the empty ATR crystal, before initiating collection of the following background.

### 2.3. Chemometric analysis

All data visualization and processing was carried out in Matlab (The Mathworks, Cambridge, UK). Scripts for double cross-validated Partial Least Squares (PLS) regression following the method of Smit et al. [33] were written in-house. In advance of chemometric analysis, the datasets were pre-processed as follows.

#### 2.3.1. NMR

Regions of empty baseline were discarded; the region around the chloroform reference peak was discarded, and the data were re-aligned by simple linear sideways shifting using the glyceride-peak maxima as the reference point. Each resulting spectrum was normalized to unit integrated area.

#### 2.3.2. MIR

Analysis was carried out on the "fingerprint" region (600–1800 $cm^{-1}$) only, as this corresponds to a window of high optical throughput for the diamond ATR crystal. Spectra were truncated to 2500 data points in this region, baseline corrected at 1800 $cm^{-1}$, and normalized by setting the integrated area of each spectrum to unity.

## 3. Results and discussion

### 3.1. NMR spectra: assignment and visual inspection

The complete set of spectra from the 'pure' oils and oil mixtures is shown in Fig. 2. The chloroform component of each sample gives rise to a prominent resonance at a chemical shift of 7.23 ppm, and this peak was used to align the spectra shown here. Also, arising from chloroform are two small peaks just visible at ~9.0 ppm and ~5.5 ppm, which are the two $^{13}C$ satellite resonances.

From consideration of higher field NMR spectra in the literature [17,30–32,34] various other resonances can be assigned to specific chemical groups. In particular, the peaks at ~4.2 ppm arise from $^{1}H$ nuclei attached to carbon at positions 1 and 3 on the glycerol backbone. This is a useful group of peaks, as it provides a measure of the overall glyceride concentration in the sample, and can be employed as a reference signal. The set of peaks at ~5.2 ppm arises largely from the $^{1}H$ nuclei attached to carbons involved in a double bond, usually referred to as olefinic. This signal is related to the total number of unsaturated bonds in a triglyceride, regardless of whether these are located within mono-unsaturated or poly-unsaturated chains. The very small signals at ~2.7 ppm arise from bis-allylic protons from the $-CH_2-$ group located between pairs of unsaturated bonds and thus provide a measure of the number of poly-unsaturated fatty acid chains present in the sample.

In Fig. 2, it is difficult to discern differences between the oil types simply by visual inspection. However, closer scrutiny of certain spectral regions reveals clear, systematic differences between olive oils and hazelnut oils. Fig. 3 shows the region 3.6–5.8 ppm in the spectra from the pure oils only (20 x olive, 10 x hazelnut). This region contains the olefinic and glyceride groups of resonances. The data are shown scaled so that the height of the glyceride-peak maximum is equal to unity in all spectra. It can be seen that the hazelnut oils all exhibit larger olefinic peaks than the olive oils. This is consistent with the literature; studies show that hazelnut oils typically contain slightly more unsaturated fatty acids than olive oils [9,17,35].

Fig. 4 shows an expansion of the large methylene resonance at ~1.3 ppm, again for the pure oils only. For clarity, the spectra are vertically offset with respect to one another, with an additional offset separating the data from the two oil types. Notwithstanding some variation within each oil type reflecting the fact that these are natural products, there is a clear, consistent difference between the shape of this peak for hazelnut oils and olive oils. This large feature is a series of overlapping peaks that can be assigned to $-CH_2-$ groups in a fatty acid chain more than two bonds away from a double bond or the ester coupling the glyceryl backbone and the fatty acid chains [27]. The overall shape represents a distribution of chemical shifts resulting from subtle differences in the $^{1}H$ chemical shift along the fatty acid chain: the presence of double bonds in a chain shifts the resonance of these protons to a slightly higher chemical shift. Thus, these observations suggest that the olive oils and hazelnut oils differ systematically with regard to their unsaturated contents.

### 3.2. NMR spectra: quantitative data analysis

The visible differences between the spectra of the pure oils suggest that simple data analysis based on internal ratios of integrated peak areas may prove useful. Fig. 5(a) summarizes the ratios of the

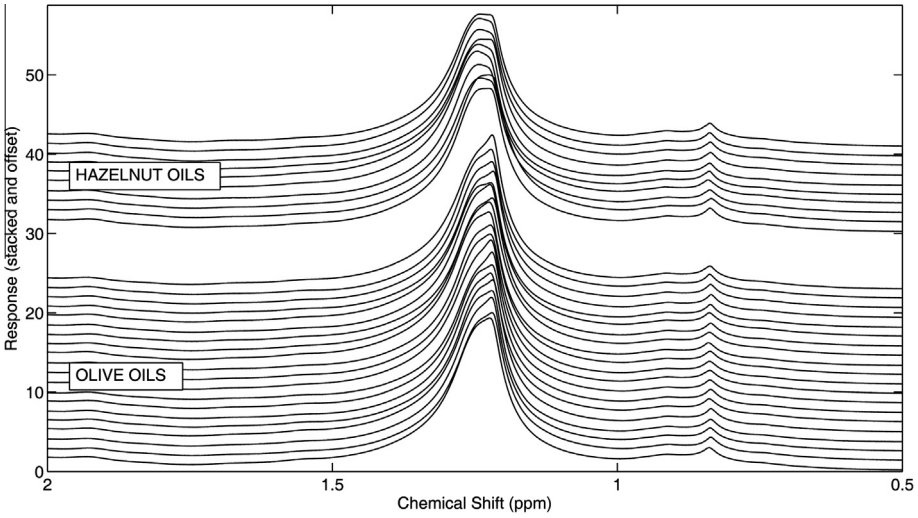



**Fig. 4.** Expansion of the region 0.5–2 ppm for 60 MHz $^{1}H$ NMR spectra of the pure olive and pure hazelnut oils only. The spectra have been aligned and the spectral intensity scaled to the glyceride peak maximum. For clarity, the spectra are offset relative to one another, with an additional offset between the olive oil and hazelnut oil groups. The central peak dominating the figure is the methylene peak at ~1.3 ppm. For hazelnut oil, the methylene peak shape noticeably differs from that of olive oil and is shifted towards a higher chemical shift.

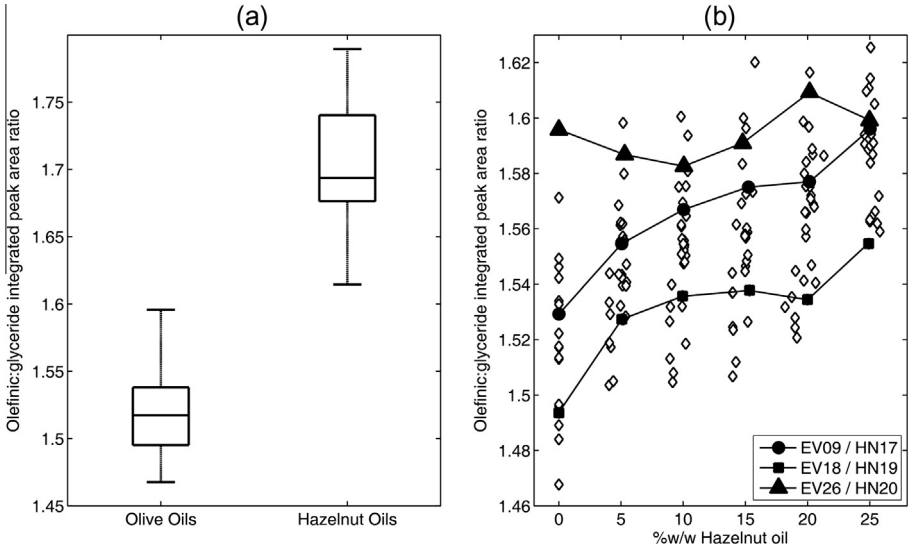



**Fig. 5.** Ratios of NMR olefinic to glyceride integrated peak areas summarized by a box-plot for: (a) the pure olive and hazelnut oils, and (b) the olive oil and hazelnut oil mixtures plus pure olive oils. In (a), for each box, the central mark is the median, the lower and upper edges of the box are the 25th and 75th percentiles, and the whiskers extend to the most extreme data points. In (b), the solid symbols indicate three representative sets of oil combinations to highlight the systematic link between composition and area ratio. The codes EVxx and HNxx identify the oil samples (Table 1).

integrated olefinic and glyceride peaks, calculated for each of the pure oils. The median ratios from each group differ significantly ($p_{Kruskal\text{-}Wallis} = 10^{-5}$). Furthermore, this single internal ratio alone appears to be sufficient to distinguish completely between the 20 olive oils (for which the ratio < 1.6) and the 10 hazelnut oils (for which the ratio > 1.6).

The same ratio was calculated from the remaining 144 mixture spectra, and is plotted against the quantity of hazelnut oil present in each sample in Fig. 5(b). The ratios for the pure olive oils are also included on this plot (although data from the pure hazelnut oils are omitted for clarity). Although much of the overall variability in the olefinic:glyceride peak area ratios can be attributed to natural variation within each oil type, there is also a systematic relationship with mixture composition: greater hazelnut oil levels are associated with larger olefinic:glyceride ratios. This is best illustrated by examining the results from individual series of mixtures prepared from discrete olive/hazelnut combinations. Fig. 5(b) highlights three of these (30 such series are present in the dataset as a whole, see Table 1). Simple linear regression of the %w/w hazelnut oils onto the area ratios was carried out for each series separately as well as the whole data collection (in all cases omitting the hazelnut oils to avoid excessive leverage) and the results are shown in Table 2. The large majority of the discrete series yield significant ($p < 0.01$) regressions with R-squared values >0.9. The median of the residual standard deviations is 6.5%w/w. The regression across the entire data collection also yields a residual standard deviation of 6.5%w/w. The 95% ($2\sigma$) confidence interval in prediction is thus ±13%w/w, which can be taken as an approximate indication of the limit of detection (LOD) of hazelnut oil in an olive oil by this approach.

By way of comparison, using 500 MHz $^1$H NMR to examine olive oil and hazelnut oil mixtures, Fauhl et al. [15] were able to recognize as suspicious adulteration of olive oil with 25% or more hazelnut oil by applying discriminant analysis to internal ratios. In a different study, Mannina et al. [14] used 600 MHz $^1$H NMR and five internal signal intensity ratios combined with PCA to detect hazelnut oil adulteration in olive oil at the 10% level. They highlighted the limitation of their approach at lower field strengths: even at 400 MHz, the specific peak due to squalene, an important discriminant in their methodology due to its relatively high abundance in olive oil compared to hazelnut oil, is lost.

**Table 2**
Summary of regressions for each olive hazelnut series, and for the entire set collectively

| Olive/ hazelnut combination | | *r*-squared value | *p*-value | Standard deviation of residuals |
|---|---|---|---|---|
| EV01 | HN03 | 0.94 | <0.01 | 9.3 |
| EV02 | HN03 | 0.91 | <0.01 | 11.6 |
| EV04 | HN03 | 0.86 | 0.01 | 14.2 |
| EV05 | HN07 | 0.73 | 0.03 | 19.5 |
| EV06 | HN07 | 0.96 | <0.01 | 7.7 |
| EV08 | HN07 | 0.94 | <0.01 | 9.2 |
| EV09 | HN17 | 0.99 | <0.01 | 3.1 |
| EV10 | HN17 | 0.97 | <0.01 | 5.7 |
| EV11 | HN17 | 0.97 | <0.01 | 6.2 |
| EV12 | HN17 | 0.96 | <0.01 | 6.6 |
| EV13 | HN17 | 0.96 | <0.01 | 6.7 |
| EV14 | HN17 | 0.99 | <0.01 | 3.7 |
| EV15 | HN17 | 0.98 | <0.01 | 5.0 |
| EV16 | HN17 | 0.95 | <0.01 | 7.4 |
| EV18 | HN19 | 0.98 | <0.01 | 4.7 |
| EV26 | HN20 | 0.27 | 0.29 | 7.9 |
| EV26 | HN23 | 0.02 | 0.79 | 9.1 |
| EV26 | HN24 | 0.17 | 0.42 | 8.7 |
| EV27 | HN21 | 0.78 | 0.02 | 4.5 |
| EV27 | HN22 | 0.80 | 0.02 | 4.2 |
| EV27 | HN25 | 0.90 | <0.01 | 3.0 |
| EV28 | HN22 | 0.88 | 0.01 | 3.2 |
| EV28 | HN24 | 0.68 | 0.04 | 5.3 |
| EV28 | HN25 | 0.90 | <0.01 | 3.0 |
| EV29 | HN20 | 0.55 | 0.09 | 6.4 |
| EV29 | HN22 | 0.27 | 0.29 | 8.1 |
| EV29 | HN23 | 0.96 | <0.01 | 1.8 |
| EV30 | HN21 | 0.07 | 0.61 | 9.0 |
| EV30 | HN22 | 0.53 | 0.10 | 6.5 |
| EV30 | HN23 | 0.27 | 0.29 | 8.0 |
| All EV and EV/ HN mixtures | | 0.35 | <0.01 | 6.5 |

### 3.3. Chemometric analysis of NMR data

As demonstrated in sub-section 3.2 above, straightforward univariate analysis of internal peak-area ratios (in this instance, olefinic-to-glyceride peak areas) is sufficient to provide a reasonable

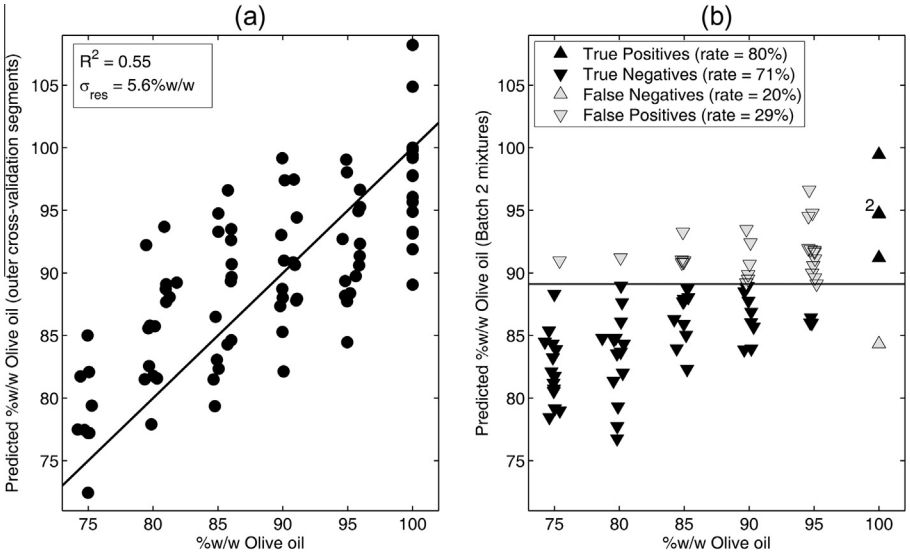



**Fig. 6.** Results of chemometric analysis on NMR data. PLS1 predictions *versus* actual olive oil content for: (a) the Batch 1 olive oils and mixture samples, where $R^2$ indicates the squared correlation between the predicted (from the outer cross-validation segments) and actual %w/w olive oil, and $\sigma_{res}$ the standard deviation of the residuals, and (b) applying the Batch 1 model (using variance-scaling and 12 PLS factors) to the Batch 2 data. In (b), the '2' for 100% actual olive oil denotes two overlapping data points. Plot (b) reports the outcome of testing for pure olive oil at the 95% confidence threshold: the upright filled triangles are the 80% true 'pure olive oil' hits and the upside down filled triangles are the 71% correct assignments of samples as 'impure'.

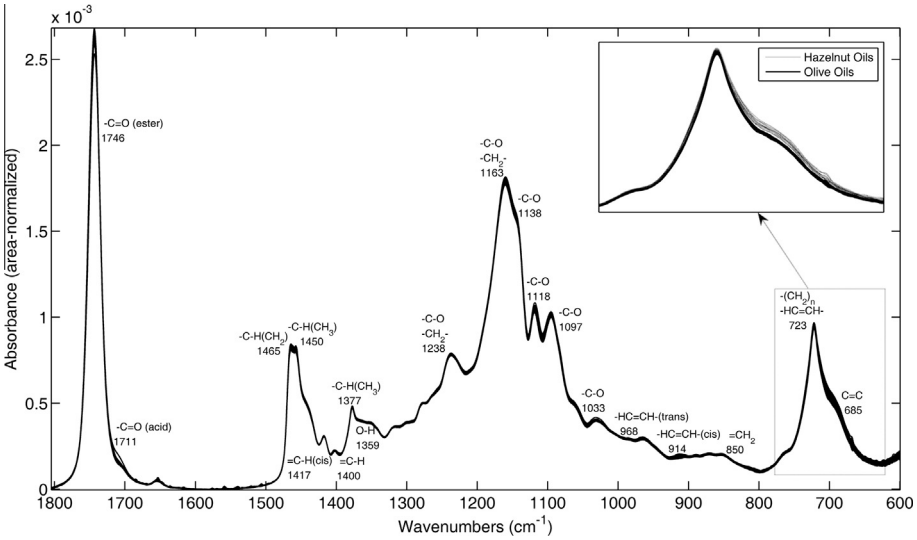



**Fig. 7.** Mid-infrared spectra of pure olive oils and hazelnut oils. The data has been baseline corrected and area normalized. The annotation is based on Lerma-García et al. [27]. The inset shows the region 780–640 $cm^{-1}$, a part of the spectrum where the differences between olive and hazelnut oils are more pronounced.

screening tool, with an LOD of hazelnut oil in olive oil of ~13%w/w. However, it is also common practice to use chemometric methods to analyze NMR data, as chemometrics can offer improved quantitative outcomes by making use of information content from across the whole spectrum [19]. In preparation for the multivariate modelling, the original spectra as shown in Fig. 2 were treated as described in Section 2 above. The final pre-treatment step, normalization to unit integrated area, was included because the variance in overall spectral intensity arising from differential chloroform evaporation would otherwise present an additional source of unwanted and likely confounding variation. None of these pretreatments involve any kind of statistical modelling, so they were applied directly to both Batch 1 and Batch 2 data.

PLS regression was used to model the %w/w of olive oil present in each sample. Model development was carried out using data from the Batch 1 olive oils and mixtures only (again omitting the pure hazelnut oils to avoid excessively leveraging the regressions). The algorithm used was non-orthogonalized PLS with one dependent variable, employing variance scaling and a double cross-validation procedure [33]. This involves apportioning the data into $m$ cross-validation segments or blocks, and utilizing these in a cyclical model/validate procedure. In the "inner" cross-validation cycle, $m-1$ of the blocks are used to develop a PLS model via leave-block-out cross-validation: each of the $m-1$ blocks acts as a test segment in turn, producing a set of cross-validation results, from which an optimal model is chosen. In the "outer" cross-validation cycle, the

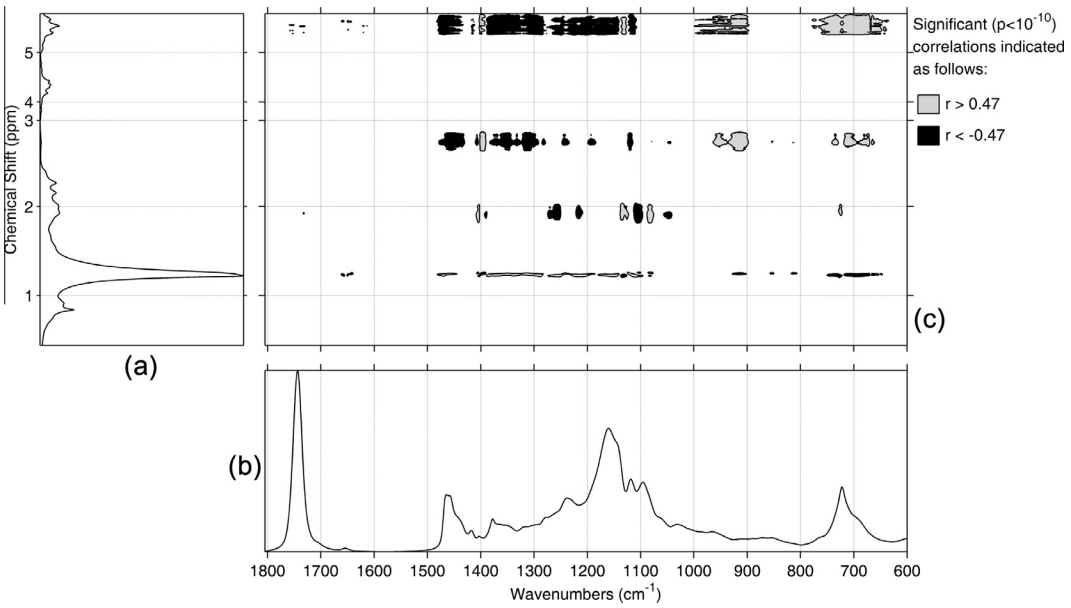



**Fig. 8.** Simplified heat-map showing Pearson's correlation coefficient *r* between NMR and MIR data. Panel (a) is NMR data (Fig. 2) and panel (b) is MIR data (Fig. 7). Regions of the NMR chemical-shift axis have been excised to eliminate blank areas and to create a more compact figure. On correlation panel (c), to account for multiple testing, only correlations for which $p < 10^{-10}$ are indicated. Light regions are significant positive correlations ($r > 0.47$) and dark regions are significant negative correlations ($r < 0.47$). The blank background corresponds to areas of no significant (anti)correlation at the $p < 10^{-10}$ level.

optimal model from the inner cycle is applied to the *m*th block. The complete outer cycle thus returns results from all of the *m* blocks, each acting as an independent test segment in turn. For application to further, unseen data (in the present work, from the Batch 2 samples), a single "final" model is created using all *m* blocks, again using conventional leave-block-out cross-validation.

Double cross-validation is a rigorous procedure which guards against overfitting, provided it is implemented with appropriate choices of blocks. In our implementation, the *m* cross-validation segments corresponded to the four distinct hazelnut oils used to prepare the Batch 1 mixtures (see Table 1) (i.e., a "leave-adulterant-out" strategy). This ensured that, at all times, samples acting as test segments in both the "inner" and "outer" cycles are, in terms of their constituent-oil composition, truly independent of the samples involved in model development.

The prediction results for the outer cross-validation segments are shown in Fig. 6(a), plotted against the actual olive oil content. The standard deviation of the residuals is 5.6%w/w. This corresponds to a 95% confidence interval of ±11.2%w/w in predicting the olive oil content and, in line with this, it is seen that the predictions for the pure olive oils all lie within the range 100 ±11.2%w/w. This naturally leads to the possibility of screening olive oils for the presence of hazelnut oil, accepting as "authentic" only those samples for which the predicted olive oil content falls within this range.

The 12-factor PLS model developed using the Batch 1 data was applied to the entire collection of Batch 2 spectra. The predictions obtained are plotted against the actual olive oil content in Fig. 6(b). The pattern of results is very similar to that obtained from the Batch 1 data, showing that the model has good general applicability when challenged with unseen, unrelated samples. It further indicates that the bench-top NMR instrument has exhibited sufficient stability over the time interval of several weeks that elapsed between the acquisition of spectra from the Batch 1 and Batch 2 samples.

The lower boundary of the 95% confidence interval for olive oils is marked on Fig. 6(b), to be used as a threshold for determining whether or not an individual sample is "authentic". The symbols are coded by the outcome of this binary classification test. Also, the Batch 2 hazelnut oils all resulted in true negatives, although these are omitted from the plot for clarity. The choice of confidence level is of course arbitrary and, in practice, will depend on whether one is more interested in minimizing false positives (type I errors) than false negatives (type II errors). Furthermore, the error rates depend on the nature of the samples with which the model is challenged: if Batch 2 had contained only the pure oils, say, then the true positive rate (sensitivity) and true negative rate (specificity) would have been found to be 100%. Nevertheless, the error rates calculated from the full range of Batch 2 samples, as presented in Fig. 6(b), provide a snapshot of the performance of the simulated authentication protocol, which can be informatively compared like-for-like with the analogous IR analysis presented in the following sub-section.

### 3.4. Comparison with IR data

Fig. 7 shows the collection of MIR spectra from the pure olive and hazelnut oils, baseline corrected and area normalized on the fingerprint region only. The peaks are annotated using nominal peak centers, after the assignment table given in Lerma-Garcia et al. [27]. As noted by other authors {e.g., [26]}, the MIR spectra of hazelnut oils and olive oils are found to be very similar to one another. However, in good-quality spectra obtained with low instrumental noise, it is possible to discern systematic differences between the two oil types, as shown by the expanded region of the spectral window shown in the upper right of Fig. 7.

Since the IR spectra were acquired from precisely the same collection of samples that were used in the NMR analysis, an avenue exists for exploring correlations between the two datasets. Fig. 8 shows a simplified heat-map showing the significant pairwise correlations between points in the NMR and MIR datasets. The lower of the three panels in Fig. 8 shows the mean of the pre-treated MIR data, while the left-hand panel shows the mean of the pre-treated NMR spectra (with sections of empty baseline removed to aid clarity). These are aligned with the axes scales of the main heat-map panel, which indicates regions where the Pearson's correlation coefficient *r* is found to be significant. The correlation heat-map

was interpreted on the basis of NMR assignments in Alonso-Salces et al. [30] and MIR assignments in Lerma-García et al. [27]. Some alternative heat-maps (not shown) were also considered, prepared from differently pre-treated spectra (for example, MIR data normalized to the height of the ester peak at ~1750 $cm^{-1}$; NMR data normalized to the height of the glyceride peaks).

It is worth emphasizing that the features seen in this panel arise from strong positive or negative correlations, and may bear little relation to size of the features in either the MIR or NMR spectra. For example, the bis-allylic contributions to the NMR spectra at ~2.8 ppm are so small as to be almost invisible on the NMR spectral plot, yet they generate significant heat-map features. This is because even a very low intensity signal at a particular NMR chemical shift that varies across the set of oil samples in the same way as the intensity at some MIR wavenumber will give rise to a light grey positive correlation feature (or a dark anti-correlation region, if the variation in the NMR spectra is inversely related to that in the MIR).

Central to interpreting the correlation plot is the observation that in the MIR, peaks attributed to double bonds dominate at lower wavenumbers (<1050 $cm^{-1}$), and these rise with increasing hazelnut oil content. This is consistent with an increase in unsaturated fatty acid content at the expense of saturated fatty acids relative to olive oil, as might be expected from some literature (e.g., [17], though acknowledging the wide ranges of measured fatty-acid content, e.g., [9,36–38]). Since the data are area normalized, a rise across one wavenumber region must be accompanied by a fall elsewhere to maintain constant area. Broadly speaking, our MIR data show a rise in intensity with increasing hazelnut oil concentration for wavenumbers less than ~1050 $cm^{-1}$ along with a decrease in intensity for wavenumbers greater than ~1090 $cm^{-1}$.

The glyceride peaks at ~4.2 ppm do not give rise to any features in the correlation panel. This is because the data pre-treatment results in these peaks having approximately constant magnitude across the dataset, since the glyceride content is itself unvarying across the collection of what are largely tri-glyceride samples. Since there is little variation, one would not expect to see any significant correlation. In a similar way, other 'blanks' on the heat-map that are linked to the unvarying glyceride include the ~5.1 ppm (>CHOCOR) and the peaks at ~2.3 ppm and ~2.1 ppm (–OCO–$CH_2$–). Likewise, the methyl group –$CH_3$ at ~0.8 ppm does not give rise to any features on the correlation plot, since there is little change in this feature across the range of oil samples.

The remaining NMR signals are all involved in significant correlations with the MIR data, occurring broadly in four horizontal 'Bands' corresponding to different ppm ranges. Moving up the heat-map in order of increasing ppm values, these are as follows.

### 3.4.1. Band 1 (~1.3 ppm)

The NMR signal arises from –$CH_2$ groups more than two chemical bonds away from a functional group, such as –C=C– (or the ester coupling group linking the fatty acid chain to the triglyceride backbone). The number of such 'conventional' –$CH_2$ groups will vary with the number of double bonds in the fatty acid chains, so the positive correlation with 1490–1450 $cm^{-1}$ in the MIR can be assigned to vibrations of –C–H (of $CH_2$), since this will track the decrease in –$CH_2$ bonds with increasing hazelnut oil content. There is a small anti-correlation with 1405–1395 $cm^{-1}$ that can be assigned to MIR vibrations of =C–H. Note that there are positive correlation features at the same location in the other three Bands. The long feature (correlated with 1390–1150 $cm^{-1}$) arises from the overall drop in MIR intensity that correlates with the falling –$CH_2$ NMR signal intensity. At 1150–1140 $cm^{-1}$ is a small anti-correlation feature (with heat-map counterparts in Bands 2 and 4) that is suggestive of a double-bond feature, presumably either =C–H or –CH=CH–, though the MIR annotation table does not list a candidate functional group or vibration. The small anti-correlation with 940–900 $cm^{-1}$ is related to vibrations involving –CH=CH– (note alignment with features in Bands 3 and 4), as is the extended anti-correlation with 750–650 $cm^{-1}$.

### 3.4.2. Band 2 (1.8–2.1 ppm)

The relevant NMR spectral feature for this Band is –$CH_2$–CH=CH (the allylic protons). Progressing from olive oil through mixtures to hazelnut oils is likely to be associated with a relatively weak increase in this signal, since both olive oil and hazelnut oil contain similar high levels of monounsaturated oleic acid. The positive correlation with 1410–1400 $cm^{-1}$ is due to MIR vibrations involving =C–H. Another positive correlation, with 1130–1115 $cm^{-1}$, has counterparts in Bands 1 and 4, and is likely to be

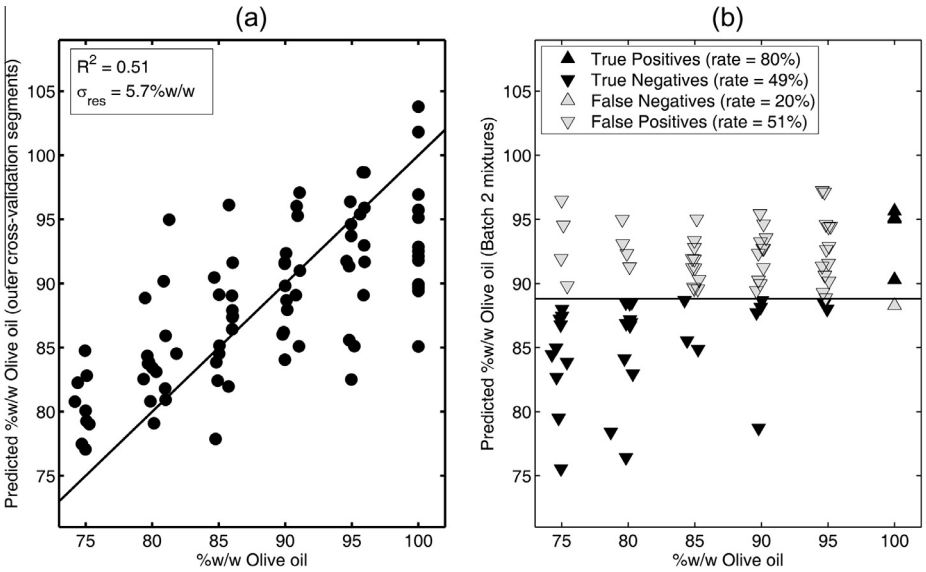



**Fig. 9.** Results of chemometric analysis on MIR data. PLS1 predictions *versus* actual olive oil content for: (a) the Batch 1 olive oils and mixture samples, where $R^2$ indicates the squared correlation between the predicted (from the outer cross-validation segments) and actual %w/w olive oil, and $\sigma_{res}$ the standard deviation of the residuals, and (b) applying the Batch 1 model (using variance-scaling and 8 PLS factors) to the Batch 2 data. Plot (b) reports the outcome of testing for pure olive oil at the 95% confidence threshold: the upright filled triangles are the 80% true 'pure olive oil' hits and the upside down filled triangles are the 49% correct assignments of samples as 'impure'.

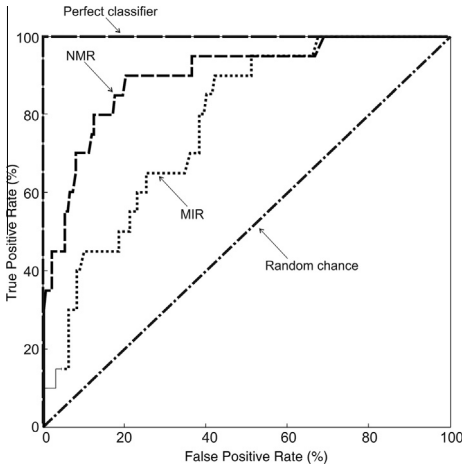



**Fig. 10.** ROC curves that show the true positive rate *versus* the true negative rate (at all possible threshold values) obtained from all NMR and MIR prediction outcomes (Batch 2 data and Batch 1 outer segment predictions). Also marked are theoretical ROC curves that would be obtained from "perfect" and "random" classifiers.

double-bond related. A further positive correlation feature, with 1095–1085 $cm^{-1}$ lacks counterparts in Bands 3 and 4, and awaits interpretation. There is also uncertainty surrounding the region 960–900 $cm^{-1}$, which shows blank, yet would appear ripe for a positive correlation based on other features both in this Band and in corresponding locations in Bands 3 and 4.

*3.4.3. Band 3 (2.6–2.9 ppm)*

The relevant NMR spectral feature is =CH–$CH_2$–CH= (the bis-allylic protons). Although the NMR feature is low intensity, it can be expected to vary relatively strongly with increasing hazelnut oil content in oil mixtures. The anti-correlation with 1480–1430 $cm^{-1}$ arises from the –C–H (of $CH_2$). The correlation feature at 1400–1390 $cm^{-1}$ is attributed to vibrations involving =C–H, and two further features (at 960–900 $cm^{-1}$ and 740–660 $cm^{-1}$) are associated with –HC=CH–.

*3.4.4. Band 4 (5.2–5.6 ppm)*

The relevant NMR spectral feature for this band is largely due to –CH=CH– (olefinic protons). Mirroring Band 3, there is an anti-correlation with 1480–1430 $cm^{-1}$ associated with vibrations of –C–H (of $CH_2$), and a correlation at 1400–1390 $cm^{-1}$ associated with =C–H. The large anti-correlation feature stretching across 1390–1150 $cm^{-1}$ is the converse of the correlation feature in Band 1, and arises from the relative suppression of this part of the MIR spectrum as the low wavenumber end increases with hazelnut oil content. There is a double-bond-related correlation feature at ~1130 $cm^{-1}$ with counterparts in Bands 1 and 2. Finally, two large correlation features (with 1000–900 $cm^{-1}$ and 780–640 $cm^{-1}$) are associated with vibrations of –HC=CH–.

To summarize, the correlation plot reveals an overall pattern of increasing double-bond content and commensurate reduction in –$CH_2$– content with increasing hazelnut oil content across the dataset. The NMR contributes to a more refined interpretation of the MIR spectrum. In return, the MIR spectrum data help to establish relationships between different features in the NMR spectrum. Even small-intensity signals and small changes in intensities that might otherwise be ignored or overlooked can exhibit strong effects if they are sufficiently correlated, as evidenced by the small but highly significant bis-allylic NMR signals. Finally, analysis of the heat-map strongly suggests the existence of a double-bond-related feature in the MIR spectrum at around ~1130 $cm^{-1}$, that has not previously been annotated in the literature, residing between two documented vibrations of –C–O at 1118 $cm^{-1}$ and 1138 $cm^{-1}$.

*3.5. Chemometric analysis of MIR data*

Unlike the NMR spectra, there are no isolated peaks in the MIR spectra that can be usefully integrated and related to the sample composition via simple internal ratios. However, an analogous PLS1 analysis (employing variance scaling and double cross-validation) was developed using the Batch 1 MIR data, and the resulting model applied to the Batch 2 spectra. The results are presented in Fig. 9, as a direct analogue of Fig. 6. From Fig. 9a, it is seen that the $R^2$ values and the standard deviations of the residuals are broadly comparable to those for the NMR data. Fig. 9b summarizes the performance of the model as applied to the Batch 2 samples. As for the NMR analysis, a threshold is marked on the plot corresponding to the lower boundary of the 95% confidence interval for olive oils, and the symbols coded according to the outcome of the binary test for oil type. Here, it is seen that, although the true positive rate (sensitivity) is 80%, equal to that for the NMR analysis, the true negative rate (specificity) is substantially poorer at 49% (compared with 71% for the NMR dataset). This implies that NMR is a more effective tool upon which to base a screening protocol than MIR, producing a marginally better calibration model and, crucially, a substantially better outcome at the application stage.

A further means of comparison of the NMR and MIR performances is through receiver-operating characteristic (ROC) curves, which compare the true positive and true negative rates for not just one but a range of values of the binary classifier threshold. ROC curves were constructed from the complete sets of prediction values (the Batch 2 predictions along with the Batch 1 outer cross-validation segments) obtained by each technique, and these are compared in Fig. 10. Also marked on the plot are the "perfect classifier" and "random chance" ROC curves. It is seen that the NMR ROC is closer to that of the perfect classifier than the MIR curve for virtually all threshold values. We can conclude that, for the simulated authentication problem studied, and the experimental protocol and acquisition conditions, 60 HMz NMR has outperformed MIR by a substantial margin.

## 4. Conclusions

In this work, we explored the potential of a new 60 MHz $^1$H NMR bench-top spectrometer, by applying it to a simulated problem of detecting olive oil adulterated with hazelnut oil. The spectra obtained from the pure oils and mixtures were as anticipated, in the sense that their spectral features corresponded to low-field versions of known high-field outputs reported elsewhere. The quality of the spectra was such that qualitative differences between the two oil types could be seen, consistent with a higher unsaturated component of hazelnut oil compared to olive oil, as anticipated from published values. Furthermore, even a single internal ratio of two peak areas, the olefinic-to-glyceride peaks, was enough to detect hazelnut oil adulteration at the ~13%w/w level for the given set of oils. In contrast to high-field $^1$H NMR data, where other candidate peaks exist, there are no other peak areas that can readily be used in this way due to a lack of clearly-defined isolated peaks, although the bis-allylic peaks may be accessible with care.

The scarcity of isolated peaks encourages the application of chemometric techniques sensitive to the entire spectrum. By applying chemometric methods to one batch of the data, and applying the model developed to a second, independent set of oils and mixtures, we have shown that it is possible to determine the level of olive oils and hazelnut oils in binary mixtures to ±11.2%w/w at 95% confidence. In turn, this implies that bench-top 60 MHz $^1$H NMR is useful as a means of screening for pure olive oil, in that it should be capable of detecting adulteration with hazelnut oil at levels in excess of ~11%w/w.

The widely-available technique of FTIR spectroscopy operating in the MIR is comparable to low-field bench-top NMR in terms of instrument cost, experiment duration and simplicity. This invites comparison, and we have carried out a direct correlation between bench-top NMR and FTIR, which has aided in annotating the MIR spectral plots and suggested the existence of a double-bond-related feature at ~1130 $cm^{-1}$, previously undocumented.

Comparing the two methods as screening tools, we have demonstrated that, for the given instrumentation, samples, and a like-for-like chemometric analysis of a simulated adulteration problem, the 60 MHz $^{1}H$ NMR has delivered comparable sensitivity and better specificity than FTIR.

## Acknowledgements

The authors acknowledge the support of the UK's Technology Strategy Board (Project number 101250) and the Biotechnology and Biological Sciences Research Council (Grant Number BBS/E/F/00042674).

## References

[1] 'Pulsar', Oxford Instruments, www.oxford-instruments.com, 2013.
[2] D.S. Wishart, Characterization of biopharmaceuticals by NMR spectroscopy, TrAC, Trends Anal. Chem. 48 (2013) 96–111.
[3] H. Martens, T. Naes, Multivariate Calibration, Wiley-Blackwell, New York, 1991.
[4] T.M. Alam, M.K. Alam, Chemometric analysis of NMR spectroscopy data: a review, in: G.A. Webb (Ed.), Annual Reports on NMR Spectroscopy, vol. 54, Academic Press Ltd–Elsevier Science Ltd, London, 2005, pp. 41–80.
[5] L. Eriksson, H. Antti, J. Gottfries, E. Holmes, E. Johansson, F. Lindgren, I. Long, T. Lundstedt, J. Trygg, S. Wold, Using chemometrics for navigating in the large data sets of genomics, proteomics, and metabonomics (gpm), Anal. Bioanal. Chem. 380 (2004) 419–429.
[6] A.M. Weljie, J. Newton, P. Mercier, E. Carlson, C.M. Slupsky, Targeted profiling: quantitative analysis of H-1 NMR metabolomics data, Anal. Chem. 78 (2006) 4430–4442.
[7] J.C. Moore, J. Spink, M. Lipp, Development and application of a database of food ingredient fraud and economically motivated adulteration from 1980 to 2010, J. Food Sci. 77 (2012) (1980) R118–R126.
[8] T. Mueller, Extra Virginity: The Sublime and Scandalous World of Olive Oil, Atlantic Books, London, 2013.
[9] P.L. Benitez-Sanchez, M. Leon-Camacho, R. Aparicio, A comprehensive study of hazelnut oil composition with comparisons to other vegetable oils, particularly olive oil, Eur. Food Res. Technol. 218 (2003) 13–19.
[10] UCLAF, Fight against fraud: annual report 1997, European Commission, Brussels, 1998.
[11] UCLAF, Fight against fraud: annual report 1998, European Commission, Brussels, 1999.
[12] E.N. Frankel, Nutritional and biological properties of extra virgin olive oil, J. Agric. Food Chem. 59 (2011) 785–792.
[13] A. Agiomyrgianaki, P.V. Petrakis, P. Dais, Detection of refined olive oil adulteration with refined hazelnut oil by employing NMR spectroscopy and multivariate statistical analysis, Talanta 80 (2010) 2165–2171.
[14] L. Mannina, M. D'Imperio, D. Capitani, S. Rezzi, C. Guillou, T. Mavromoustakos, M.D.M. Vilchez, A.H. Fernandez, F. Thomas, R. Aparicio, H-1 NMR-based protocol for the detection of adulterations of refined olive oil with refined hazelnut oil, J. Agric. Food Chem. 57 (2009) 11550–11556.
[15] C. Fauhl, F. Reniero, C. Guillou, H-1 NMR as a tool for the analysis of mixtures of virgin olive oil with oils of different botanical origin, Magn. Reson. Chem. 38 (2000) 436–443.
[16] D.L. Garcia-Gonzalez, L. Mannina, M. D'Imperio, A.L. Segre, R. Aparicio, Using H-1 and C-13 NMR techniques and artificial neural networks to detect the adulteration of olive oil with hazelnut oil, Eur. Food Res. Technol. 219 (2004) 545–548.
[17] G. Vigli, A. Philippidis, A. Spyros, P. Dais, Classification of edible oils by employing $^{31}P$ and $^{1}H$ NMR spectroscopy in combination with multivariate statistical analysis. A proposal for the detection of seed oil adulteration in virgin olive oils, J. Agric. Food Chem. 51 (2003) 5715–5722.
[18] L. Mannina, A.P. Sobolev, High resolution NMR characterization of olive oils in terms of quality, authenticity and geographical origin, Magn. Reson. Chem. 49 (2011) S3–S11.
[19] P. Dais, E. Hatzakis, Quality assessment and authentication of virgin olive oil by NMR spectroscopy: a critical review, Anal. Chim. Acta 765 (2013) 1–27.
[20] G. Gurdeniz, B. Ozen, Detection of adulteration of extra-virgin olive oil by chemometric analysis of mid-infrared spectral data, Food Chem. 116 (2009) 519–525.
[21] G. Gurdeniz, B. Ozen, F. Tokatli, Comparison of fatty acid profiles and mid-infrared spectral data for classification of olive oils, Eur. J. Lipid Sci. Technol. 112 (2010) 218–226.
[22] A. Rohman, Y.B.C. Man, Fourier transform infrared (FTIR) spectroscopy for analysis of extra virgin olive oil adulterated with palm oil, Food Res. Int. 43 (2010) 886–892.
[23] H.S. Tapp, M. Defernez, E.K. Kemsley, FTIR spectroscopy and multivariate analysis can distinguish the geographic origin of extra virgin olive oils, J. Agric. Food Chem. 51 (2003) 6110–6115.
[24] N. Vlachos, Y. Skopelitis, M. Psaroudaki, V. Konstantinidou, A. Chatzilazarou, E. Tegou, Applications of Fourier transform-infrared spectroscopy to edible oils, Anal. Chim. Acta 573 (2006) 459–465.
[25] N. Sinelli, E. Casiraghi, D. Tura, G. Downey, Characterisation and classification of Italian virgin olive oils by near- and mid-infrared spectroscopy, J. Near Infrared Spectrosc. 16 (2008) 335–342.
[26] B.F. Ozen, L.J. Mauer, Detection of hazelnut oil adulteration using FT-IR spectroscopy, J. Agric. Food Chem. 50 (2002) 3898–3901.
[27] M.J. Lerma-Garcia, G. Ramis-Ramos, J.M. Herrero-Martinez, E.F. Simo-Alfonso, Authentication of extra virgin olive oils by Fourier-transform infrared spectroscopy, Food Chem. 118 (2010) 78–83.
[28] V. Baeten, J.A.F. Pierna, P. Dardenne, M. Meurens, D.L. Garcia-Gonzalez, R. Aparicio-Ruiz, Detection of the presence of hazelnut oil in olive oil by FT-Raman and FT-MIR spectroscopy, J. Agric. Food Chem. 53 (2005) 6201–6206.
[29] R.M. Maggio, L. Cerretani, E. Chiavaro, T.S. Kaufman, A. Bendini, A novel chemometric strategy for the estimation of extra virgin olive oil adulteration with edible oils, Food Control 21 (2010) 890–895.
[30] R.M. Alonso-Salces, M.V. Holland, C. Guillou, H-1-NMR fingerprinting to evaluate the stability of olive oil, Food Control 22 (2011) 2041–2046.
[31] G. Knothe, J.A. Kenar, Determination of the fatty acid profile by H-1-NMR spectroscopy, Eur. J. Lipid Sci. Technol. 106 (2004) 88–96.
[32] M.D. Guillen, A. Ruiz, Edible oils: discrimination by H-1 nuclear magnetic resonance, J. Sci. Food Agric. 83 (2003) 338–346.
[33] S. Smit, M.J. van Breemen, H.C.J. Hoefsloot, A.K. Smilde, J. Aerts, C.G. de Koster, Assessing the statistical validity of proteomics based biomarkers, Anal. Chim. Acta 592 (2007) 210–217.
[34] M.D. Guillen, A. Ruiz, H-1 nuclear magnetic resonance as a fast tool for determining the composition of acyl chains in acylglycerol mixtures, Eur. J. Lipid Sci. Technol. 105 (2003) 502–507.
[35] M. Lisa, M. Holcapek, M. Bohac, Statistical evaluation of triacylglycerol composition in plant oils based on high-performance liquid chromatography-atmospheric pressure chemical ionization mass spectrometry data, J. Agric. Food Chem. 57 (2009) 6888–6898.
[36] N.K. Andrikopoulos, Triglyceride species compositions of common edible vegetable oils and methods used for their identification and quantification, Food Rev. Int. 18 (2002) 71–102.
[37] J. Parcerisa, D.G. Richardson, M. Rafecas, R. Codony, J. Boatella, Fatty acid, tocopherol and sterol content of some hazelnut varieties (*Corylus avellana* L.) harvested in Oregon (USA), J. Chromatogr. A 805 (1998) 259–268.
[38] C. Crews, P. Hough, J. Godward, P. Brereton, M. Lees, S. Guiet, W. Winkelmann, Study of the main constituents of some authentic hazelnut oils, J. Agric. Food Chem. 53 (2005) 4843–4852.