Application of the parallel factor analysis to visible near infrared data for the control of the malting process

Nathalie Allosio-Ouarnier,[a] Patrick Boivin,[a] Philippe Courcoux[b] and Dominique Bertrand[c]

[a]*IFBM, Pôle Technologique de Brabois, 7 rue du Bois-de-la-Champelle, F-54512 Vandoeivre Cedex, France.*

[b]*ENITIAA/INRA, Laboratoire de Statistique Appliquée à la Caracterisation des Aliments, rue de la Géraudière, BP 82225, F-44322 Nantes Cedex 3, France.*

[c]*INRA, PHYTEC Laboratory, rue de la Géraudière, F-44316 Nantes Cedex 3, France.*

Introduction

The malting process consists of three main steps. Barley grains are firstly steeped in order to increase their water content. Thereafter, malting involves germination of grains in order to modify the endosperm and to accumulate various hydrolytic enzymes. Finally, the grain has to be kilned in order to check its growth, to develop the future flavours and colour which contribute to the final beer and to create a dry and stable product to allow malt storage. The quality of a beer depends on the quality of the malt used for its production. The malting process has, therefore, to be carefully controlled. The maltster has to perform several analyses to determine the state of its products. Reference analyses are time-consuming and are impossible to achieve regularly during the process. Numerous works have dealt with the use of near infrared (NIR) spectroscopy as a means to measure the biochemical content of barley and malt[1–6] or to predict reference analysis results from barley spectra.[7] However, few authors have been interested in controlling the malting process with this technique.[8,9]

In this study, we intend to classify samples collected daily during the malting process using NIR. The spectral collection could be stored in a three-way data table where the three modes are respectively the batches, the wavelengths and time. This data structure led us to use a multi-dimensional analysis for studying the spectra. Three-way factor analysis has been used for various kinds of spectral studies such as the interpretation of fluorescence spectra where the three modes were the samples, the excitation wavelengths and the emission wavelengths.[10–12] The parallel factor analysis (PARAFAC) was used to study the effect of instrumental settings such as excitation and emission slit widths on the fluorescence spectra recorded. In this case, the data constituted a three-way array with modes (excitation slit widths) × (emission wavelengths) × (emission slit widths).[13] For the analysis of three mode tables, PARAFAC has proved to be very efficient and relevant.

The present work is aimed at evaluating the PARAFAC model as a means of separating samples according to their malting step and barley varieties.

Near Infrared Spectroscopy: Proceedings of the 9th International Conference

Experimental

Sample collection

Studies were performed on four varieties of spring barley, named Pyramid, Ferment, Alexis and Celinka and one variety of winter barley, named Tulip. The samples were collected daily from a pilot plant during the malting process from the end of steeping until the end of germination. The steeping consisted of three wet steeps. The germination process lasted five days at 16°C for the spring barleys and at 18°C for the winter barley. It was possible to collect samples from 13 batches of 600 kg of barley. The codes of the studied batches are given in Table 1.

Before recording spectra, two aliquots (15 g) of each collected sample were processed in a laboratory grinder (A10 IKA grinder, Bioblock, France) for one minute.

Reference analyses

The water content was measured daily for each batch by drying an aliquot (5 g) of ground germinating barley in an oven at 130°C for 12 hours. For each batch the results obtained daily were averaged to assess the average water content of the considered batch. The other reference analyses were performed on resulting malts using the EBC Analytica procedures.[14] The friability reveals the physical extent of the modification. The extract and attenuation values, respectively, reveal the amount of substances that would be solubilised and the amount of fermentescible sugars in a wort produced from the studied malt.

Spectral acquisition

Spectra were recorded in the reflection mode between 400 and 2500 nm with a 2 nm resolution using the spectrometer NIRSystems™ 6500 (NIRSystems, Silver Spring, USA). Three spectra were recorded by aliquot. The spectral collection, therefore, included 468 spectra (2 ground aliquots × 3 spectra × 6 malting steps × 13 batches).

Mathematical procedure

The PARAllel FACtor analysis (PARAFAC) model was first formulated by Harshman[15] and Caroll and Chang[16] for the analysis of three-mode data. In the present application, as we have collected the absorbance values for P wavelengths for the N products at T different periods, data can be stored in a three-way array $\mathbf{X}$ of dimensions $N \times P \times T$ (Figure 1). The PARAFAC model performs a trilinear decomposition of the order R of this array such as:

$$x_{ijk} = \sum_{r=1}^{R} w_{kr} u_{ir} v_{jr} + e_{ijk} \qquad i = 1,...,N,\ j = 1,...,P,\ k = 1,...T$$

Table 1. Batch codes.

Batch codes	Variety
321/322/323	PYRAMID
324/326	FERMENT
327/330	ALEXIS
328/329/331/332	CELINKA
333/334	TULIP

If $\mathbf{X}_k$ is $N \times P$ matrix, representing the k^{th} slice of $\mathbf{X}$, a matricial expression of the PARAFAC model can be written as:

$$\mathbf{X}_k = \mathbf{U}\mathbf{W}_k\mathbf{V}' + \mathbf{E}_k$$

where $\mathbf{U}$ is an $N \times R$ matrix of scores u_{ir}, of the samples, $\mathbf{V}$ is a $P \times R$ matrix of the loadings v_{jr} of the variables, $\mathbf{W}_k$ a diagonal $R \times R$ matrix of

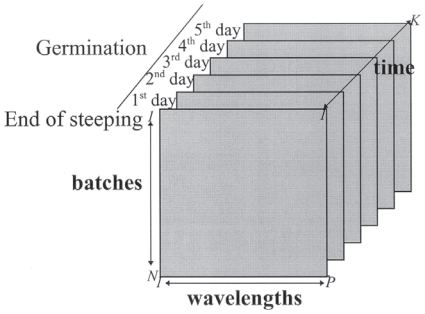

Figure 1. Data structure.

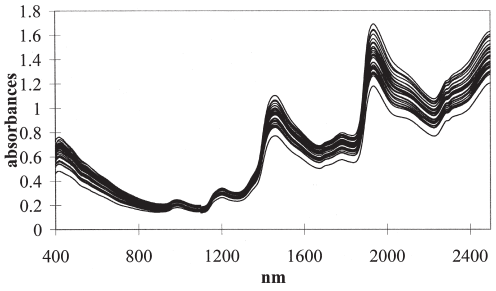

Figure 2. Spectral collection.

weights w_{kr} and $\mathbf{E}_k$ is the $N \times P$ matrix of the residuals e_{ijk}. The scores **U** and the loadings **V** are common but weighted differently for each period k. For a given number R of components, **U**,**V**, and $\mathbf{W}_k$ are estimated so as the norm of **E** is minimised using an alternating least squares algorithm.[17]

In the application presented, matrices **U** and **V** were constrained to be column-wise orthogonal to aid the interpretation of scores and loadings. The data were centred. The centring procedure consisted of subtracting from each absorbance recorded at a given wavelength for a given sample and the mean of the absorbances recorded for all the samples at this wavelength. The trilinear composition was of order 3.

Results and discussion

Reference analyses

The results are shown in Table 2. CELINKA batches 329, 331 and 332 contained more water than the other batches. Indeed, their water content was higher than 44.9%. The friability was relatively correlated with the extract level ($R = 0.65$) and the extract level was relatively correlated with the attenuation level ($R = 0.69$). Lower friabilities were recorded for the TULIP batches (333; 334). This was in accordance with the current knowledge. As a matter of fact, winter barleys generally show a lower modification degree than spring barleys. The two ALEXIS batches showed different physical alteration degrees. Batch 330 was less modified than batch 327. This could be explained by a lower steeping temperature for batch 330. The extract and attenuation values made it possible to set batches 328, 329, 331, 332, 330, 333, 334 against batches 321, 322, 323, 324, 326 and 327. The former have led to malts with extract levels lower than 82.9% and attenuation levels lower than 81.5% whereas the latter have led to malts with extract levels higher than 83% and attenuation levels higher than 82.5%.

Spectral interpretation

Spectral collection.

The recorded spectra are shown on Figure 2. The spectra obtained showed a similar shape for every sample, whatever the batch they belonged to and whatever their malting step. Between 400 and 1100 nm, few spectral variations were observed. They could mainly be due to colour modification or to the rootlet appearance. The small peaks at 1200 nm could be due to lipid absorption.[18] The two intense peaks, 1450 nm and 1950 nm, characteristic of the O–H group of water, were observed and revealed the high moisture content of the samples. Absorption areas of starch were observed between

Table 2. Biochemical analysis results.

Batches	Average water content %	Friability (%)	Extract (%)	Attenuation (%)
321	43.5	96	84.7	84.1
322	42.3	86	84.7	83.5
323	44.7	87	85	84.6
324	44.2	93	83.5	83.1
326	43.4	93	83.8	82.5
327	44.11	91	83	83
330	43.3	88	82.9	84
328	44.3	85	82.5	79.5
329	45.1	92	82.1	81.5
331	44.9	92	81.8	81.5
332	45.6	89	81.8	80
333	43.9	77	79.6	81.3
334	43.3	77	80	81.3
mean	44.2	88.15	82.72	82.30
std	1.26	3.31	1.62	1.53
range	5.77	19	5.4	5.10

1400 and 1600 nm and at 2100 nm. The 1400 to 1600 nm spectral range was, however, difficult to study because of the large amount of water present in the samples. Indeed the water peak at 1450 nm might conceal spectral modifications due to starch. A shoulder at 2100 nm can be observed and used for investigating modifications of the starch content of the samples. It was difficult to see the protein bands because of the presence of large amounts of water. In the 2270–2380 nm region two weak shoulders were seen. This area is reported in the literature as being characteristic of the β-glucans. Allison *et al*. have reported that bands of β-glucans can be seen between 2330 and 2350 nm.[6] Szczodrak *et al*.[4] have found that the 2260 to 2380 nm region was typical for β-glucans. This region might help us to analyse the beginning of the malting process where the cell wall breakdown takes place.

PARAFAC

Among the available methods, Unfold-PCA (U-PCA) and PARAFAC are commonly used. PARAFAC presents several advantages over U-PCA. PARAFAC is a trilinear decomposition and a symmetric model. Indeed, in the PARAFAC model, each mode plays the same role whereas in the U-PCA the mode which is left intact in the unfolding operation undergoes special treatment. Moreover the numbers of parameters to be estimated in the PARAFAC model is considerably lower than for the U-PCA case. This makes the results easier to interpret. All these considerations led us to choose the PARAFAC model.

The spectral range spanning between 400 and 1100 nm makes it possible to exclusively study modifications of grain appearance such as colour variations or rootlet rising. In order to limit ourselves to the study of internal changes, the PARAFAC model was applied between 1100 and 2500.

The map obtained using the scores of the first two components is shown in Figure3. The first component allowed us to identify three groups of batches. The first one corresponded to Celinka batches

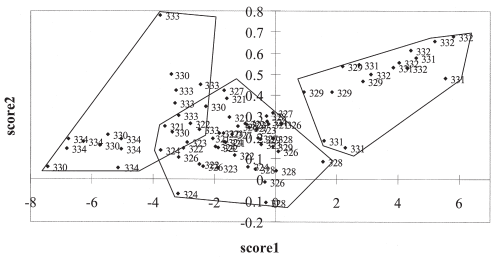

Figure 3. Map using the scores of the first two components.

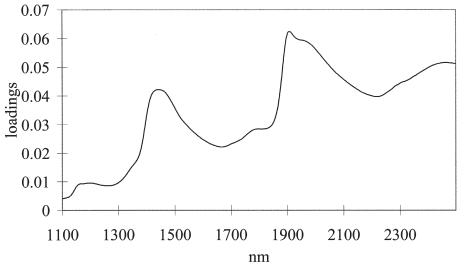

Figure 4. Loading vector associated with the first component.

329, 331 and 332. The second one included PYRAMID batches 321, 322, 323, FermenT batches 324, 326, ALEXIS batch 327 and Celinka batch 328. Then, ALEXIS batch 330 and TULIP batches 333 and 334 constituted the third group. The loading vector (on the wavelength-mode) associated with the first component (Figure 4) was similar to every spectra of the recorded collection. The differences between batches were mainly due to increases in the absorbance values. Batches 330, 333 and 334 were less modified and, on average, contained less water than the other batches. The spectral intensity seemed to be related to the water content and to the modification degree. The greater the spectral intensity, the better the modification occurring in malting. The weights, w_{k1} (Figure 5) associated with the first component decreased as the germination and, therefore, the modification progressed. The decrease in w_{k1} revealed the modification progress and particularly between the first and the second day of germination and between the third and the fourth day of germination. The second component did not allow the discrimination between batches. However the associated w_{k2} (Figure 5) decreased steadily and principally between the first and the third day of germination. The loading vector (on the wavelength-mode), associated with the second component, showed two main peaks (Figure 6). The first one, at 1410 nm could be due to water changes or to the modification of the cristallinity of one or several sugars.[19] The second one, around 1940 nm, can be split up into a peak situated at 1900 nm and a shoulder at 2000 nm. That can be due to variations of the spectral intensity combined with a shift of the water absorption band characteristic of soluble substance appearance.[20]

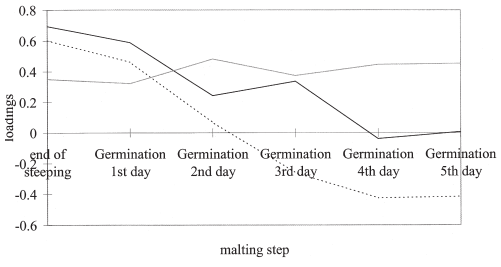

Figure 5. Weights on the time-mode associated with the first (___), the second (---) and the third (___) components.

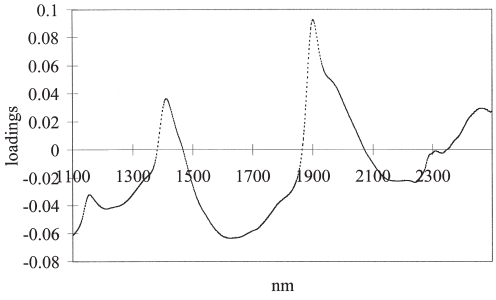

Figure 6. Loading vector associated with the second component.

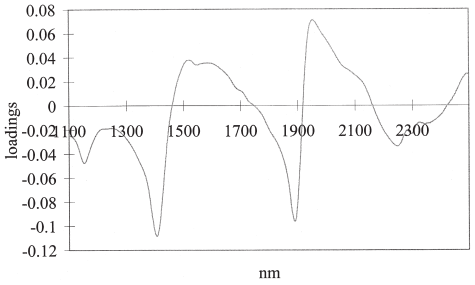

Figure 7. Map using the scores of the first and third components.

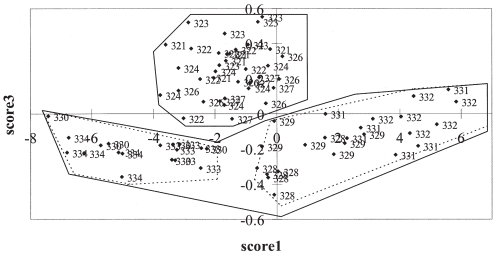

Figure 8. Loading vector associated with the third component.

The map obtained using the scores of the first and the third component is shown in Figure 7. The third component made it possible to set batches 328, 329, 331, 332, 330, 333, 334 against batches 321, 322, 323, 324, 326 and 327. As was seen earlier, the former have led to malts with extract levels lower than 82.9% and attenuation levels lower than 81.5% whereas the latter have led to malts with extract levels higher than 83% and attenuation levels higher than 82.5%. This revealed that the starch granules of batches 321, 322, 323, 324, 326 and 327 seemed to be more accessible to amylolytic enzymes or that these batches contained more starch. If this statement is true, the loading vector associated with the third component (Figure 8) should show a characteristic region of starch. A shoulder can actually be observed at 2100 nm. The associated w_{k3} (Figure 5) did not show any significant increase or decrease during malting. The starch composition seemed relatively stable during germination.

Conclusion

Near infrared spectroscopy combined with a three-way factor analysis, PARAFAC, was used to study the variation of barley during its transformation into malt. It was possible to observe differences between the studied batches and, within a batch, it was possible to study the variation of samples collected at different steps of the process. However, the presence of water in significant amounts and spectral intensity variations concealed characteristic biochemical phenomena of the barley modification and prevented us from determining the actual origins of the differences observed between batches and malting steps.

References

1. P.C. Williams, K.R. Preston, K.H. Norris and P. M. Starkey, *J. Food Sci.* **49(1),** 17 (1984).
2. R.J. Henry, *Journal of the Institute of Brewing* **91,** 393 (1985).
3. Z. Czuchajowska, J. Szczodrak and Y. Pomeranz, *Cereal Chem.* **69(4),** 413 (1992).
4. J. Szczodrak, Z. Czuchajowska and Y. Pomeranz, *Cereal Chem.* **69(4),** 419 (1992).
5. S.A. Halsey, *Journal of the Institute of Brewing* **93,** 407 (1987).
6. M.J. Allison, I.A. Cowe and R. McHale, *Journal of the Institute of Brewing* **84,** 153 (1978).
7. R.J. Henry, *Carbohydrate Research* **141,** 13 (1985).
8. M.J. Allison and A.P. Maule, in *Proceedings of the 23rd EBC Congress*. Lisbon, Portugal, Ed by IRL Press at Oxford University Press, Oxford, UK, pp. 133–138 (1991).
9. J. Tenhunen, M. Svinhufvud, K. Sjöholm and K. Pietilä, in *Proceedings of the 25th EBC Congress*, Brussels, Belgium, Ed by IRL Press at Oxford University Press, Oxford, UK, pp. 585–592 (1995).
10. S. Leurgans and R.T. Ross, *Statistical Science* **7(3),** 289 (1992).

11. R. Bro, *Journal of Chemometrics.* **10,** 47 (1996).
12. P. Courcoux, M.F. Devaux, E. Vigneau and B. Novales, TRICAP, Washington, USA, (1997).
13. L. Norgaard, *Journal of Chemometrics* **10,** 615 (1996).
14. ANALYTICA-EBC. Brauerei- und Getränke-Rundschau, Zurich, Switzerland, 4th Edn (1987).
15. R.A. Harshman, *UCLA Working Papers in Phonetics* **16,** 1(1970).
16. J.D. Caroll and J.J. Chang, *Psychometrika* **35,** 283 (1970).
17. P. M.Kroonenberg and J. De Leeuw, *Psychometrika* **45,** 69 (1980).
18. R.H Wilson, *Spectroscopic Techniques for Food Analysis*, Ed. by R.H. Wilson. VCH Publishers, Weinheim, Germany, p 70 (1994).
19. A.M.C Davies and C.E Miller, *Appl. Spectrosc.* **42(4),** 703 (1988).
20. N. Pedretti, D. Bertrand, M. Semenou, P. Robert and R. Giangiacomo, *J. Near Infrared Spectrosc.* **1(3),** 174 (1993).