MODEL-BASED DISCRIMINANT ANALYSIS OF NEAR-INFRARED SPECTROSCOPIC DATA IN FOOD AUTHENTICITY.

Deirdre Toher*1, Gerard Downey2 and Thomas Brendan Murphy1

* Teagasc, The National Food Centre, Ashtown
1 Department of Statistics, Trinity College Dublin.

After the food scares of recent years, food authenticity studies have become an increasingly important tool for boosting consumer confidence. Honey is a highly variable natural substance that is relatively expensive to produce, therefore is more likely to be the subject of intentional adulteration. As honey is amongst the most difficult of foodstuffs to authenticate, it should be possible to generalise techniques developed specifically for honey to use for other foodstuffs.

Near-infrared spectroscopy is an expensive, non-destructive analytical technique, thus it is ideally suited for food authenticity studies. The technique is useful in preliminary studies where the aim is to develop methods to reliably classify samples into “pure” and “needs further testing, but probably adulterated”. As can be seen from figure 1a, authentic honey, even when from the same geographic region, is extremely variable – the wider the line, the more variable the honey samples are at that wavelength. Indeed the region of the near-infrared spectrum with the greatest variability happens to represent the natural sugars in honey. Honey is often adulterated by these sugars, or by compounds which echo the spectral composition of these sugars.

Thus visually examining the spectra alone is unable to provide satisfactory classification results. Therefore, statistical procedures are required in order to ensure that the analysis of the data is performed in a consistent manner.

25 honey samples from throughout Ireland were adulterated with adulterant solutions consisting of fructose and glucose in the following ratios: 0.7:1, 1.2:1 and 2.3:1 weight/weight (w/w), each of the following three consistent manners.

The dimension reduction techniques demonstrated above yield quite promising results, given that only one approach towards classification was examined. Using the F-statistic proved to be a surprising effective technique, for one that can be explained to those with limited knowledge of statistics.

The authors would like to thank Dr. William J. Coen for helpful discussions and support during this project.

This work was funded by the Teagasc, under the Walsh Fellowship Scheme.

References