ELSEVIER

Food Chemistry 53 (1995) 295–298

0308-8146/95/$9.50

Characterisation of flour by means of pattern recognition methods

Paweł Zagrodzki,[a,b] Małgorzata Schlegel-Zawadzka,[b] Mirosław Krośniak,[b] Przemysław Malec,[c] Andrzej Bichoński[d] & Erazm Dutkiewicz[a]

[a] *H. Niewodniczański Institute of Nuclear Physics, Radzikowskiego 152, 31-342 Cracow, Poland*
[b] *Department of Food Chemistry and Nutrition, Collegium Medicum Jagiellonian University, Podchorażych 1, 30-084 Cracow, Poland*
[c] *Department of Physiology and Biochemistry of Plants, Zurzycki Institute of Molecular Biology, Jagiellonian University, Al. Mickiewicza 3, 31-120 Cracow, Poland*
[d] *Institute of Cultivation and Acclimatization of Plants, Zawiła 4, 30-442 Cracow, Poland*

(Received 6 May 1994; revised version received and accepted 15 August 1994)

Samples of 20 varieties of flour were examined for contents of toxic elements (Pb, Cd) and of some macro- and microelements (Mg, Zn, Cu, Fe, Mn). Chemical composition of flour and its baking characteristics found in the course of routine analysis were also studied. The relationship between both groups of parameters was assessed by the canonical correlation method.

INTRODUCTION

Pattern recognition methods used for analysing multivariate data are considered as a tool for generating a new set of variables, being the linear combination of the original ones. The transformation is designed in such a manner that it enhances hidden properties of the original data. In many cases it allows the analyst to gradually reduce a multidimensional dataset to only a few dimensions and to present it in a new coordinate system. The resulting data set is easier to inspect and gives a better opportunity to recognise patterns within the data structure (Chambers & Kleiner, 1987). The main aim of the present study was to simplify the description of different varieties of flour by applying a pattern recognition approach and to check, simultaneously, if there was any correlation between the typical baking parameters (e.g. sedimentation value, falling number, calorimetric value, protein content, gluten index) and elemental composition.

MATERIALS AND METHODS

The samples of winter wheat originated from the fields of the Plant Experimental Stations in Smolice and Oleśnica Mała, situated in south-west and north-west Poland, respectively. Samples from the same Plant Experimental Station represented several varieties. The wheat milling process was carried out in the Institute of Cultivation and Acclimatization of Plants (Cracow, Poland). Baking parameters were determined by the standard methods of the routine analysis of flour (Wise *et al.*, 1965; Axford *et al.*, 1978).

Metal levels were measured by atomic absorption spectrometry (AAS). The CEM Corporation's Microwave Digestion System MDS-2000 was used for mineralisation of the samples prior to the analysis. The atomic absorption analyses of Cd, Cu, Fe, Mn and Pb were performed with the Perkin–Elmer 5100 PC atomic absorption spectrometer equipped with the 5100 ZL Zeeman Furnace Module. Mg and Zn were measured with the same spectrometer using its flame device. More details regarding the analytical procedure and the validity of the method were given in our earlier work dealing with elemental analysis of organic materials (Zagrodzki *et al.*, 1993).

Multivariate statistical approach

The original data set obtained for 20 samples of flour consisted of features including typical baking parameters and the content of selected minerals. Since the number of baking parameters exceeded (significantly) the number of the objects studied divided by three, some procedures of feature extraction had to be employed to reduce the dimensionality of the parameter data subset (Kowalski & Wold, 1982). An algorithm of feature definition was chosen similar to one of those given by Jain and Dubes, (1977), which was composed of two phases of data treatment. During the first phase the correlation coefficients were calculated as a measure of the features' proximity. The proximity (similarity) matrix of features obtained in such a way was subsequently clustered (*R*-mode cluster analysis). The single clustering method was used to achieve a smaller number of separated clusters grouping the highly correlated features. In the second phase each resulting cluster (if it consisted of more than one variable) was compressed into only one most representative feature. This was done with the help of principal components analysis and consideration of the subsequent statistics. The sta-

Table 1. Means, minimum and maximum values, standard deviations and medians of baking and elemental composition parameters

Parameter	Unit	Mean	Min	Max	Standard deviation	Median
Baking parameters						
Wet gluten content	%	42·9	32·0	54·9	5·6	42·0
Dry gluten content	%	14·3	11	17·4	1·7	14·7
Protein content	%	15	12·6	17·2	1·3	14·9
Water absorbability	%	69·9	64·0	73·9	2·3	69·8
Development of flour	min	2·2	1·5	3·3	0·5	2·0
Valorimetric value	ju	52	44	64	5	52
Stability of dough	min	5·3	2·9	10·6	2·0	4·7
Gluten index		73·7	54·1	99·0	13·7	70·6
Sedimentation value	cm^3	67	60	79	6	67
Falling number	sec	318	70	418	108	366
Volume of bread[a]	cm^3	553	536	562	6	553
Flour extraction	%	61	56	66	3	61
Softening of dough	jB	52	30	90	18	50
Elemental composition parameters						
Fe	ppm	20·6	13·4	55·5	9·1	18·9
Cu	ppm	1·63	0·98	2·53	0·39	1·53
Mn	ppm	14·8	10·9	19·9	1·96	14·7
Cd	ppm	0·058	0·043	0·075	0·008	0·059
Pb	ppm	0·15	0·04	1·45	0·31	0·08
Zn	ppm	14·8	11·1	19·9	2·38	14·3
Mg	ppm	325	240	366	35·1	336

[a]Baked from 100 g of flour.

Table 2. Canonical correlations data

Number	Eigenvalue	Canonical correlation	Wilks Lambda	χ^2	DF	Significance level
1	0·878	0·9368	0·0070	61·94	36	0·0046
2	0·763	0·8735	0·0575	35·69	25	0·0764
3	0·576	0·7589	0·2427	17·70	16	0·3419
4	0·359	0·5995	0·5724	6·97	9	0·6398
5	0·106	0·3257	0·8935	1·41	4	0·84
6	0·0005	0·0230	0·9995	0·01	1	0·94

Table 3. Coefficients for the first two canonical variables of baking and elemental composition parameters

Coefficients for canonical variables of baking parameters		
Development of flour	−0·31	−0·86
Sedimentation value	−0·02	−0·18
Falling number	−0·59	0·68
Volume of bread after baking from 100 g of flour	−0·19	−0·19
Flour extraction	−0·04	0·13
Softening of dough	−0·79	−0·28
Coefficients of canonical variables of elemental composition parameters		
Fe	0·41	−0·70
Mn	0·29	0·23
Mg	−0·71	−0·15
Zn	−0·10	−0·14
Pb	−0·06	−0·14
Cd	−0·22	−0·75

tistical analysis was identical to that used by Moret *et al.* (1986) for the same purpose. The latter procedure was also applied to reduce the subset of the elemental composition parameters from seven to six dimensions. Such selection produced a smaller number of essentially uncorrelated features. Finally, canonical correlation analysis was used to study the relationship between two sets of features, extracted as described above. All calculations were performed using STATGRAPHICS and SYSTAT statistics packages.

RESULTS AND DISCUSSION

The main descriptive characteristics of both baking and elemental composition parameters are summarised in Table 1. The scatter about the mean of baking parameters was smaller than that of elemental composition parameters. *R*-mode cluster analysis provided an interfeature correlation pattern consisting of six clusters: (1) wet gluten content, dry gluten content, protein content, water absorbability, development of flour; (2) calorimetric value, stability of dough, gluten index, sedimentation value; (3) falling number; (4) volume of bread baked from 100 g of flour; (5) flour extraction; (6) softening of dough (the latter four features formed separate classes themselves). The development of flour and sedimentation value were revealed as the most important parameters in clusters 1 and 2, respectively. Thus, the flour description was limited to only six parameters.

Results of canonical correlation analysis are summarized in Tables 2–3 and shown in Fig. 1. The first canonical correlation was 0·9368 with 0·0046 probability level for the hypothesis that all the canonical correlations were 0. The second canonical correlation was less significant. This indicated strong and clear correlations between some parameters of both subsets. From the first canonical correlation it was concluded that the influence of magnesium on flour characteristics is expressed mainly by the softening of dough and the falling number. Magnesium is known as a very important nutrient for plants, because it plays a unique role in plant cells activating a large number of enzymes involved in phosphate transfer (Douce & Day, 1983; Lauddi & Bieleski, 1983). In particular, the magnesium ion is indispensable to energetic processes involved in seed germination and starch mobilisation in plants (Murray, 1984). A positive correlation between some of the baking parameters and the content of high-molecular-weight glutenin in wheat flour has been found (Subda, 1992). Other authors have shown the influence of proteolytic enzymes and α-amylase activities on baking characteristics of wheat flour (Sandstedt, 1961; McDonald, 1969). The results reported here indicate that some of baking parameters correlate positively with magnesium concentration in the flour as well. The findings suggest that nutritional value of flour might be connected with the manifold role of magnesium in plants. Therefore, it is proposed that some technological properties of the flour could be dependent on the magnesium level. However, the nature of this relationship needs further investigation. It is also worth noting that toxic elements having antropogenic origins — cadmium and lead — played a much less important role in the revealed correlation, which is in accordance with expectations.

ACKNOWLEDGEMENT

This work was partly supported by the State Committee for Scientific Research, Grant no. 202599101. The authors are indebted to Mr Władysław Chłopicki for improving the English.

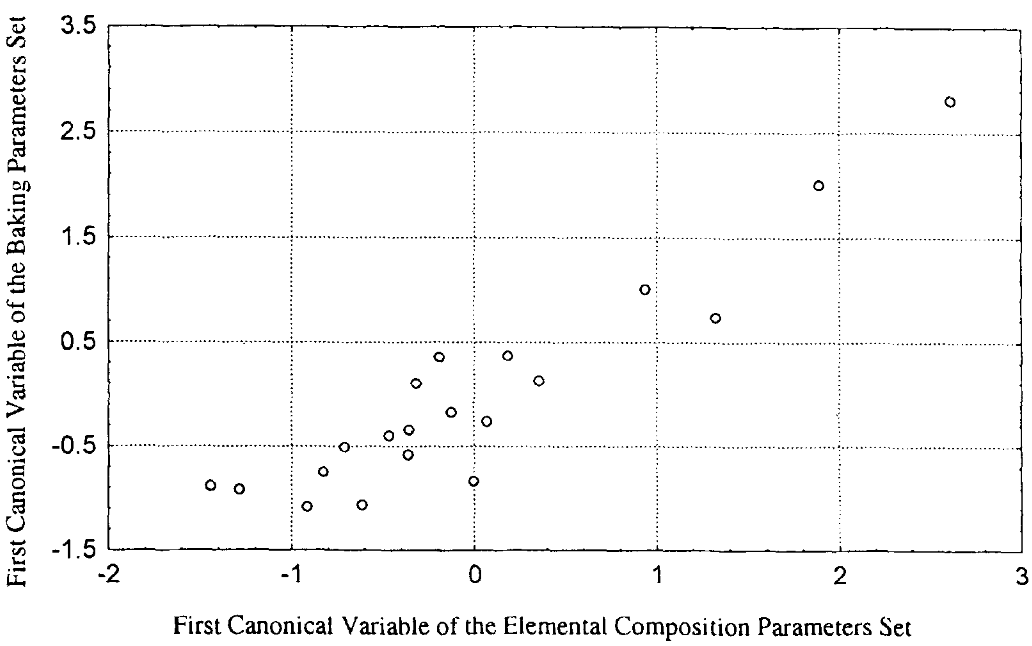

Fig. 1. The scores of flour samples on the first canonical correlation variables.

REFERENCES

Axford, D. W. E., McDermott, E. E. & Redman D. G. (1978). Small scale tests of breadmaking quality. *Milling Feed and Fertilizer*, **66**, 5–18.

Chambers, J. M. & Kleiner B. (1987). Graphical techniques for multivariate data for clustering. In *Handbook of Statistics* (Vol. 2), eds P. R. Krishnaiah & L. N. Kanal. North-Holland Publishing Company, Amsterdam, The Netherlands, pp. 209–44.

Douce, R. & Day, D. A. (eds) (1983). Higher plant cell respiration. In *Encyclopaedia of Plant Physiology*, (Vol. 18). Springer, Berlin, Germany.

Jain, A. K. & Dubes, R. (1977). Feature definition in patternecognition with small samples size. *Pattern Recognition*, **10**, 85–97.

Kowalski, B. R. & Wold, S. (1982). Pattern recognition in chemistry. In *Handbook of Statistics* (Vol. 2), eds P. R. Krishnaiah & L. N. Kanal. North-Holland Publishing Company, Amsterdam, The Netherlands, pp. 673–97.

Lauddi, A. & Bieleski, R. L. (eds) (1983). Inorganic plant nutrition. In *Encyclopaedia of Plant Physiology* (Vol. 15 A, B). Springer, Berlin, Germany.

McDonald, E. C. (1969). Proteolytic enzymes of wheat and their relation to baking quality. *Baker's Dig.*, **43**, 23–32.

Moret, I., Capodaglio, G. & Scarponi, G. (1986). Statistical evaluation of the group structures of five Venetian wines from chemical measurements. *Anal. Chim. Acta.*, **191**, 331–50.

Murray, D. R. (ed.) (1984). *Seed Physiology* (Vol. 2). Germination and reserve mobilization. Academic Press, New York, USA.

Sandstedt, H. R. (1961). The function of starch in the baking of bread. *Baker's Dig.*, **35**, 36–41.

Subda, H. (1992). Assessment of chemical composition of flour and its effect on baking characteristics of wheat. *Polish J. Food. Nutr. Sci.*, **42**, 15–20.

Wise, M., Sneed, E. & Pope, W. K. (1965). Modification in the sedimentation test for the evaluation of small samples of wheat. *Argon. J.*, **57**, 93.

Zagrodzki, P., Dutkiewicz, E. M., Malec, P., Krośniak, M., Knap, W. & Bichoński, A. (1993). *Instrumental Methods for Analysis of Some Elements in Flour.* Raport INP No. 1648/CA, Cracow, Poland.