Food Chemistry **29** (1988) 233–238

Technical Note

Discriminant Analysis of Black Tea by Near Infrared Reflectance Spectroscopy

ABSTRACT

A new method for the discrimination of commercial black tea samples using near infrared (NIR) reflectance spectroscopy has been investigated. NIR data at four wavelengths (1660, 1720, 2050 and 2300 nm) corresponding to maximum variation in the intensity of absorption bands in the spectra of black teas were used to calculate Mahalanobis distances. A cut-off point in the values of these distances was determined by means of which two sets of black teas with differing sensory properties could be discriminated with a 91% success rate.

INTRODUCTION

Near infrared (NIR) reflectance spectroscopy has become established as a rapid method of food analysis (Osborne & Fearn, 1986) but its progress has sometimes been limited by dependence on calibration against reference manual procedures. In the food industry, however, sample composition often varies over very narrow ranges and it may be preferable to classify samples; using NIR, it may not be necessary to determine the concentration of a constituent in order to achieve this. For example, it has been shown that it is possible to discriminate between regular and decaffeinated coffee by use of NIR spectra without calibration against caffeine content (Davies & McClure, 1985), between substances used in the manufacture of pharmaceutical products (Mark & Tunnell, 1985; Mark, 1986) and between wheat varieties on the basis of their bread baking quality (Bertrand *et al.*, 1985; Devaux *et al.*, 1986).

In the tea industry, quality is assessed by professional tasters who take

Food Chemistry 0308-8146/88/$03·50

into account the appearance of the dry tea as well as various characteristics of the brewed liquor. This procedure is clearly slow, labour intensive and subjective. NIR has been investigated as a means of achieving a more rapid, automated and objective test for black tea quality by calibration against sensory profile data (Hall, 1986). However, the success of this approach was limited by the imprecise nature of the calibration data. We set out therefore to test the feasibility of discriminating black teas of differing quality by use of NIR without calibration against sensory profiles.

MATERIALS AND METHODS

Samples

For the selection of wavelengths a set of 38 widely differing black teas representing seven growing areas (China, Northern India, Southern India, Bangladesh, Sri Lanka, East Africa, Central Africa) was used. For the discriminant analysis experiment two sets, A and B, comprising 28 and 27 samples, respectively, were prepared by collecting commercial teas (set A) and mixing each with 15% by weight of an adulterant tea (set B).

NIR spectroscopy

The NIR instrument used was a Pacific Scientific Mk I 6350 Research Composition Analyzer which was allowed a 1 h warm-up time before use. Each sample was well mixed before measurement but no sample preparation was carried out. The sample was packed into a special holder which contains a sample thickness of 10 mm between a spring-loaded back plate and a glass window. The sample holder was placed in the instrument in a fixed orientation for measurement of the NIR spectrum. Data were recorded as $\log(1/R)$, where R is the reflectance, at 2 nm intervals in the range 1200–2400 nm. R is defined as P_s/P_0 where P_s is the power of radiation reflected from the sample and P_0 the power of radiation reflected from a ceramic standard measured prior to each sample. Each spectrum (including that of the ceramic standard) was recorded as an average of 50 scans taken over a period of 30 s and samples were scanned in random order.

Statistical analysis

The $\log(1/R)$ data for the wavelength selection set were transformed to second derivative $d^2 \log(1/R)/d\lambda^2$ using a gap size of 20 nm and distance between gaps of 20 nm. The standard deviation, expressed as percentage of the greatest, of the second derivative data for the 38 samples at each wavelength was plotted. Log $(1/R)$ data for the 55 teas in sets A and B at the four wavelengths of maximum variance corresponding to absorption bands in the spectra were used to calculate Mahalanobis distances (Gnanadesikan, 1977).

RESULTS AND DISCUSSION

The use of NIR for qualitative or discriminant analysis is in its infancy and several methods of accomplishing this have been proposed. Methods for the actual discriminant analysis are fairly standard but require some means of prior data reduction or wavelength selection for application to NIR. This data reduction has been accomplished by Fourier transform (Davies & McClure, 1985), principal components analysis (Bertrand *et al.*, 1985) or wavelength selection by the nearest neighbour method (Mark & Tunnell, 1985). Of these methods, the first two allow most of the spectral variation in the samples to be used for discrimination but suffer from the disadvantage that they utilize artificial variables which cannot be directly related to chemical absorption bands in the spectra. Similarly, the nearest neighbour method of wavelength selection is a purely mathematical process which takes no account of spectroscopic factors. In this paper it is proposed that a small number of wavelengths for discriminant analysis be selected on the basis of maximum variation in the intensity of absorption bands in the spectra of the samples.

In order to find the areas of maximum variance in the NIR spectra of black tea, a set of samples incorporating as much natural variation as possible was selected. The NIR data were transformed to the second derivative of the raw $\log(1/R)$ to minimise effects due to differences in particle size between the samples so that wavelengths corresponding to absorption rather than scattering differences could be selected. Figure 1 shows the mean spectrum of the 38 samples in both the $\log(1/R)$ and second derivative forms and Table 1 lists the wavelengths corresponding to relative standard deviations of greater than 30% in the standard deviation spectrum. By comparison of the data presented in Fig. 1 and Table 1 it is possible to identify the wavelengths of maximum variance which correspond to absorption bands in the NIR spectra of the tea samples. It should be noted at this point that minima in the second derivative spectrum correspond to maxima in the $\log(1/R)$ spectrum and it is these that are to be identified. Six wavelengths emerged from this analysis (Fig. 1) of which 1440 and 1920 nm are known to correspond to absorption bands due to water. Since their use would result in discrimination on the basis of differences in moisture

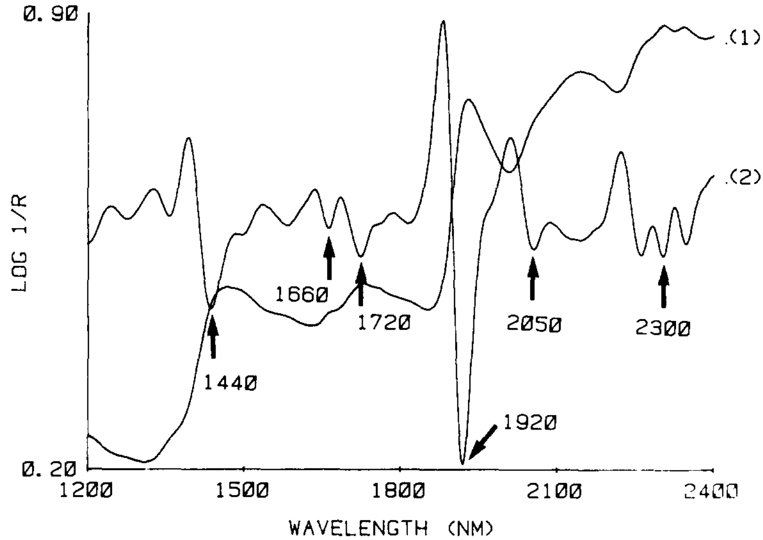

Fig. 1. Mean NIR spectrum for 38 samples of black tea plotted as log (1/*R*) (1) and its second derivative (2).

content, they were rejected and the remaining four used for discriminant analysis. These wavelengths are tentatively assigned to aromatic (1660 nm), methylene (1720, 2300 nm) and hydroxyl (2050 nm) functional groups.

The next stage in the investigation was to test the discriminant properties of NIR data at these wavelengths using samples with a much narrower range of variability, such as would be encountered in commercially available black teas. Since the object of the work was to discriminate between samples

TABLE 1
Standard Deviation of NIR Spectra of 38 Black Tea Samples at Given Wavelengths as % of Greatest

Wavelength (nm)	*Relative standard deviation* (%)
1 200	31
1 320	33
1 390	50
1 440	64
1 660	38
1 720	30
1 880	75
1 920	100
2 010	37
2 050	32
2 240	40
2 300	64

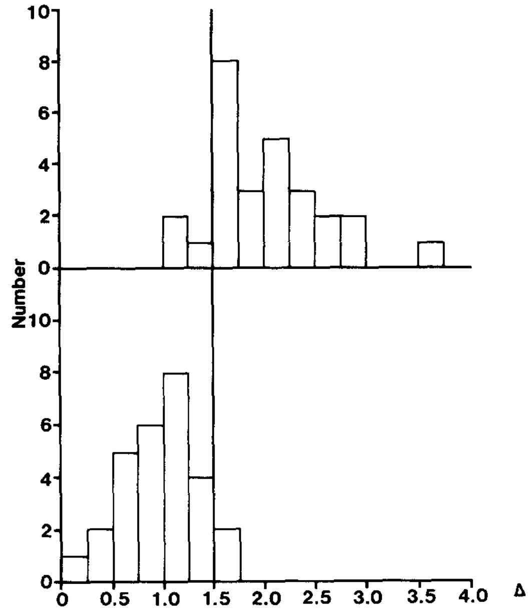

Fig. 2. Histograms of Mahalanobis distances for two sets of black tea samples; bottom diagram represents set A and top diagram set B.

without recourse to calibration against sensory profiles, model mixtures were prepared of representative commercial black teas adulterated with an extra constituent in a quantity known adversely to affect the sensory profile. Using log $(1/R)$ data at the $p = 4$ selected wavelengths and a metric defined by $1/p$ times the variance matrix from set A, the Mahalanobis distance (Δ) (Gnanadesikan, 1977) of each of the 55 samples from the mean for set A was calculated. Figure 2 shows the results for A and B in the form of histograms of the Δ values together with a suggested cut off point of 1·50. Use of this cut off point, in the discrimination of the two sets, results in five errors out of 55 samples or a success rate of 91%. The performance here is being assessed on a calibration set and prediction results might be expected to be slightly worse. Nevertheless it may be concluded that there is a reasonable chance of successful discrimination between black teas of differing sensory profile on the basis of wavelengths corresponding to absorption bands in the spectra of teas. Further, since log $(1/R)$ data at only four wavelengths were required, such discrimination would be possible using a simple filter instrument rather than the sophisticated spectrophotometer used in this study. However, since only one adulterant was examined, the results presented should be regarded as a preliminary to more extensive studies involving a wider range of variables. It is also considered important that a chemical basis be established to justify a relationship between the reported absorption bands and black tea quality.

REFERENCES

Bertrand, D., Robert, P. & Loisel, W. (1985). Identification of some wheat varieties by near infrared reflectance spectroscopy. *J. Sci. Food Agric.*, **36**, 1120.

Davies, A. M. C. & McClure, W. F. (1985). Near infrared analysis in the Fourier domain with special reference to process control. *Anal Proc.*, **22**, 321.

Devaux, M. F., Bertrand, D. & Martin, G. (1986). Discrimination of bread-baking quality of wheats according to their variety by near infrared reflectance spectroscopy. *Cereal Chem.*, **63**, 151.

Gnanadesikan, R. (1977). *Methods for statistical data analysis of multivariate observations.* John Wiley, New York, Chapter 4.

Hall, M. (1986). The application of NIR analysis to foodstuffs with special reference to tea. *Proc. IV International NIR User Conference*, Pacific Scientific, Marlow.

Mark, H. (1986). Normalized distance for qualitative near infrared analysis. *Analyt. Chem.*, **58**, 379.

Mark, H. & Tunnell, D. (1985). Qualitative near infrared reflectance analysis using Mahalanobis distances. *Analyt. Chem.*, **57**, 1449.

Osborne, B. G. & Fearn, T. (1986). *Near infrared spectroscopy in food analysis.* Longman Group, Harlow.

B. G. Osborne and T. Fearn
Flour Milling and Baking Research Association,
Chorleywood, Herts., WD3 5SH, UK

(Received 6 February 1987; revised version received 13 November 1987; accepted 13 November 1987)