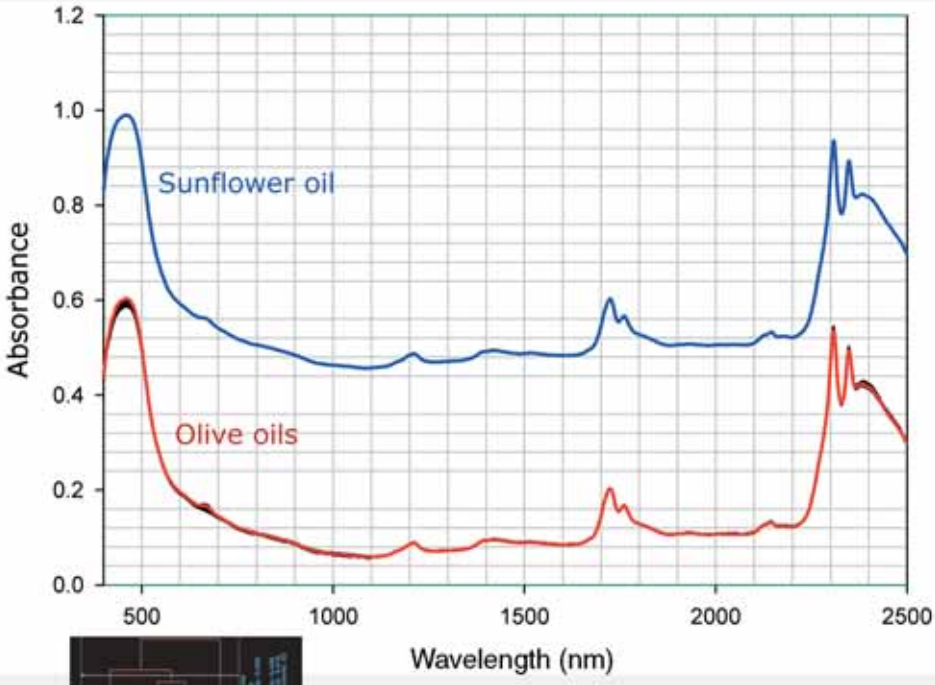
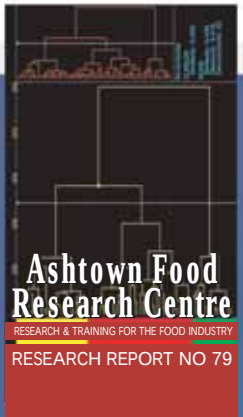
Teagasc
Agriculture and Food Development Authority

FINAL REPORT

Project RMIS No. 4907

Food Authentication using Infrared Spectroscopic Methods





Ashtown Food Research Centre

RESEARCH & TRAINING FOR THE FOOD INDUSTRY

RESEARCH REPORT NO 79

FOOD AUTHENTICATION USING INFRARED SPECTROSCOPIC METHODS

Editor-in-Chief: Dr Gerard Downey

Authors

Gerard Downey BSc PhD DSc
J. Daniel Kelly BSc PhD
Ashtown Food Research Centre,
Teagasc, Ashtown, Dublin 15

ISBN 1 84170 460 X
June 2006

Teagasc Oak Park Carlow Co. Carlow

CONTENTS

SUMMARY

Confirmation of the authenticity of a food or food ingredient is an increasing challenge for food processors and regulatory authorities. This is especially the case when an added-value claim, such as one relating to geographic origin or a particular processing history, is made on the food label. Regulatory agencies are concerned with the prevention of economic fraud while the food processor needs confirmation of such claims in order to protect a brand, the image of which could be severely damaged should an adulterated ingredient make its way into the branded food product.

To be of greatest value, any analytical tool deployed to confirm authenticity claims needs to be portable, easy to use, non-destructive and accurate. Infrared spectroscopy, near and mid-infrared, is a tool which has been demonstrated to possess these properties in a wide range of situations. While some applications in food authenticity have been reported, the work undertaken in this project was designed to explore their capabilities regarding a number of products and authenticity issues of particular interest to the Irish agri-food industry *i.e.* olive oil, honey, soft fruit purées and apple juice.

In the case of olive oil, near infrared spectroscopy in combination with factorial discriminant analysis was able to classify oils from Crete, the Peloponese and other regions of Greece with an overall correct classification rate of about 96%. A sunflower oil adulterant could be detected and quantified with a prediction error equal to 0.8% w/w; this suggests a minimum detection level of approximately 3% w/w.

Raspberry and strawberry purées were adulterated with apple purée. In the case of raspberry, a near infrared model was able to detect apple adulteration at levels >10% w/w; for strawberry, unambiguous detection was not possible below 20% w/w. Quantitation of apple content in adulterated purées was possible with a prediction error equal to 3.4 and 5.5% w/w in raspberry and strawberry respectively.

Honeys were adulterated with fructose and glucose (FG) mixtures, beet invert syrup (IB) and high fructose corn syrup (HFCS) and then analysed using near infrared spectroscopy. Regarding IB and HFCS adulterants, 90.9% and 100% of authentic and adulterated honeys were correctly classified; in the case of FG, 96% of authentic and 99% of adulterated samples were correctly identified. A parallel study involved the application of mid-infrared spectroscopy to detect honey adulteration by beet sucrose, dextrose syrup, partially-inverted cane syrup, high fructose corn syrup and solutions containing fructose and glucose. Regarding the latter adulterant, 87.9 and 96.4% of authentic and adulterated samples respectively were correctly classified. For the larger group of adulterants, a combination of classification techniques permitted identification of authentic honey and honey adulterated by beet sucrose, dextrose syrups and partially inverted corn syrup with correct classification rates of 96.2, 97.5, 95.8 and 91.7% respectively. This method was not able to unambiguously detect adulteration by high fructose corn syrup or fully-inverted beet syrup.

Apple juice was adulterated using high fructose corn syrup and a mixture of fructose, glucose and cane sucrose. Using a near infrared model for detecting either type of adulteration, a correct classification rate of 96% for adulterated juices but a lower value (86%) for authentic juices was obtained. Models specific to each adulterant achieved correct identification rates of 91 and 98% for authentic juice with sugar mixtures and high fructose corn syrup adulterants respectively; corresponding adulterated juices were accurately identified at rates of 93 and 91%. In a mid-infrared study, adulterants investigated were partially inverted cane syrup (PICS), beet sucrose (BS), high fructose corn syrup (HFCS) and a solution of fructose, glucose and sucrose (FGS). For all adulterants except FGS, the overall correct classification figures obtained were 92%; most incorrect classifications of adulterated juices involved samples adulterated at the lowest level *i.e.* 10% w/w. This method was unable to detect FGS adulterants in apple juice.

INTRODUCTION

Globalisation is increasingly a feature of the food industry. Among the benefits of this development are the wider range of ingredients and foods available to the Western consumer and the year-round availability of normally seasonal ingredients. Greater opportunities for travel to increasingly distant countries and the consequent exposure to an eclectic range of cuisines have fuelled consumer demand for non-traditional foods. Many of these foodstuffs are expensive and the consumer, food processor and regulatory agency all have an interest in ensuring their authenticity. Apart altogether from environmental concerns arising from the considerable distances which these generally perishable items must cover, the large geographical separation of producer and consumer in combination with a lack of familiarity by Western consumers with many of the foods involved raises the possibility of their economic adulteration (Lees, 2003).

Food adulteration has a long history with reports of its occurrence dating back at least as far as Roman times. The basic principle is that a high-value food or ingredient is adulterated through the addition of a lower-value substance which the consumer or food processor is unlikely to be able to detect without resort to chemical or physical analysis. This extension of such a high-value material by an inferior analogue is an easy way to increase financial return by an unscrupulous trader but it is a fraudulent practice. The types of adulteration which may arise are to a large extent food specific although certain general categories may be mentioned. These include, for example, mis-labelling of geographic region of origin (*e.g.* in fruit juice), species (*e.g.* of meat), variety (*e.g.* of coffee) and type of production process involved (*e.g.* in extra virgin olive oil).

Detection of adulteration is a complex matter and very many highly-sophisticated analytical techniques have been studied for their efficacy in identifying adulterated foods. These have included mass spectrometry, gas chromatography, high performance liquid chromatography, DNA-based

methods, isotope ratio mass spectroscopy etc. While such techniques have delivered subtle and powerful tools for the analyst, for the food processor they may not be practical options since they require lengthy analyses which must be performed in dedicated analytical laboratories by highly-skilled personnel. In contrast, infrared spectroscopic methods, specifically mid-infrared (MIR) and near infrared (NIR) spectrophotometry, are characterised by their speed, absence of any requirement for chemical reagent purchase or disposal and ease-of-use after initial method development. For these reasons, together with the fact that they do not destroy the sample being tested, both methods represent analytical approaches which may be useful as screening tools by the food industry to check the authenticity of incoming foods or ingredients.

Building on expertise developed over many years in both acquisition and mathematical interrogation of infrared spectra, this project set out to investigate the utility of mid-IR and NIR in detecting certain adulterations of ingredients of relevance to the Irish food processing industry. In addition to soft fruit, apple juice and honey which are locally produced, a typical high-value non-traditional ingredient (olive oil) was also studied. The results for each food ingredient are reported in the following sections; given the limited number of samples involved, these results are to be regarded as indicative.

NIR and MIR spectra of foods are complex and therefore multivariate statistical techniques are required to extract useful information; this information may be applied to predict composition or to classify samples on the basis of some natural grouping. The main method used for quantitation is partial least squares (PLS) regression while for discrimination, a variety of approaches including factorial discriminant analysis (FDA), *k*-nearest neighbours (*k*-NN) analysis, discriminant PLS and soft independent modelling of class analogy (SIMCA) may be used. Hierarchical cluster analysis (HCA) is a useful exploratory tool to detect clustering in a dataset. All of these techniques significantly reduce the size of a dataset in the first instance; principal component analysis (PCA) is commonly employed for this purpose and modifies the original wavelength data so that the vast majority of the

information contained in the sample spectrum is contained in, say, 20 new axes or components. Each sample spectrum may then be represented by a single point (score) on each of these new axes and the score values may be further manipulated. Partial least squares regression differs slightly from PCA in that it also uses information from any associated chemical or physical data to generate the new axes.

Discrimination between different functional classes of any given food ingredient and confirmation of the authenticity of a food or food ingredient are two other analytical problems requiring exploration. Primarily, what is involved is the comparison of a test material to a set of values previously established as typical for the material in question. The basic assumption underlying the application of spectroscopy to this problem lies with the generation of a "fingerprint" of foods. Given that an individual food ingredient has a characteristic chemical composition, it may be expected also to have a characteristic spectrum since this is the result of the absorption of radiation by many (or all) of the chemical constituents. Statistical techniques used for identification or discrimination try to group similar material together in tight clusters and apart from other dissimilar (although perhaps related) materials. [For a fuller description of these and other techniques, the reader is referred to Downey, 1994 and Naes *et al.*, 2002].

ADULTERATION OF OLIVE OIL

Obtained from *Olea Europa sativa*, olive oils are marketed according to the process used for their extraction. Virgin olive oils (extra virgin, virgin, ordinary virgin and lampante virgin) are the most sought after and therefore most expensive grades, produced using only cold-pressing techniques. Virgin oils are in particular demand within Europe for both their sensory and nutritional properties and command a premium price. The main adulteration issue involves addition of other, cheaper oils while, since oils command different prices based on their country of origin, a significant labelling issue involves

false claims concerning the geographic origin of an olive oil. In this project, the potential of near infrared spectroscopy for (a) discriminating between extra virgin olive oils from adjacent regions of the eastern Mediterranean *i.e.* Crete, the Peloponese and other Greek locations and (b) detecting and quantifying the presence of sunflower oil was studied.

Olive oil samples were collected at random from farm lots at oil mills in the three eastern Mediterranean regions by Elais S.A. (Peireus, Greece); the oils were frozen and stored at -18°C for several months. A single sample of sunflower oil was collected and stored in the same way. Prior to spectral collection, oils were thawed and stored in screw-capped glass vials at room temperature in the dark for a period of approximately 3 months. A total of 65 samples was available from three broad regions *i.e.* Crete (n=18), the Peloponese (n=28) and other parts of Greece (n=19). Those samples originating outside Crete and the Peloponese were collected from the north, west and central regions of the Greek mainland as well as islands off the north-west and eastern coasts.

Visible and near infrared transflectance spectra (400 - 2498nm) were recorded on a NIRSystems 6500 scanning monochromator (FOSS NIRSystems, Silver Springs, MD, USA) fitted with a sample transport accessory. Samples (0.5ml approx.) were placed on the inside of the quartz window of a 0.1mm pathlength camlock cell fitted with a gold-plated backing disc. Instrument control and initial spectral manipulation were performed using WinISI II software (v.1.04a; Infrasoft International, Port Matilda, MD, USA). Oil spectra were collected at an ambient temperature of between 18 and 24°C. Between samples, the sample cell components were washed with detergent in warm water, rinsed with tepid tap water and then with distilled water at room temperature. Components were dried using paper tissue. Samples were scanned in duplicate with the mean spectra being used in all subsequent calculations.

Spectral data were examined for unusual or outlying samples by principal

Figure 1. Near infrared spectrophotometer and camlock transflectance sample cell.

component analysis (The Unscrambler v 7.6; CAMO A/S, Oslo, Norway) and hierarchical cluster analysis (Pirouette v.3.02; Infometrix Inc., Woodinville, WA, USA). Classification techniques investigated in this work were partial least squares regression (PLSII - The Unscrambler), factorial discriminant analysis (FDA – SAISIR; D. Bertrand, INRA, Nantes, France) and *k*-nearest neighbour analysis (*k*-NN - Pirouette). Models were developed using three wavelength ranges - 400-2498nm, 400-750nm (visible) and 1100-2498nm (near infrared) with and without data pre-treatment.

The mean spectra of all 65 oil samples are shown in Figure 2. Absorption maxima are clearly evident at 450, 672, 1210, 1722, 1760, 2310, 2346 and 2386nm. Smaller absorption bands may be seen at 2124 and 2144nm. Bands around 1200nm arise from overtones of C-H stretching vibrations while those at 1722 and 1760nm arise from the first overtone of C-H stretching vibrations of methyl, methylene and ethylene groups. Oleic acid has been reported to absorb at 1725nm while saturated and *trans*-unsaturated triglycerides have absorption maxima at 1725 and 1760nm. Absorbances at 2310, 2346 and 2386nm are due to combination bands arising from C-H stretching vibrations and other vibrational modes. Smaller peaks at 2124 and 2144nm have been ascribed to the presence of *cis* double bonds. Following a hierarchical cluster analysis of the same data set using the 1100-2498nm wavelength range (Figure 3), it was observed that samples #5 and #8 clustered together at one

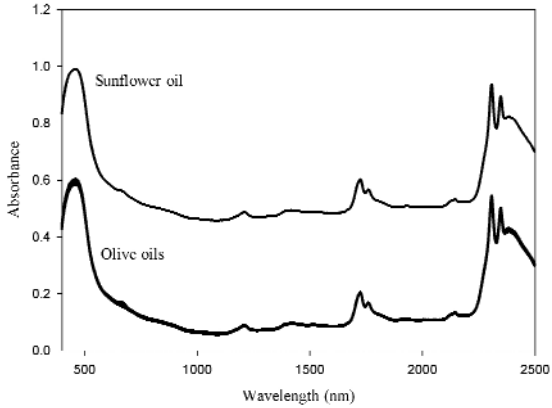


Figure 2. Average visible and near infrared spectra of extra virgin olive oils (n=65) and sunflower oil (n=1; offset for clarity).

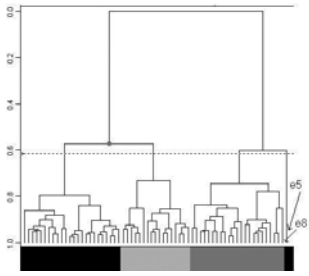


Figure 3. Hierarchical cluster analysis of near infrared spectra of olive oils; leftmost block mainly Peloponese, centre block, mainly Other Greece and rightmost block, mainly Crete.

extreme end of the calculated dendrogram with #5 having the largest distance of any sample to any cluster. This positioning suggested that the samples may be outliers but given the fact that (i) these samples did not show any anomalous behaviour when the complete (400-2498nm) wavelength range was studied, (ii) these diagnostic statistics are advisory and not prescriptive

and (iii) no other evidence existed to corroborate their potential outlier status, they were retained as valid samples for data analysis purposes.

Classification of the three olive oil types on the basis of their geographic origin was attempted using partial least squares regression (PLSII), factorial discriminant analysis (FDA) and *k*-nearest neighbours (*k*-NN) analysis. The best overall classification was obtained using FDA (Table 1) on raw spectral data, producing correct classification rates of 97 and 94% in calibration and prediction sample sets respectively.

Table 1. Classification of olive oil samples by factorial discriminant analysis

Wavelength range (nm)	PCs[1] in optimal model	Correct classification rate[2] (%)	
		Calibration samples	Prediction samples
1100-2498	4(#2,4,5,6)	84	73
400-750	3(#1,3,4)	78	73
400-2498	4(#1,3,7,12)	97	94

[1]principal components [2]percentages rounded to nearest integer

The most significant wavelengths in this best discriminant model are in the visible wavelength range (457, 541, 579 and 671nm arising from carotenoids, pheophytins and carotenoid pigments) and those in the near infrared at 2310 and 2350nm.

Regarding detection and quantification of sunflower oil in olive oils, 46 pure extra virgin olive oils were each separately adulterated with 1% ($^{w}/_{w}$) and 5% ($^{w}/_{w}$) sunflower oil, producing a sample set of 138 oils. In terms of adulterant detection, the strategy pursued was to develop a mathematical model which described pure olive oils and apply this to the spectra of pure and adulterated

samples. A technique called SIMCA (soft independent modelling of class analogy) was used for this purpose; as before, a number of different spectral ranges and pre-treatments were studied. The most accurate model was produced using a 1st derivative of spectral data from 400-2498nm; in this case, 100% of pure olive oils were correctly identified as were the adulterated samples. Of particular interest was the fact that no false positive predictions were made *i.e.* no adulterated samples were wrongly identified as authentic. Once detected, the next step was quantification of the adulterant sunflower oil. PLS regression was applied to this problem; using raw spectral data in the range 400-2498nm, it was possible to predict sunflower oil content with a standard error of prediction equal to 0.8%. This suggests a minimum detection limit of approximately 3% w/w.

ADULTERATION OF SOFT FRUIT PURÉES

Soft fruits such as strawberry (*Fragraria ananassa)* and raspberry *(Rubus idaeus)* are key raw materials for the manufacture of a wide range of foods including jams, preserves, yoghurts and pie-fillings. Due to their fragile and perishable nature, internationally-traded strawberries and raspberries are normally transported either in block frozen form or as purées. In purée form, the physical structure of the fruit is lost while the colour intensity is considerable reduced, making visual identification difficult. For this reason, economic adulteration of strawberry and raspberry purées is likely to occur; such adulteration may be by the addition of cheaper fruit (*e.g.* apple), water, sugar or vegetable matter. The work reported in this section describes the detection and quantification of a specific type of adulteration of strawberry and raspberry purées *i.e*. addition of apple.

Samples of strawberries (class 1 from Holland; n=32), raspberries (class 1 from Ireland and England; n=32) and cooking apples (England; n=22) were purchased in local supermarkets during June and July 2001. Samples were stored at +4ºC in the dark for 24 hours after purchase before conversion to purées. Apples were peeled and cored, with only the flesh being processed

further. Lots of individual fruit types were puréed in a blender (Braun Hand-held Liquidiser, 250 watt) for 30 seconds. Prior to mixing or further treatment, purées were stored in covered plastic beakers at +4ºC. Adulterated strawberry and raspberry samples were produced by adding apple purée to the pure fruit purées to produce apple concentrations of 10, 20, 30, 40, 50 and 75% by weight.

Reflectance spectra were recorded and analysed statistically as described in the previous section. Typical spectra are shown in Figure 4; the main features are absorptions around 1450 and 1940nm which arise from water. Principal component models were made for the unadulterated strawberry (n=32) and raspberry (n=32) purée samples using a number of wavelength ranges. Each of these models was evaluated on a spectral file containing the same unadulterated purées (n=64) and the adulterated strawberry and raspberry material (n=280 samples); a summary of the results obtained is shown in Table 2. In interpreting these results, the number of false positive and false negative identifications is calculated. A false positive results when, for

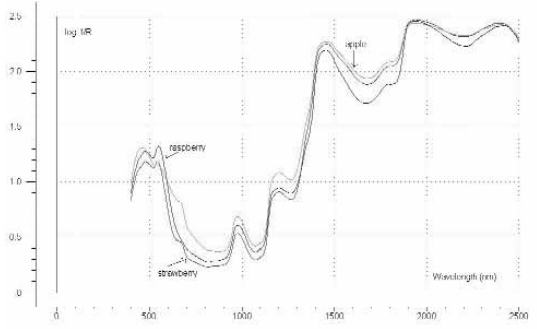


Figure 4. Representative reflectance spectra of apple, strawberry and raspberry purées collected using a transflectance cell with a gold-plated reflector (0.1mm sample thickness).

Table 2. Summary results of fruit purée classifications using spectra collected in a camlock cell fitted with a gold-plated reflector (0.1mm sample thickness; most accurate models in bold)

Fruit Type	Wavelength range (nm)	FP[1]	FN[2]	PCs[3]
Raspberry	750-1100	118	0	2
	400-1880	52	0	3
	1100-1880	111	0	2
	400-750	9	0	3
	400-1100	13	0	3
Strawberry	750-1100	103	1	1
	400-1880	43	0	2
	1100-1880	126	0	2
	400-750	46	0	2
	400-1100	36	0	2

[1]false positive classifications [2]false negative classifications
[3]number of principal components in model

example, a non-raspberry (in this case a strawberry or adulterated raspberry or strawberry) is wrongly identified as pure raspberry; a false negative result occurs when a true raspberry sample is wrongly identified as either adulterated or non-raspberry material. Ideally, the value for false positive and negative results should be zero. The best model for raspberry classification used spectral data in the visible wavelength region only and displayed good predictive accuracy. All of the 9 false positive results were adulterated at the

lowest (*i.e.* 10% $^{w}/_{w}$) level indicating that, in the present study, raspberry purées containing more than 10% apple can be identified as adulterated. Information in the visible range of the spectrum is important in performing this classification; this is believed to arise from coloured anthocyanin components, particularly cyanidin-3-sophoroside, which are present in raspberry but absent from apple. In the case of strawberry, the optimum model used spectral data in the 400-1100nm wavelength range; error rates were higher than for raspberry, with 36 false positives reported. Twenty-five of these were at the 10% apple adulteration level and 11 were at 20% adulteration. This indicates that, in practical terms, detection of apple adulteration in strawberry purée at less than 20% $^{w}/_{w}$ is not possible. Wavelengths of 482 and 546nm were the most important in the strawberry model, likely to be due to anthocyanins present in strawberry, mainly pelargonidin-3-glucoside. No strawberry sample (unadulterated or adulterated) was mis-classified as raspberry (unadulterated or adulterated) or *vice versa* revealing good separation between the fruit models.

Quantification

Regression models were developed to predict the percentage apple in adulterated strawberry and raspberry samples and a summary of the results obtained is shown in Table 3. For raspberry, the most accurate model used spectral data in the wavelength range 400-1100nm and predicted apple content with a root mean square error of prediction (RMSEP) equal to 3.4%. For strawberry, the optimum model used data between 400 and 1880nm and had a prediction error of 5.5%. Comparing the classification and quantification tables, it is apparent that the optimal wavelength ranges selected were not identical in both cases, although for raspberry they were close. This is likely to be a consequence of the different data compression strategies applied in each case *i.e.* principal components analysis *vs* partial least squares regression.

Table 3. Summary results for quantification of apple content in adulterated raspberry and strawberry purées.

Fruit	Wavelength range (nm)	Factors[1]	RMSEP[2]	R[3]	a[4]
Raspberry	750-1100	9	13.6	0.85	0.74
	400-1880	4	3.7	0.99	0.98
	1100-1880	16	7.7	0.95	0.94
	400-750	3	4.3	0.99	0.97
	400-1100	3	3.4	0.99	0.98
Strawberry	750-1100	8	15.4	0.80	0.66
	400-1880	5	5.5	0.98	0.96
	1100-1880	10	11.3	0.90	0.82
	400-750	3	7.5	0.96	0.92
	400-1100	3	6.7	0.97	0.94

[1]number of partial least squares factors in regression model; [2]root mean square error of prediction; [3]correlation coefficient ; [4]slope of regression line

ADULTERATION OF HONEY

Honey is a natural biological product of complex and variable composition. Within the European Union, it is defined by Council Directive 2001/110/EC as " ..the natural, sweet substance produced by *Apis mellifera* bees from the nectar of plants or from secretions of living parts of plants or excretions of plant-sucking insects on the living parts of plants, which the bees collect,

transform by combining with specific substances of their own, deposit, dehydrate, store and leave in honeycombs to ripen and mature." Central to this definition is the statement that honey is a completely natural product to which nothing may be added or removed. The concepts of purity and naturalness are important properties much-valued by consumers and certain groups are prepared to pay a price premium for honey produced by small-scale, artisanal producers. The main components of honey are carbohydrates (~80% w/w) and water (~17%); minor components such as oligosaccharides, gluconic acid, amino acids and minerals account for ~3% w/w. Fructose and glucose, typically present in the ratio 1.2:1, account for 85-95% of the total carbohydrate in honey. Sucrose accounts for about 1.5% with disaccharides typically present at a level close to 6.5%. Extension of honey with sugar solutions or syrups is therefore a technologically simple adulteration method which has been used since Roman times (when concentrated grape juice was used as an adulterant) and through the Middle Ages. Such extension could involve the use of any commercially-available sweet substance such as high fructose corn syrup, invert cane sugar or even simple mixtures of fructose and glucose. Work undertaken in this project involved a study of the utility of NIR spectroscopy for the detection of adulteration of honeys with fructose and glucose (FG) mixtures, beet invert syrup (IBS) and high fructose corn syrup (HFCS). A parallel study involved the application of MIR spectroscopy to detect honey adulteration by partially-inverted cane syrup (PICS), beet invert syrup (IBS), beet sucrose (BS), dextrose syrup (DS), HFCS and solutions of fructose and glucose (FG).

Honey samples were obtained directly from beekeepers throughout the island of Ireland during the years 2000 and 2001; they were stored unrefrigerated from time of production until scanning and were not filtered after receipt in the laboratory. Immediately prior to spectral collection, honeys were incubated at 40°C overnight to dissolve any crystalline material, manually stirred to achieve homogeneity and adjusted to a standard solids content (70° Brix) with distilled water. Adulterant solutions were also adjusted to 70° Brix.

Near infrared spectroscopy

In the first study, commercially-sourced, fully-inverted beet syrup (IB; 50:50 fructose:glucose; Siúcra Eireann Teo, Carlow, Ireland) and high fructose corn syrup (HFCS; 45% fructose and 55% glucose; Unilever Ireland Ltd.) were used to adulterate honeys (n=83). Ten pure honeys were adulterated by HFCS at 10, 30, 50 and 70 % w/w while 8 were adulterated by IB at levels of 7, 10, 14, 21, 30, 50 and 70 % w/w. This produced 96 adulterated samples in total. Raw and 2nd derivative spectra of honey, adulterated honey and the individual adulterants are shown in Figure 5. The similarity between the 2nd derivative spectra is obvious; the main features are large minima at 1442, 1930, 2278, 2324 and 2490nm with smaller troughs evident at 1696 and 2112nm and in the range between 2340 and 2450nm (Figures 5b and 5d). Following principal component analysis of the entire raw spectral data set, the score plot for components 1 and 2 revealed the clear presence of several groups (Figure 6). In the case of beet invert sugar (Figure 6a), there is a clear separation into three groups - one of these represents authentic honey (H) while the others contain honey plus beet invert syrup at various adulteration levels. For high fructose corn syrup adulteration (Figure 6b), samples are also clearly divided into two groups - authentic (H) and adulterated. Discrimination between authentic and adulterated honeys was investigated using SIMCA and results are summarised in Table 4. The best model, involving 5 principal components, was obtained using raw spectral data. In this case, all the adulterated honeys and 23 of the 27 unadulterated honeys were correctly identified.

The robustness of this honey model was evaluated by adding 52 different authentic Irish honey samples to the existing validation set to produce a total of 79 authentic samples. Thirteen of these samples were not classified as authentic honeys, producing an overall correct classification rate of 78% (62 out of 79). Two-thirds of these extra authentic samples were then added to the original calibration file to develop a new authentic model; when this was tested on the validation set produced by adding the remaining 1/3 of the extra samples to the previous validation file, the number of authentic samples

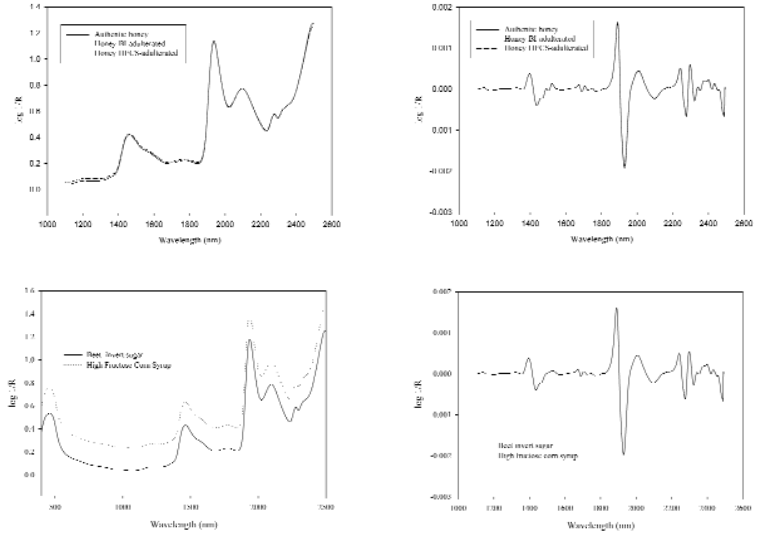


Figure 5. Raw and 2nd derivative transflectance spectra of honey + adulterated honeys (a + b) and adulterant solutions (c + d).

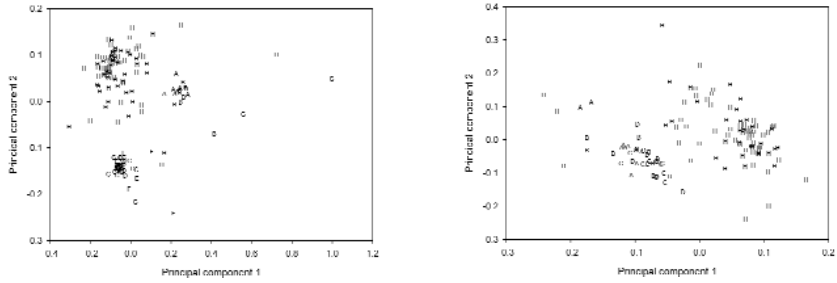


Figure 6. Scores plots for authentic and adulterated honeys: (a) IB adulterant [authentic (H) and beet invert syrup-adulterated honeys at 7 (A), 10 (B), 14 (C), 21 (D), 30 (E), 50 (F) and 70% w/w (G)]; (b) HFCS adulterant [authentic (H) and high fructose corn syrup-adulterated samples at 10 (A), 30 (B), 50 (C) and 70% w/w (D)].

correctly identified emerged as 90.9%. No adulterated honey was classified as authentic. This result indicates that the greater variability incorporated by the extra honey sample may be successfully handled by SIMCA to produce a model of higher accuracy. The absence of any false positive results is also very important for a screening technology such as NIR spectroscopy.

Table 4. Classification of authentic and adulterated (IB plus HFCS) honey by SIMCA

Data pre-treatment	No. of correctly-classified samples		
	Authentic (n=27)	Adulterated IB[1] (n=56)	Adulterated HFCS[2] (n=40)
None	23	56	40
MSC[3]	22	56	40
2Der[4]	23	40	34

[1]beet invert syrup; [2]high fructose corn syrup; [3]multiplicative scatter correction; [4]2nd derivative (Savitzky-Golay, 10 datapoint gap)

A second study involved honey adulterated with solutions of glucose and fructose. This involved 75 Irish honeys adulterated by aqueous solutions of fructose and glucose in the following ratios: 0.7:1, 1.2:1 and 2.3:1 w/w. In the initial part of the study, 25 of the pure honeys were adulterated with each of the three adulterant solutions at three levels *i.e.* 7, 14 and 21% w/w, thus producing 225 adulterated honeys. In an extension to the study, a further 50 unadulterated honeys (standardised to 70° Brix) were added to the sample set and the predictive models developed in step 1 were then applied to the enlarged dataset. A number of mathematical techniques were studied to develop discriminant models and the best performance was by discriminant PLS. The optimum model involved the use of spectral data between 1100 and 2498nm after a 1st derivative data pre-treatment; 21 out of 25 authentic honeys and 222 out of 225 adulterated samples were correctly identified. In

the extended study, 72 out of the 75 (96%) unadulterated honeys and 223 of the 225 (99%) adulterated samples were correctly identified.

Mid-infrared spectroscopy

In a study involving honey adulteration by glucose and fructose solutions, the experimental design was as described above. In this case, 99 authentic honeys and 221 adulterated samples were collected and scanned. MIR spectra were collected at room temperature on a BIO-RAD Excalibur series FTS 3000 spectrometer (Figure 7: Analytica Ltd., Dublin, Ireland); instrument control and spectral collection was performed using WIN-IR Pro (v 3.0) software supplied by the equipment manufacturer. Spectra were recorded on an in-compartment benchmark ATR trough top plate using a 45° ZnSe crystal with 11 internal reflections. Plots of typical spectra of honey and varying concentrations of fructose are shown in Figure 8. This plot reveals the spectral changes which occur from increasing concentrations of this sugar in water and suggest that infrared spectroscopy may be a useful tool to use in detection of it and other sugars. Applying discriminant PLS to this dataset, the best results were obtained using normalised spectra data; this produced correct-classification rates of 87.9 and 96.4% for authentic and adulterated samples respectively.

A subsequent MIR study involved a large collection of Irish honeys (n=580) and a subset adulterated with fully-inverted beet syrup (n=280), high fructose corn syrup (n=160), partially inverted cane syrup (n=120), dextrose syrup (n=160) and beet sucrose (n=120). Using a decision-tree approach (*i.e.* rather than trying to classify samples as one of 6 groups, to proceed through a series of binary classification decisions until a definitive identification is arrived at) combined with soft independent modelling of class analogy (SIMCA) and partial least squares (PLS) classification, authentic honey and honey adulterated by beet sucrose, dextrose syrups and partially inverted corn syrup could be identified with correct classification rates of 96.2, 97.5, 95.8 and 91.7% respectively. This combination of spectroscopic technique and

Figure 7. Mid-infrared spectrophotometer with attenuated total reflectance sample cell.

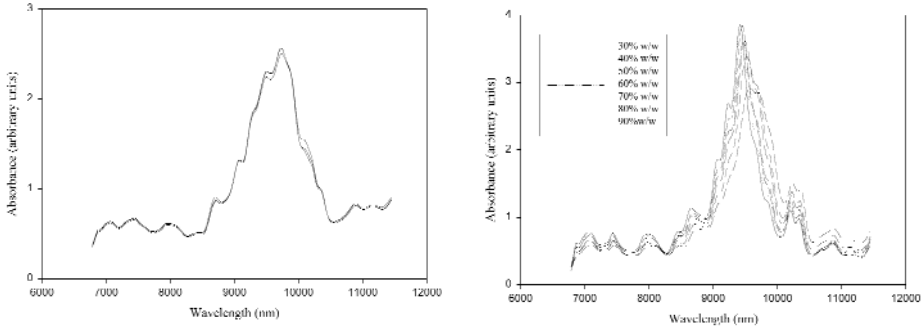


Figure 8. Mid-infrared attenuated total reflectance spectra of authentic honeys (a) and (b) varying concentrations (30-90% w/w) of fructose in water.

chemometric methods was not able to unambiguously detect adulteration by high fructose corn syrup or fully-inverted beet syrup.

Baseline study of Irish honey composition

During the course of the honey authentication work described above, it became apparent that, while much has been published on the characterisation of honey produced in a number of countries, little corresponding, independent information on the characterisation of honey produced in Ireland was available. Therefore, a study was performed to provide baseline compositional data for a small number (n=50) of Irish artisanal honeys. Analyses were performed to quantify moisture and ash contents, conductivity, acidity, hydroxymethylfurfural (HMF), pH and mineral content. HMF content is an indicator of a honey's thermal history since it increases after excessive heating; it may also be used as an indirect indicator of honey freshness. Palynological (*i.e.* pollen) analyses were also performed on a subset (n=25) of these samples to obtain pollen profiles which would be characteristic of Irish honeys.

Honeys were collected over two harvest seasons (2000 – 2002 and 2002 – 2003) each providing 25 samples. The thermal history of the two sample sets differed somewhat. Those sampled earlier, year 1, had undergone two periods of warm-holding (40°C overnight) while the second set, year 2, had not undergone any warm-holding prior to analysis. Palynological analysis (Lutier and Vaissière, 1993) was only performed on samples collected during the 2000 – 2001 season. Physico-chemical parameters were determined using standard methods (Downey *et al.*, 2005). Results of the qualitative pollen analysis for the 25 honey samples are summarised in Table 5 in which results are reported as percentages of the total pollen content in each sample. Overall, 43 pollen types (present at levels ≥1%) were identified; an accurate count could not be made for sample 21 because of the very small number of pollen grains detected. The number of pollen types (present at levels ≥1%) constituting the total pollen content of each sample ranged between five (sample 13) and eighteen (sample 5).
Trifolium repens was the dominant pollen type in nineteen of the twenty-five honeys. It was present in 23 samples ranging from 5 to 78% of total pollen. *Trifolium repens* (white clover) is a very common plant on the island of Ireland

Table 5: Pollen analysis of honey samples (n = 25) from 2000 - 2001 season. Results presented as percentages of total pollen. (p= values below 1%).

	Sample No.																								
Pollen type	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
Acer spp.		1	2	1	2					2	1	p		p	5	1	p	4	4	2		6	4	2	4
Heracleum spondylium					1													p						p	
Apiaceae			8																						
Ilex aquifolium									p						1	p			p	p					
Catalpa sp.							1	p																	
Myosotis						85																			
Buddleya globosa					1				p							2									
Cistus sp																						1			
Carduus spp.							1	p	1			p			1		p							p	
Brassica spp.	3		p	2	8	1	p	1	2	2		1			2	9	1	3	1	2					
Rorippa sp.					3	1		p	1			1			p		p	2					p		
Empetrum nigrum														1									p		
Cytisus scoparius	6	3	2	1		1	1	1			1			1	1	4	4	6	2	6		1	2	2	1
Lotus spp.	1	3	4	1	2		1	16	p	1		38		6	p				7	1		1	p	14	2
Medicago lupulina		11	14	16	4		5		9	9	25	16	7	22	18	23	7	6	5	6		3	11	3	1
Trifolium pratense	18																					p			
Trifolium repens type		28	40	33	29	5	71	38	60	51	40	30	78	20	32	26	69	48	60	42		42	45	59	67
Ulex spp.			p		1																				
Quercus spp.	1	p	2						1	1					2	2		2		1		1	p		p
Castanea sativa														8	p			p		p		p			p
Hypericum spp.			1	p	1								p												
Aesculus hippocastanum	1			p	1				p		p				p										

Table 5 (contd.): Pollen analysis of honey samples (n = 25) from 2000 - 2001 season. Results presented as percentages of total pollen. (p= values below 1%).

	Sample No.																								
Pollen type	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
Allium sp.	p	1	1				p			p					p	1							1		
Moraceae											1									p		p			p
Poaceae		1	1		1		1		1									1	1					p	p
Ligustrum vulgare					2	p		1				1				1			p						
Montia fontana								4	p																
Ranunculus spp.									1		p					p									
Anemone nemorosa					2	2			p	1		p	p			8	2	p		p			1		p
Caltha palustris				4	8	1	2	1	1	1			1	1	p	1				p					
Filipendula ulmaria	1	2	1	6	6		3	2	7	8	6	2	3	13	3	5	5	9	3	4		17	8	2	3
Malus spp.	p		4	1			1		1	p	p			1	3			5	2	p					
Prunus spp.		6	p	p	p				p	p		p	p		4	3		4	3	11		4	1	2	13
Rosa type								2																	
Geum sp.	1																								
Rosa spp.								p	p						1	2	2	1					p		
Rubus fruticosus	64	38	17	24	25	3	10	30	11		22	6	9	26	22	11	6	8	9	12		20	24	14	6
Potentilla sp.										p	p						1					p			
Fragaria sp.				8	1																				
Fragraria vesca												2			p	p								p	
Other Rosaceae		2	p							20															
Salix spp.	1	2	1	2			p	2	p	1				p	p	p	1	p		10		p	2		1
Verbena officinalis							2																		
Total % Pollen	97	98	98	99	98	99	99	98	96	97	96	97	98	99	95	98	99	99	97	97	0	97	98	98	98

and its presence in Irish honey in large amounts is to be expected. *Trifolium pratense* (red clover) was present at a significant level (18% of total pollen) in sample no. 1 (Table 5). *Rubus fructicosus* (blackberry) was the second most abundant pollen type identified, being the dominant pollen in three of the twenty-five honeys tested. It was present in 23 samples with values ranging from 3 to 64% of all pollen types present. Such large differences in percentage pollen content can be attributed to a number of factors: the amount of pollen present in the nectar can be very variable, pollen can be filtered out in the bee's honey sac or the bee may take pollen without taking nectar.

Pollen from a number of other species was present in a large number of the honey samples, albeit at generally low levels. These included *Brassica* spp. (cabbage family), *Lotus* spp. (*e.g. Lotus corniculatus* or Bird's foot trefoil), *Malis* spp. (*e.g.* crab apple), *Prunus* spp. (*e.g.* blackthorn, wild cherry or dwarf cherry) and *Acer* species (*e.g.* sycamore or maples). A number of specific plant varieties were similarly represented *i.e. Cytisus scoparius* (scotch broom), *Filipendula ulmaria* (meadowsweet) and *Caltha palustris* (marsh marigold). *Medicago lupulina* (black medick) was also present in 21 of the 25 honeys at levels ranging from 1 to 25% of total pollen. All of these are relatively common plants in Ireland and their presence in the artisanal honeys is to be expected.

Some unusual pollen types were identified in this sample set. Sample 6 contained >85% of total pollen from a single species (*Myosotis*; forget-me-not); it was the only sample to contain pollen from this plant. *Catalpa* is an ornamental woody plant not common in Ireland but its pollen was present in two samples (samples no. 7 and 8) although at very low levels, *i.e.* 1 and >1% respectively. Sample no. 12 was found to contain mostly *Lotus* spp. (Bird's foot trefoil). On the basis of this small sample set, therefore, it would appear *Apis mellifera* bees in Ireland feed mainly on a diet of nectar from white clover and blackberry.

Physico-chemical parameters

Honeys from year 1 had an average moisture content in the range 15.6 to 18.4% w/w, indicating optimum harvesting and a good degree of maturity (Table 6). Year 2 moistures ranged from 16.5 to 20.6%, with one sample slightly above the upper limit (20%) laid down in EU Council Directive 2001/110/EC (EU, 2001). Overall, year 2 samples had slightly higher moistures than year 1, but both sets when averaged were within the prescribed limits.

Ash content is a quality criterion of particular relevance for honey of stated botanical origin; blossom honeys have a lower ($\leq$ 0.6%) ash content than honeydew honeys ($\leq$ 1.2%). All honeys analysed in this work had ash contents below 0.6%, indicating that they were more likely to be of floral than honeydew origin. Electrical conductivity varies with botanical origin and conductivity values recorded ranged from 0.17 to 0.40 mS/cm^{-1} (year 1) and 0.11 to 0.48 mS/cm^{-1} (year 2); the mean conductivity value of 0.3 mS/cm^{-1} obtained for the 50 samples in this study is similar to published values for Spanish honeys of 0.25 and 0.21 mS/cm^{-1} respectively (Serra Bonheví and Granados Tarrés, 1993).

Acidity in honey is calculated as free, lactonic and total acidity. EU regulations (EU, 2001) specify a free acidity of not more than 50 milli-equivalents acid per 1000g (meq/kg). The average values for free acidity in samples from year 1 were between 23.8 and 42.1 meq/kg but in year 2, one value (50.9 meq/kg) very slightly exceeded EU limits. The mean free acidity values for the remaining year 2 honeys ranged from 20.7 to 48.5 meq/kg. Lactonic acidity ranges were from 0.2 to 14.9 meq/kg in year 1 and 0.3 to 6.3 meq/kg in year 2 samples. Total acidity ranged from 26.8 to 55.9 meq/kg and 21.2 to 52.4 meq/kg in year 1 and year 2 respectively.

In year 1 samples, a large variation in HMF (mg/kg) levels was found as evidenced by a standard deviation of 45.1. Actual levels recorded ranged from 8.1 to 230.8 mg/kg, with eleven out of the twenty five samples above the

Table 6. Summary of honey physico-chemical parameters.

Samples	Moisture % w/w	Ash % w/w	Conductivity mS/cm^{-1}	Free Acidity meq/kg	Lactonic Acidity meq/kg	Total Acidity meq/kg	HMF mg/kg	pH
Year 1 (2000-2001) (n=25)								
Range	15.6 - 18.8	0.07 - 0.36	0.17 - 0.40	23.8 - 42.1	0.2 - 14.9	26.8 - 55.9	8.1 - 230.8	3.85 - 4.28
Mean	17.2	0.2	0.3	32.6	4.5	37	19.7	4.1
SD	0.7	0.1	0.1	5.3	2.8	6.3	3.8	0.1
Year 2 (2002-2003) (n=25)								
Range	16.3 - 20.6	0.03 - 0.46	0.11 - 0.48	17.7 - 48.5	0.3 - 6.3	21.2 - 50.7	0.9 - 37.3	3.75 - 4.55
Mean	18.0	0.2	0.3	32.6	2.1	34.7	7.0	4.1
SD	1.1	0.1	0.1	10.0	1.7	9.7	8.6	0.2
All Data (Year 1 and 2) (n = 50)								
Range	15.6 - 20.6	0.03 - 0.46	0.11 - 0.48	17.7 - 48.5	0.2 - 14.9	21.2 - 55.9	0.9 - 230.8	3.75 - 4.55
Mean	17.6	0.2	0.3	32.7	3.4	36.1	26.1	4.1
SD	1.0	0.1	0.1	7.9	3.4	8.7	38	0.2

limit contained in the EU regulations (EU, 2001) *i.e.* not more than 40 mg/kg. The heat treatments and the long storage time (two heat treatments, two years storage) experienced by these samples prior to analysis would most likely explain these results. Year 2 honeys received no heat treatment and experienced only a very short storage time prior to analysis *i.e.* a maximum of 6 months. No sample in this group exceeded EU regulations (EU, 2001). Honey pH values are of great importance during extraction and storage as they influence texture, stability and shelf-life. Values recorded for this parameter in the current study ranged between 3.85 and 4.28 (year 1) and 3.75 to 4.61 (year 2); standard deviations of 0.1 and 0.2 respectively were recorded.

Mineral content is an important indicator of possible environmental pollution and a potential indicator of honey geographical origin. In this study, eight elements were quantified *i.e.* iron (Fe), copper (Cu), zinc (Zn), calcium (Ca), magnesium (Mg), manganese (Mn), sodium (Na) and potassium (K). Mean, standard deviation and range values of the data obtained are shown in Table 7. Calcium, sodium and magnesium levels exhibited average values of 10.8, 9.3 and 3.0 mg/100g honey respectively. Average values for iron, zinc, manganese and copper were 1.0, 0.7, 0.4 and 0.2 mg/100g respectively. One interpretation of this observation is that honeys of Irish origin may be less exposed to industrial pollution than those from other geographical locations. Apart from HMF values in year 1, which were due to heating and storage treatments not likely to be encountered in artisanal honey as normally purchased, these Irish honeys tested complied with the relevant European standards. Evidence from all parameters tested is consistent with the fact that they were generally of floral origin.

ADULTERATION OF FRUIT JUICES

Apple juice was selected as a model food in this section of the project. Its main authenticity issues are those that arise from the substitution of less-costly or lower quality ingredients for those declared on the label of the fruit

Table 7. Summary of honey mineral content (mg/100g of honey).

Samples	Fe	Cu	Zn	Ca	Mg	Mn	Na	K
Year 1 (2000-2001) (n=25)								
Range	0.17 - 1.32	0.14 - 0.23	0.16 - 0.85	7.93 - 15.15	1.89 - 5.33	0.17 - 1.02	6.01 - 15.82	41.02 - 69.34
Mean	0.7	0.2	0.3	11.3	3.2	0.4	10.2	55.5
SD	0.3	0.0	0.2	2.1	0.8	0.2	2.7	7.7
Year 2 (2002-2003) (n=25)								
Range	0.25 - 3.63	0.10 - 0.23	0.17 - 2.25	7.49 - 17.54	2.01 - 3.93	0.09 - 1.00	4.13 - 19.57	44.71 - 71.40
Mean	1.0	0.2	0.7	10.8	3.0	0.4	9.3	57.7
SD	0.6	0.0	0.5	2.4	0.5	0.2	4	7.1
All Data (Year 1 and 2) (n = 50)								
Range	0.17 - 3.63	0.10 - 0.23	0.16 - 2.25	7.49 - 17.54	1.89 - 5.33	0.09 - 1.02	4.13 - 19.57	41.02 - 71.40
Mean	0.8	0.2	0.5	11.1	3.1	0.4	9.8	56.6
SD	0.5	0.0	0.5	2.3	0.7	0.2	3.5	7.5

SD: standard deviation

beverage. Water and carbohydrates are the main components of apple juice; fructose, glucose and sucrose are the main carbohydrates with average concentrations of 5.6, 2.5 and 1.7 % w/w respectively. As a result, the majority of apple juice adulterations involve the substitution of juice solids with different types and combinations of sugar solutions or syrups. The objective of this study was to investigate the usefulness of NIR analysis for the detection and quantification of apple juice adulteration by added sugars.

Near infrared spectroscopy

Two different adulterants were studied *i.e.* (1) high fructose corn syrup (a commercial sweetener that approximates the carbohydrate composition of apple juice) and (2) a mixture of sugars (fructose, glucose and cane sucrose) reflecting the average composition of apple juices. Predictive models were developed for each adulterant separately and for both combined. Apple samples from 19 different varieties were collected from orchards throughout the main cultivation areas in Ireland. Two batches of samples were used in this study. A first group of 68 samples was collected in December 2002 and January 2003 from the following varieties: Bramley (14), Elstar (7), Fiesta (1), Gala (1), Golden Delicious (5), Ida Red (7), Ingrid Marie (1), Jonagold (16), Jonagored (6), Jupiter (1), Karminjn de Sonne Ville (5), Pinova (1) and Red Prince (3). Samples were stored refrigerated (4°C) until February when the apple juices were extracted from the fruit using a L'equip model 110.5 centrifugal juicer (L'equip, Lemoyne, PA, USA) without filtering. Apple juices were kept frozen (-20°C) in screw-capped glass vials and spectra were recorded in July 2003. Samples (100 ml) were held overnight in a cold (4°C) dark room for defrosting and then in the lab at room temperature (20-25°C) to equilibrate before spectral collection. A second group of 82 samples was collected in October-November 2003 from: Braeburn (1), Bramley (20), Cox Pippin (4), Elstar (13), Fiesta (5), Gala (3), Golden Delicious (10), Hongate Wonder (1), Ida Red (2), Jonagold (10), Jonagored (8), Jupiter (1), Karminjn de Sonne Ville (2), Lord Lambourne (1) and Redwood Elstar (1). Apple juices were immediately obtained and these juices were kept frozen until collection of spectra in December 2003.

Adulterant solutions were produced at 12°Brix, the average Brix content in apple juice samples. High fructose corn syrup (HFCS) contained 45% fructose and 55% glucose while the prepared solution of sugars (SUGARS) comprised 60% fructose, 25% glucose and 15% cane sucrose (the average content of these sugars in apple juice). Adulterated juice samples were prepared with both types of adulterants. In the first batch of 68 apple juices, odd-numbered samples were adulterated at 20 % with HFCS and 40% with SUGARS, and even-numbered samples at 40% with HFCS and 20% with SUGARS. In the second batch of 82 apple juices, odd-numbered samples were adulterated at 10% with HFCS and 30% with SUGARS, and even-numbered samples at 30% with HFCS and 10% with SUGARS. A total of 450 pure and adulterated juices were thus prepared and scanned using a camlock cell fitted with a gold-plated backing plate (0.2mm pathlength).

Average raw and second derivative near infrared transflectance spectra of the authentic apple juice samples are shown in Figure 9. The raw spectrum reveals clear peaks only at 464, 1452, 1932 and 2500nm. After transformation, the second derivative spectrum reveals minima (*i.e.* absorption maxima in the original spectra) at 484, 1414 and 1906nm. Other absorption bands in the derivatised spectrum are observable at 430, 664, 1154, 1462, 1504 and especially above 2250nm at 2276, 2320, 2356, 2388, 2418, 2446 and 2484nm. Peaks around 1400 and 1900nm arise from water absorption while features in the wavelength region around 2200-2400nm are characteristic of sugar bands. Principal component analysis (PCA) was performed for a preliminary examination of the entire spectral dataset. The score plot for components 1 and 2 shows that, although no special grouping can be observed for pure and adulterated apple juices, two groups can be clearly seen (Figure 10). This grouping is on the basis of the batch of apples used to produce the authentic juices. Different procedures used to obtain the apple juices may have significantly affected the composition of both groups of samples.

In the first batch, the samples were stored refrigerated (4°C) for 1-2 months before the apple juices were extracted whereas in the second batch the apple

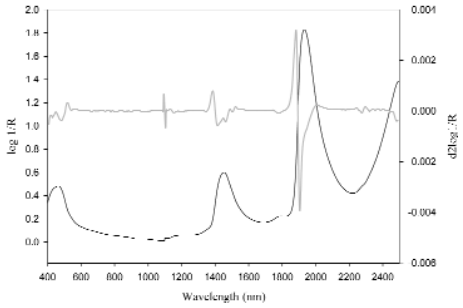


Figure 9. Average raw (black) and second derivative (grey) transflectance spectra of authentic apple juices (the feature at 1100nm in the second derivative spectrum arises from a detector change in the instrument).

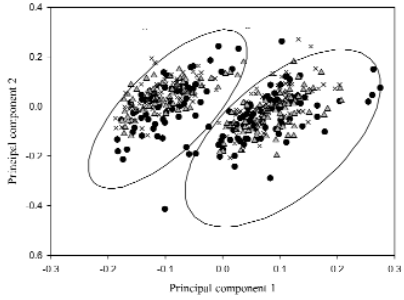


Figure 10. Principal component scores plot of 150 authentic apple juice samples (•) and the same samples adulterated with HFCS (o) and SUGARS (+) solutions. Ellipses representing batch 1 (left) and 2 (right) are indicative of cluster boundaries only.

juices were obtained immediately after the collection of fruits. Apple juices from each batch were also kept frozen (-20°C) for different periods of time before spectral collection although resultant compositional effects are unlikely to be significant. Significant differences in the sugar composition of juices made from fresh and stored apples have been reported: the sucrose content decreased while both the fructose and glucose concentrations increased on storage, indicating that sucrose was inverted. Some sugar can also be lost through respiration in intact apples, thereby reducing the total carbohydrate content.

Discriminant PLS regression was studied as a method for discriminating between authentic and adulterant apple juices. This tool was applied separately for each adulterant type and for combined adulterants using the whole wavelength range (400-2498nm) after multiplicative scatter correction of data. A summary of the results obtained is shown in Table 8. The most useful model is that for detecting either type of adulteration *i.e.* combined; this reveals a correct classification rate of 96% for adulterated juices but a lower value (86%) for authentic juices. Higher correct classification rates for

Table 8. Apple juice classification by discriminant PLS models developed separately for each adulterant type and for both adulterants in combination.

		Unadulterated apple juice		Adulterated apple juice	
Adulterant	# of PLS loadings	# of samples	% correctly classified	# of samples	% correctly classified
HFSC	11	150	98	150	91
SUGARS	11	150	91	150	93
Combination	13	150	86	300	96

authentic juice were reported for individual adulterants. Nonetheless, this represents a commercially-useful outcome given that the greatest danger for a food processor or retailer is the inadvertent use of an adulterated juice. Samples identified as adulterated by this model may be referred for further analysis to confirm their authenticity.

Mid-infrared spectroscopy

An investigation into the utility of MIR spectroscopy for apple juice authentication involved 224 authentic apple juices and 480 adulterated samples. Adulterants studied were partially inverted cane syrup (PICS), beet

sucrose (BS), high fructose corn syrup (HFCS) and a synthetic solution of fructose, glucose and sucrose (FGS). Adulteration was carried out on individual apple juice samples at levels of 10, 20, 30 and 40% w/w. Apples were gathered from orchards throughout the island of Ireland during the 2002 (n=69) and 2003 (n=155) harvest seasons (total=224). Nineteen apple varieties were included *i.e.* Braeburn (n=2), Bramley (n=50), Cox (n=8), Elstar (n=30), Fiesta (n=11), Gala (n=5), Golden Delicious (n=20), Granny Smith (n=1), Howgate Wonder (n=1), Idared (n=11), Ingrid Marie (n=1), Jonagold (n=32), Jonagored (n=22), Jupiter (n=3), Karmijn de Sonne Ville (n=7), Lord Lambourne (n=1), Pinova (n=1), Red Prince (n=3), Redwood Elstar (n=1) and unknown samples (n=14). In the case of HFCS and sugar mixture adulterants (FGS), levels of 10, 20, 30, and 40% w/w adulteration were used; a total of 150 samples of authentic apple juice were adulterated, each at one level only. In the case of the PICS and BS adulterants, 30 samples of authentic apple juice were adulterated at levels of 10, 20 and 30% w/w. Therefore, the total number of samples studied was 704 *i.e.* 224 plus 480 adulterated.

Fourier transform infrared spectra of two randomly-selected apple juice samples and the individual adulterant solutions are shown in Figure 11. The

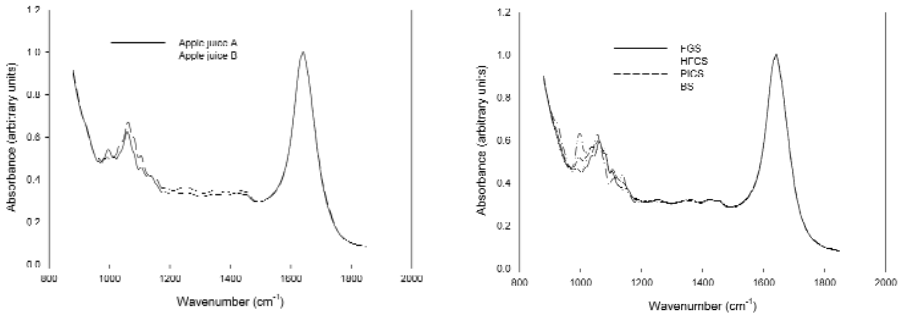


Figure 11. Mid-IR spectra of (a) two randomly-selected authentic apple juice samples and (b) adulterant solutions (FGS = fructose:glucose:sucrose; HFCS = high fructose corn syrup; PICS = partially inverted corn syrup; BS = beet sucrose).

spectrum of apple juice is dominated by water and sugar absorptions; bands appearing between 1150-1470cm^{-1} are attributed to bending modes of C-C-H, C-O-H and O-C-H groups while more intense peaks in the region between 900 and 1150cm^{-1} arise mainly from C-O and C-C stretching modes, with a peak around 1020-1060cm^{-1} due to O-H vibrations. At lower energies, bands due to C-H and O-H bending vibrations are also useful for discrimination and quantification purposes. At 1725cm^{-1}, absorption from organic acids (C=O stretch) appears as a shoulder on the large water absorption band at 1641cm^{-1}.

Discriminant models were developed using a variety of mathematical approaches for each adulterant separately. A summary of the results obtained is shown in Table 9. Discriminant partial least squares emerged as the best discriminant tool for each adulterant; for all adulterants except FGS, the overall percentage correct classification figures exceeded 92%, a high accuracy level likely to be of commercial use if confirmed on a larger sample set. Closer examination of the prediction results reveals that for PICS, HFCS and BS, the incorrect false positive identifications of adulterated juices arose from the lowest adulteration level used. It is not surprising that the most difficult adulterant to detect is FGS; this solution was prepared to mimic closely the natural composition of apple juices and, unlike the other adulterants, lacks any other oligosaccharide material which might have a detectable spectroscopic signal.

Quantification of each adulterant was attempted using PLS regression and the results are summarised in Table 10. The model developed for PICS adulterant produced a root-mean-square error of cross-validation (RMSECV) value of 4.9% with a correlation coefficient equal to 0.89. The 95% confidence level for the prediction error associated with this model is ±9.7% adulteration (*i.e.* ±1.98 * RMSECV). Similar prediction accuracies were obtained for BS and HFCS adulterants. This procedure was not, however, able to effectively quantify adulteration by FGS solution as revealed by the values of the relevant RMSECV, slope and intercept (9.5, 0.58 and 3.92 respectively).

Table 9. Summary of classification results for apple juice adulterated by partial invert cane syrup (PICS), beet sucrose (BS), high fructose corn syrup (HFCS) and fructose:glucose:sucrose mixture (FGS) achieved using mid-infrared spectroscopy

Adulterant	Discriminant Method	% Correct Classification		
		Authentic Juice	Adulterated Juice	Overall
PICS	PLS	96.0	97.8	96.5
BS	PLS	94.6	92.2	93.9
HFCS	PLS	94.2	89.3	92.2
FGS	PLS	89.3	72.0	82.4

Table 10. Statistical descriptors for linear regressions of predicted *vs* actual sugar adulterants in apple juice based on mid-infrared spectra.

Adulterant	Correlation coefficient	RMSECV*	Slope of regression line	Intercept of regression line
PICS	0.89	4.9	0.80	1.28
BS	0.90	4.6	0.80	1.13
HFCS	0.94	4.6	0.90	0.98
FGS	0.76	9.5	0.58	3.92

*root-mean-square error of cross-validation

CONCLUSIONS

Near and mid-infrared spectroscopies are powerful analytical tools at the disposal of the food industry. Their important attributes include speed, relative ease-of-use, non-destructive nature and the lack of any requirement for

chemical reagents. In the current project, they have been applied to detect and quantify a range of adulterants in a number of food types *i.e.* olive oil, honey, soft fruit purées and apple juice. In each case, one or other of the techniques, in conjunction with multivariate mathematical tools, has demonstrated a capability to discriminate between authentic and adulterated foods although the sensitivity of this detection varied from food to food and adulterant to adulterant. The results of these feasibility studies strongly support the extension of this work to involve larger sets of commercially-sourced samples; these are best undertaken on a case-by-case basis with major industry involvement.

RECOMMENDATIONS TO INDUSTRY

Near and mid-infrared spectroscopy together with multivariate mathematics have the potential to address food adulteration issues in a number of foods and food ingredients.

Each food adulteration issue requires a specific investigation to develop accurate and robust discriminant models.

The food industry needs to acquire the expertise and equipment to enable it to police the authenticity of its high-value raw materials. This is especially relevant from the point-of-view of brand protection.

Staff and equipment at Ashtown Food Research Centre are available as a national resource to assist with advice on instrumentation, data analysis and acquisition of expertise.

ACKNOWLEDGEMENTS

The authors acknowledge the assistance of individual beekeepers, apple and soft fruit growers in the donation of samples for this work. Thanks are also due to Laure Contal, Vanessa Fouratier, Lorenzo León and Cristina Petisco Rodriguez for some data collection and analysis; Karen Hussey and Thomas

Walshe are thanked for skilled technical assistance. This work was carried out as part of a research project entitled “Screening techniques for authenticity and safety confirmation in raw materials and food products” funded by the FIRM programme of the Department of Agriculture and Food (contract no. 00/R&D/TN/89). University College Dublin was the collaborating partner in this project.

REFERENCES

Downey, G. (1994). Qualitative analysis in the near infrared region. *Analyst*, 119, 2369-2375.

Downey, G., Hussey, K., Kelly, J.D., Walshe, T.F. and Martin, P.G. (2005). Preliminary contribution to the characterisation of artisanal honey produced on the island of Ireland by palynological and physico-chemical data. *Food Chemistry*, 91(2), 347-354.

European Union (2001). Council Directive 2001/110/EC relating to honey.

Lees, M. (2003). Food authenticity and traceability. Woodhead Publishing Limited. ISBN 1 85573 526 1.

Lutier, P. and Vassière, B. (1993). An improved method for pollen analysis of honey. *Review of Paleobotany and Palynology,* 78, 129-144.

Naes, T., Isaksson, T., Fearn, T. and Davies, A.M.C. (2002). A user-friendly guide to multivariate calibration and classification. NIR Publications, Chichester, UK. ISBN 0 9528666 2 5.

Serra Bonheví, J. and Granados Tarrés, E. (1993). Physicochemical properties, composition and pollen spectrum of ling heather (*Calluna vulgaris* (L) Hull) honey produced in Spain. *Apidologie*, 24, 586-596.

PUBLICATIONS FROM THIS RESEARCH

Downey, G., McIntyre, P. and Davies, A.N. (2002). Detecting and quantifying sunflower oil adulteration in extra virgin olive oils from the eastern Mediterranean by visible and near infrared spectroscopy. *Journal of Food and Agricultural Chemistry,* 50(20), 5520-5525.

Contal, L., Léon, V. and Downey, G. (2002). Detection and quantification of apple adulteration in strawberry and raspberry purées using visible and near infrared spectroscopy. *Journal of Near Infrared Spectroscopy,* 10, 289-299.

Downey, G., McIntyre, O. and Davies, A.N. (2003). Geographic classification of extra virgin olive oils from the eastern Mediterranean by chemometric analysis of visible and near-infrared spectroscopic data. *Applied Spectroscopy,* 57(2), 158-163.

Kelly, J.D., Downey, G. and Fouratier, V. (2004). Initial study of honey adulteration by sugar solutions using mid-infrared (MIR) spectroscopy and chemometrics. *Journal of Agricultural and Food Chemistry,* 52(1), 33-39.

Downey, G. and Kelly, J.D. (2004). Detection and quantification of apple adulteration in diluted and sulphited strawberry and raspberry purées using visible and near infrared spectroscopy. *Journal of Agricultural and Food Chemistry,* **52**(2), 204-209.

Reid, L.M., O'Donnell, C.P. and Downey, G. (2004). The potential of SPME-GC and chemometrics to detect adulteration of soft fruit purées. *Journal of Agricultural and Food Chemistry,* 52(3), 421-427.

Downey, G., Fouratier, V. and Kelly, J.D. (2004). Detection of honey adulteration by addition of fructose and glucose using near infrared transflectance spectroscopy. *Journal of Near Infrared Spectroscopy,* 11(6), 447-456.

Reid, L.M., O'Donnell, C.P., Kelly, J.D. and Downey, G. (2004). Preliminary studies for the differentiation of apple juice samples by chemometric analysis of solid-phase microextraction-gas chromatographic data. *Journal of Agricultural and Food Chemistry*, 52(23), 6891 - 6896.

Downey, G., Hussey, K., Kelly, J.D., Walshe, T.F. and Martin, P.G. (2005). Preliminary contribution to the characterisation of artisanal honey produced on the island of Ireland by palynological and physico-chemical data. *Food Chemistry*, 91(2), 347-354.

León, L., Kelly, J.D. and Downey, G. (2005). Detection of apple juice adulteration using near infrared transflectance spectroscopy. *Applied Spectroscopy,* 59 (5), 63-69.

Kelly, J.F.D. and Downey, G. (2005). Detection of sugar adulterants in apple juice using Fourier transform infrared spectroscopy and chemometrics. *Journal of Agricultural and Food Chemistry*, 53, 3281-3286.

Reid, L.R., Woodcock, T., O'Donnell, C.P., Kelly, J.D. and Downey, G. (2005). Differentiation of apple juice samples on the basis of heat treatment and variety using chemometric analysis of MIR and NIR data. *Food Research International*, 38, 1109-1115.

Kelly, J.D., Petisco, C. and Downey, G. (2006). Potential of near infrared transflectance spectroscopy to detect adulteration of Irish honey by beet invert syrup and high fructose corn syrup. *Journal of Near Infrared Spectroscopy*, 14, 139–146.

Ashtown Food Research Centre

RESEARCH & TRAINING FOR THE FOOD INDUSTRY

Ashtown, Dublin 15, Ireland.

Telephone: (+353 1) 805 9500

Fax: (+353 1) 805 9550