The virtual cup-testers' panel: near infrared transflectance spectra to forecast coffee quality

M. Petracco[a] and G. Della Riccia[b]

[a] *illycaffè S.p.A., Research and Technology Department, via Flavia 110, 34147 Trieste (Italy)*
[b]*University of Udine, Department of Mathematics and Informatics, via delle Scienze 206, 33100 Udine (Italy)*

Introduction

Coffee quality

Coffee, broadly referred to as the second most globally traded goods after petroleum, is not needed for nutritional purposes but is much appreciated for its taste appeal, along with its physiological effects on alertness and performance. The assessment of quality is of paramount importance to both of these aspects, in order to supply the customers with a pleasant and wholesome product. This applies especially to espresso coffee,[1] a way to enjoy a cup of coffee that is gaining large popularity throughout the world because of the greater sensory impact it has on the consumer, when compared with coffee brews prepared by other methods.[2]

Because of the high complexity of the raw seed matrix, and even more so when dealing with the roasted finished product, proximate analysis methods able to dig into the structure of ill-known families of compounds are required. The chemical approach to the sensory sphere benefits nowadays of the development of increasingly sophisticated analytical methods, where the parts per billion of volatile aromas are not the final frontier of detection limits.[3]

Sensory analysis

Anyway, no matter the progress of instrumental techniques, the good old cup-testing approach still remains the ultimate quality assessment tool. After all, the reason why coffee has become so popular, achieving the position of the second most largely consumed beverage after water, is its flavour or, even better, its overall impact on our senses. Sensory evaluation, which used to be considered the magic because "taste is a matter of taste", is nowadays earning the status of a highly respected analytical tool, able to produce key information with good reliability.[4]

In industrial coffee routine, some form of objective evaluation is needed to ascertain product overall quality, along with the constancy of that quality on time and on varying process conditions. The "tool" commonly put to use is a panel of assessors, who may be either coffee experts (professional cup-testers) or naïve consumers after a basic training.[5] The reason for employing more than one people is obvious: by averaging responses, the risk of incorrect judgement due to a possible bad shape or minor illness of one person is minimised. Another panel potential is the synergy that can be gained by debating coffee characteristics among the assessors during open sessions: this procedure may extract more information, since individual sensitivities and perception thresholds may be different.

Unfortunately, implementing such a practice is neither simple nor inexpensive. Cup-testing sessions cannot be too long or frequent during the day, because some fatigue develops after the first dozen of cups or so. This is particularly true for espresso tasting, due to the presence of tiny coffee oil droplets in emulsion,[6] which stick on the tongue and on mouth membranes imparting a lingering after-taste. As a consequence, industry strives to take advantage of sensory data collections, using them as raw experimental data to calibrate instrumental screening methodologies.

Instrumental testing

Near-infrared spectroscopy is a good instance of such a rapid, non-destructive fingerprinting technique. It is based on a absorption measurements of scanned monochromatic near-infrared light, whose energy is dissipated in rotational and vibrational movements of the molecular bonds of the material under examination, and ultimately transformed into heat.[7] Energy absorption patterns contain a lot of implicit information about molecular response to specific wavelengths and to their combinations. The procedure generally employed to analyse solid materials, opaque to visible light, is called diffuse reflectance. Since infrared radiation is able to penetrate some millimetres under the surface of solid materials, the diction near infrared (NIR) transflectance may apply to better describe the technique used in the present work.

NIR transflectance has been shown to be suitable for supplying simultaneous forecasts of many chemical characteristics of the sample examined, provided that a good calibration has been previously obtained by statistical correlation with traditional, time-consuming analytical methods. This secondary method has widely been used with agricultural products[8] and sometimes in the specific domain of coffee,[9–11] also for modelling sensory data.[12]

Objective of this work

In the present work, we have arranged NIR transflectance data to predict a proprietary overall sensory quality variable, called Merit, in order to allow the screening of large quantities of unground raw coffee samples as a preliminary to the orthodox cup-testing evaluation for acceptance.

Experimental

Materials

Several thousand raw coffee samples from different origins have been examined during seven years in a NIRSystems 6500 analyser (Foss Tecator), equipped with a translational sample presentation module. This module allows to average multiple NIRT scans of approximately 100 g of whole coffee beans, which happens to be the correct sample size used to roasting and brewing a cup for panel testing.

Each of those samples has been further processed and prepared for espresso cup-testing, according to a standardised protocol described elsewhere.[13] The evaluation of several organoleptic variables of the beverage, including the overall Merit, was collected.

Methods

Spectra, consisting in 1050 log 1/*R* data points from 400 to 2498 nm in 2 nm steps, have been processed by a WinISI II version 1.50 calibration software (InfraSoft International). A product library has been built with 1753 spectra of Brazilian raw coffees selected from a data set of 3438 samples, retained after the elimination of some evident spectral outliers and the streamlining of the population in the denser spectral zones (see Figure 1).

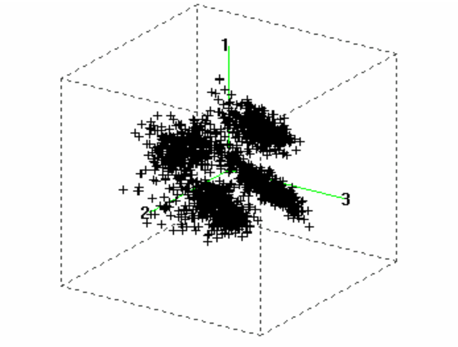

Figure 1. xyz plot of the first three principal components of spectral data point in the library

To predict Merit, we implemented the LOCAL calibration procedure, whose goal is to select from a product library those samples whose spectra are most similar to the target sample, and to use them for producing a dedicated calibration equation for each sample to be predicted. Previous experiments led us to the selection of a specified proprietary wavelength set consisting in only 84 data points in four intervals, which has been used for all predictions of 1447 test-set samples.

For the sake of comparison with a more conventional approach, a regression equation was obtained by GLOBAL calibration with the modified PLS algorithm, allowing 11 terms, cross-validated in four groups.

Results and discussion

Local calibration had been developed to evaluate large data bases of spectra and reference values, using the single sample prediction concept with the aim of performing predictions in real time.[14] It has been shown to give results as good as conventional calibration techniques with partial least squares (PLS).[15]

Calibration statistics

Statistics of our test-set LOCAL predictions, compared to GLOBAL ones, are reported in Table 1.

Table 1. Statistics of predictions obtained by the two methods

	Standar error of prediction (*SEP*)	Regression slope	*R*-squared
LOCAL	0.910	0.587	0.400
GLOBAL	0.949	0.436	0.308

The *R*-squared data may appear not exciting, if compared to those usually obtained with chemical constituents; however, this is not surprising when acknowledging that spectra of raw coffee samples are here being used to predict quality of a coffee beverage prepared through a series of steps, namely roasting, grinding, brewing and cup-testing. Moreover, one should bear in mind that they refer to sensory variables, prone to intra-panel, inter-session and inter-panel noise. Cues on how NIR chemometrics may help filtering some noise out from laboratory data can be found in the literature.[16, 17]

Prediction evaluation

The performance of our LOCAL prediction has been evaluated splitting the 1447 test-set samples in two classes, according to their predicted Merit values: the ones below a given cutoff threshold (parameter) represent those that in industrial practice would be discarded *a priori* without being submitted to the panel's assessment. Assuming that a cup-tested sample would be accepted as "good" by the panel if exceeding a fixed Merit rating, we computed the percentage of discarded samples in the test-set that would have scored "good", and defined it as "missed good coffees". Conversely, the number of samples that would have scored "no good", divided by the number of all discarded ones, was defined as "bypassed bad coffees".

Performance forecasting

The plot of "missed good coffees" and of " bypassed bad coffees" against the parameter "Merit cutoff" is reported in Figure 2 which can be used as an operating curve where one can forecast LOCAL prediction performance, in terms of panel effort savings and related quality giveaway, by entering a cutoff Merit threshold value.

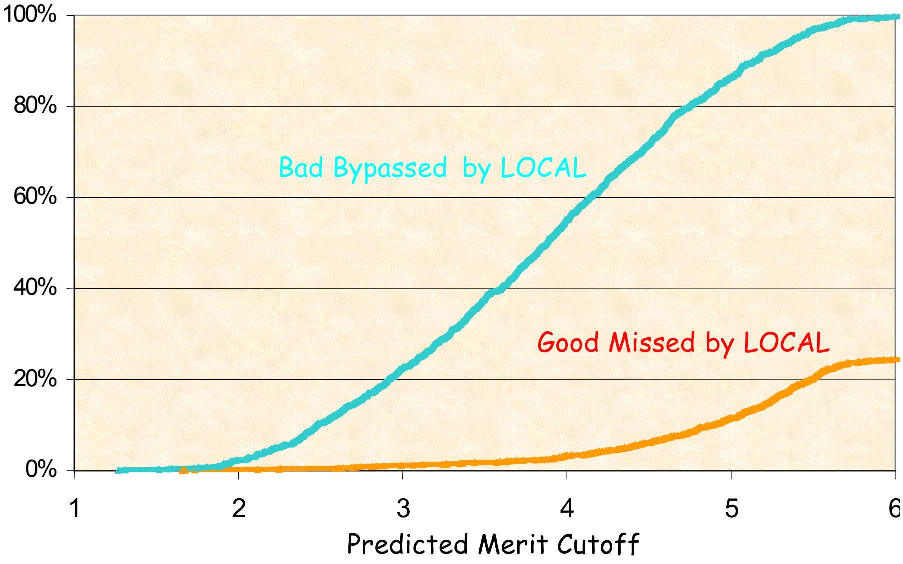

Figure 2. Operating curve to forecast prediction performance.

Comparing calibration algorithms

A similar exercise has been done with GLOBAL predictions, resulting in Figure 3, where the operating curves of both prediction algorithms are compared: LOCAL prediction consents larger savings in the range of interest (around cutoff 3), with comparable good samples' losses.

These aggregated results show that LOCAL data base regression technique performs better than GLOBAL calibration, and is suitable for coffee industrial purposes.

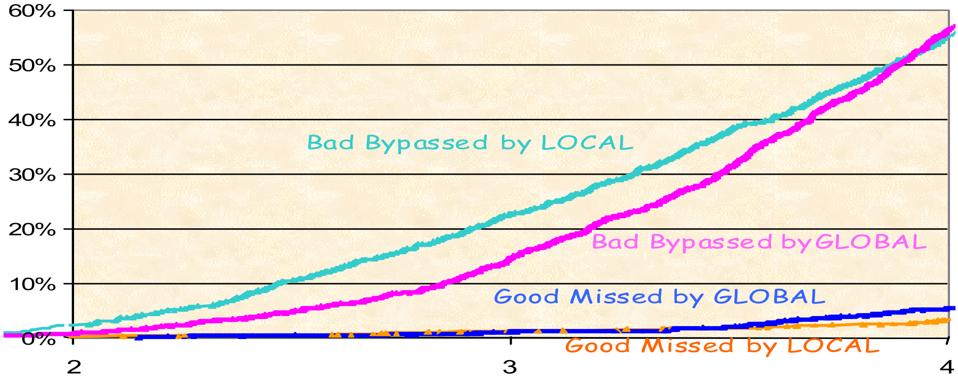

Figure 3. Operating curves from LOCAL and GLOBAL predictions, compared in the key industrial range

It is worth noticing that both savings and losses grow monotonically when the rejection cutoff is raised, as shown in Figure 4 obtained combining the two curves of Figure 2 by elimination of the Merit cutoff parameter. This means that there is no such a thing as an optimum cutoff threshold, which must be chosen by compromise to other factors, like available resources limits or acceptable losses.

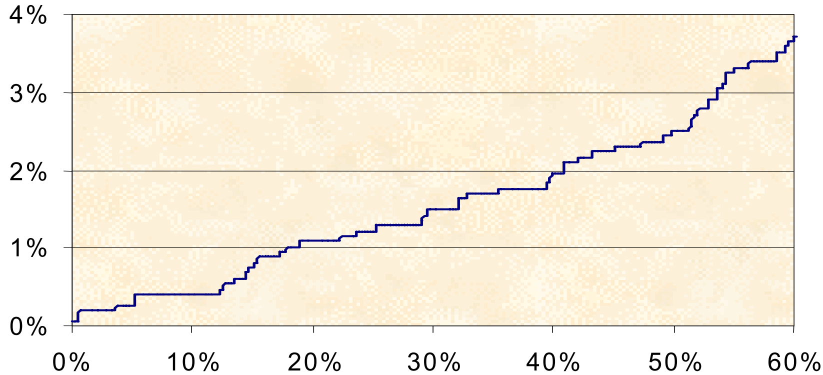

Figure 4. Prediction performance diagram, showing the percent of missed good samples (y) as a function of the percent of bypassed bad samples (x).

Conclusions

Few everyday experiences can compete with a good cup of coffee, as long as sheer sensory pleasure is considered. It is clear that most of the quality of such a beverage is determined by its overall sensory impact. In this context, espresso is the brewing method that offers the consumer the most powerful experience, even if a high quality cup is not easy to obtain.

The relevant conventional quality assurance tool, cup-testers' panel, is expensive and time consuming; this often makes the number of tested samples per day inadequate to meet with industrial needs.

A screening procedure, able to discard a priori those undesirable samples that are likely to be rejected by the panel, would be therefore of help if rapid, non destructive and worthy of being depended on. Near Infra Red Transflectance spectra acquisition fulfils the two former conditions: the last one must be checked by setting up an evaluation procedure.

In this paper we showed how such a procedure can be designed and applied to appraise the prediction of espresso beverage quality from spectral data of raw coffee samples.

References

1. A. Illy and R. Viani (Eds), *Espresso Coffee: The Chemistry of Quality.* Academic Press, London, UK (1995).
2. M. Petracco, in *Coffee Recent Developments*, Ed by R.J. Clarke and O.G. Vitzthum. Blackwell Science Ltd, Oxford, UK, p.140 (2001).
3. O.G. Vitzthum, in *Flavor chemistry. Thirty Years of Progress*, Ed by R. Teranishi *et al.* Kluver Academic/Plenum Publishers, New York, NY, USA, p.117 (1999).
4. M. Meilgaard, G.V. Civille and B.T. Carr, *Sensory Evaluation Techniques 3rd Edition.* CRC Press, Boca Raton, FL, USA (1999).
5. I.C.O., *Sensory Evaluation of Coffee.* International. Coffee Organization, London, UK (1991).
6. M. Petracco, in *Proceedings of the 13th ASIC Colloquium: Paipa (Colombia).* ASIC, Paris, France, p.246 (1989).
7. I. Murray and P.C. Williams, in *Near-Infrared Technology in the Agricultural and Food Industries,* Ed by P. Williams and K. Norris. American Association of Cereal Chemists, St Paul, MI, USA, p.17 (1987).
8. J.S. Shenk, in *Near infrared spectroscopy, Bridging the gap between Data Analysis and NIR Applications*, Ed by K.I. Hildrum, T. Isaksson, T. Naes and A. Tandberg. Ellis Horwood, Chichester, UK, p. 235 (1992).
9. F. Davrieux, S. Laberthe, B. Guyot and J.C. Manez,. in *Proceedings of the 19th ASIC Colloquium: Trieste (Italy).* ASIC, Paris, France (2001).
10. G. Downey and J. Boussion, in *Near Infrared Spectroscopy: The Future Waves*, Ed by A.M.C. Davies and P. Williams. NIR Publications, Chichester, UK, p.410 (1996).
11. B. Guyot, F. Davrieux and J.C. Vincent, in *Proceedings of the 15th ASIC Colloquium: Montpellier (France).* ASIC, Paris, France, p.626 (1993).
12. A.M. Feria-Morales, in *Proceedings of the 14th ASIC Colloquium: S.Francisco (USA).* ASIC, Paris, France, p.622 (1991).
13. M. Petracco, in *Coffee Biotechnology and Quality*, Ed by T. Sera , C.R. Soccol, A. Pandey and S. Rousssos. Kluwer Academic Publishers, Dordrecht, The Netherlands, p.347 (2000).
14. J.S. Shenk, *WinISI II Manual.* Infrasoft international, Port Matilda, PA, USA (2000).
15. F.E. Barton, J.S. Shenk, M.O. Westerhaus and D. Funk, *J. Near Infrared Spectrosc.* **8,** 201 (2000).
16. D.B. Coates, *Spectrosc. Europe*, **14(4),** 24 (2002).
17. J. Lu and W.F. McClure, *J. Near Infrared Spectrosc.* **6,** 77 (1998).