ELSEVIER

Food Chemistry 84 (2004) 29–33

Food Chemistry

www.elsevier.com/locate/foodchem

Application of partial least squares (PLS) regression to predict the ripening time of Manchego cheese

J.M. Poveda[a,*], A. García[a], P.J. Martín-Alvarez[b], L. Cabezas[a]

[a] *Departamento de Química Analítica y Tecnología de Alimentos, Universidad de Castilla La Mancha, 13071 Ciudad Real, Spain*
[b] *Instituto de Fermentaciones Industriales (CSIC), Juan de la Cierva, 3, 28006 Madrid, Spain*

Received 22 August 2002; received in revised form 21 March 2003; accepted 21 March 2003

Abstract

Partial least squares (PLS) regression was used to predict the ripening time of Manchego cheese, based on some physicochemical parameters and secondary proteolysis indices of 63 standard Manchego cheeses manufactured in different seasons (autumn, winter and spring). PLS regression demonstrated that the variables that most contributed to predict the ripening time of the cheeses were water activity (a_w), pH and dry matter (DM). The prediction model obtained yielded good results for the prediction of the ripening times of commercial Manchego cheeses manufactured in the same factory as the standard cheeses, since the root mean square error of prediction was 11.9 days. The model was also good for the prediction of the ripening times of commercial Manchego cheeses manufactured in different factories.

Keywords: Manchego cheese; Prediction; Ripening time; PLS regression

1. Introduction

Multivariate statistical techniques have been employed in recent years with the aim of analysing, describing and generally interpretting multidimensional observations (Molina & Martín-Alvarez, 1996, Pripp, Shakeel-Ur-Rehman, McSweeney, & Fox, 1999; Pripp, Shakeel-Ur-Rehman, McSweeney, Sorhaug, & Fox, 2000).

Partial least squares (PLS) regression is a multivariate statistical technique that has been applied in food research to obtain calibration models as an alternative to other statistical methods, such as multiple linear regression (MLR) and principal component regression (PCR), (Casal, Martín-Alvarez, & Herraiz, 1996; Molina, Martín-Alvarez, & Ramos, 1999; Pérez-Coello, Martín-Alvarez, & Cabezudo, 1999). In dairy products, PLS regression has been applied to different cheese varieties to find quantitative relationships between sensorial attributes and some chemical variables. Virgili et al. (1994) used PLS regression to predict the flavour in Parmigiano-Reggiano cheese, based on the volatile composition. Frister, Michaelis, Schwerdtfeger, Folkenberg, and Sorensen (2000) employed the same statistical approach in Cheddar cheese to correlate chemical parameters (peptides, caseins and casein degradation products) with the bitter taste of the cheese.

Manchego cheese is the most popular ewe's milk cheese produced and consumed in Spain, is a cured, semi-hard, enzymatically coagulated, uncooked, pressed, high fat cheese. The minimum ripening time of cheeses before marketing, required by the regulatory Board of the Manchego cheese Appellation of Origin, is 60 days. However, the date of manufacture of the cheeses is not required in the market place and the cheeses are marked into two categories: semi-hard cheeses (between 2 and 4 months) and hard cheeses (>4 months). These two categories mean that Manchego cheeses available in the market have different sensory properties and also different prices, so a model for the prediction of the ripening time for Manchego cheese could be of a great interest for marketing purposes.

García-Ruiz, Cabezas, Martín-Alvarez, and Cabezudo (1998) applied different multivariate regression techniques, namely MLR, PCR and PLS, in order to calculate the ripening time of commercial Manchego cheeses based on physicochemical and proteolysis parameters. These authors found that PLS regression yielded the best predictions of the ripening time for

* Corresponding author. Tel.: +34-926-295300; fax: +34-926-295318.
E-mail address: justamaria.poveda@uclm.es (J.M. Poveda).

0308-8146/03/$ - see front matter © 2003 Elsevier Ltd. All rights reserved.
doi:10.1016/S0308-8146(03)00163-8

Manchego cheeses. However, the results were preliminary and it was concluded that it would be necessary to check the predictive capability of the method, manufacturing standard cheese samples in different periods of the year, and using a larger number of commercial cheese samples, in order to draw a conclusion with a more statistical basis.

The aim of this work was to improve the model proposed by García-Ruiz et al. (1998), considering a higher number of samples for the prediction model, and also to reduce the number of variables in the predictive equation, with the ultimate objective of predicting the ripening time of Manchego cheeses.

2. Material and methods

2.1. Cheese samples

2.1.1. Standard cheeses

Three batches of 21 cheeses each were manufactured in three different seasons (autumn, winter and spring). They were made on a semi-industrial scale, using the Manchego cheese technology, at a cheesemaking plant located in the La Mancha region. The freshly made cheeses (1.5 kg each) were brought to the laboratory under refrigerated conditions and ripened at 13 °C and 80% RH. Three cheeses were taken for individual analysis at 15, 30, 45, 60, 90, 120 and 150 days of ripening. These 63 cheeses were considered as calibration samples for obtaining the prediction model.

2.1.2. Commercial cheeses

Twelve Manchego cheeses, manufactured in the same factory and under the same conditions as the standard cheeses, and ripened for 45, 60, 90 and 150 days, were used as evaluation samples to predict their ripening times (samples from 1 to 12). Three cheeses were analysed for each ripening time. As the ripening times of these cheeses were exactly known, these samples were used for validation of the prediction model.

Fourteen other commercial Manchego cheeses, eight of them ripened between 2 and 4 months (samples from 13 to 20) and the rest ripened for >4 months (samples from 21 to 26), as was indicated by the cheesemakers, were purchased from different cheese factories and were also considered for the prediction of their ripening times.

2.2. Analytical methods

2.2.1. Compositional analysis

Dry matter (DM) was analysed according to the standards of the AOAC (1995); pH was determined by direct reading, using an Ingold insertation eletrode; water activity (a_w) was measured directly using a Decagon Devices model CX-2 dew-point hygrometer; protein and NaCl contents were measured using the methods of AOAC (1995).

2.3. Assessment of proteolysis

The water-soluble nitrogen (WSN) determination was according to Kuchroo and Fox (1982). Phosphotungstic acid-soluble nitrogen (PTASN) was determined by the method of Bütikofer, Rüegg, and Ardö (1993) and tyrosine (Tyr) and tryptophan (Trp) concentrations were analysed following the method of Vakaleris and Price (1959).

2.4. Statistical methods

PLS regression was used for the prediction of the ripening times of commercial Manchego cheese samples, based on the parameters analysed, using an equation of the form

$$t_i = b_0 + \sum_{j=1}^{m} b_j x_{i,j}$$

where t_i was the ith cheese ripening time, b_0 was the y-intercept, and b_j was the regression coefficient for the jth prediction parameters (X_j) in the model. The contribution of each variable to the prediction of the ripening time was evaluated using the regression coefficients obtained for the standardised variables. These coefficients allow selection of those variables that most contributed to the prediction of the ripening time.

The $RMSEP_{cv}$ (root mean square error of prediction obtained by cross-validation) was used as a measure of the ability of the model to furnish accurate predictions (Martens & Naes, 1993). The value was calculated using the equation:

$$RMSEP_{cv} = \left(\sum_{i=1}^{n} \left(t_i - \hat{t}_{(i)} \right)^2 / n \right)^{1/2}$$

where t_i is the real ripening time for the ith sample of the standard cheeses, $\hat{t}_{(i)}$ is the predicted ripening time obtained with the model constructed without the ith sample, and n was the number of standard cheeses used in the calibration model ($n = 63$).

Calculations were performed using The Unscrambler program, version 7.01 (Camo, ASA, Trondheim, Norway, 1998).

3. Results and discussion

3.1. Composition and proteolysis of the standard Manchego cheeses

Table 1 shows the means and standard deviations of the analysed variables (a_w, pH, DM, protein, NaCl,

WSN, PTASN, Tyr, Trp) in the standard Manchego cheeses ($n = 63$). The results of the Student–Newman–Keuls test for comparison of the means are also shown. The a_w values decreased significantly ($P < 0.05$) throughout the ripening period. In contrast, DM values increased progressively ($P < 0.05$) from 15 to 150 days of ripening. The pH increased significantly ($P < 0.05$) between the beginning and the end of ripening. Protein and NaCl contents did not undergo any significant changes, and WSN, PTASN and Tyr contents underwent significant increases over the ripening period studied. Trp contents remained constant during the ripening period.

These results accord with those obtained by other authors for Manchego cheese (Fontecha, Peláez, & Juárez, 1994; Picón, Gaya, Medina, & Nuñez, 1995; García-Ruiz et al., 1998; González-Viñas, Poveda, & Cabezas, 2001; González-Viñas, Poveda, García-Ruiz, & Cabezas, 2001).

Table 1
Mean and standard deviation (S.D.) values ($n = 9$) of the variables analysed in the group of the 63 standard Manchego cheeses, at different ripening times

Variables		Ripening time (days)						
		15	30	45	60	90	120	150
a_w	Mean	0.97f	0.962e	0.958e	0.950d	0.942c	0.929b	0.918a
	S.D.	0.005	0.006	0.010	0.006	0.007	0.004	0.006
pH	Mean	4.95a	4.98a	5.02a,b	5.03a,b	5.06a,b	5.03a,b	5.08b
	S.D.	0.07	0.09	0.08	0.07	0.12	0.09	0.06
DM (%)	Mean	53.9a	56.23b	58.33c	60.45d	63.09e	66.89f	68.92g
	S.D.	1.05	0.72	1.31	1.15	1.24	1.00	0.58
Protein (%DM)	Mean	38.6a	38.3a	38.6a	38.3a	38.5a	38.4a	38.1a
	S.D.	2.40	2.33	1.83	2.17	1.39	1.36	1.95
NaCl (% DM)	Mean	2.52a	2.97a	2.95a	3.24a	3.18a	3.30a	3.26a
	S.D.	0.80	0.59	0.85	0.39	0.46	0.37	0.42
WSN (%TN)	Mean	13.5a	16.6b	18.4b,c	20.7c	23.4d	24.3d	25.6d
	S.D.	3.43	1.40	3.53	2.08	2.38	1.88	3.47
PTASN (%TN)	Mean	0.92a	1.51a	1.81a	2.03a	3.08b	3.14b	3.85b
	S.D.	0.60	0.50	0.44	0.68	1.42	0.87	1.35
Tyr (mg/100 g DM)	Mean	334a	349a	393b	409b	437b,c	439b,c	481c
	S.D.	47.5	52.0	53.1	26.7	15.8	32.6	37.9
Trp (mg/100 g DM)	Mean	120a	153a	189a	176a	195a	170a	164a
	S.D.	38.6	49.2	48.0	44.1	44.9	18.3	23.0

Means within rows without a common letter (a–g) are significantly different ($P < 0.05$) according to the Student–Newman–Keuls test.

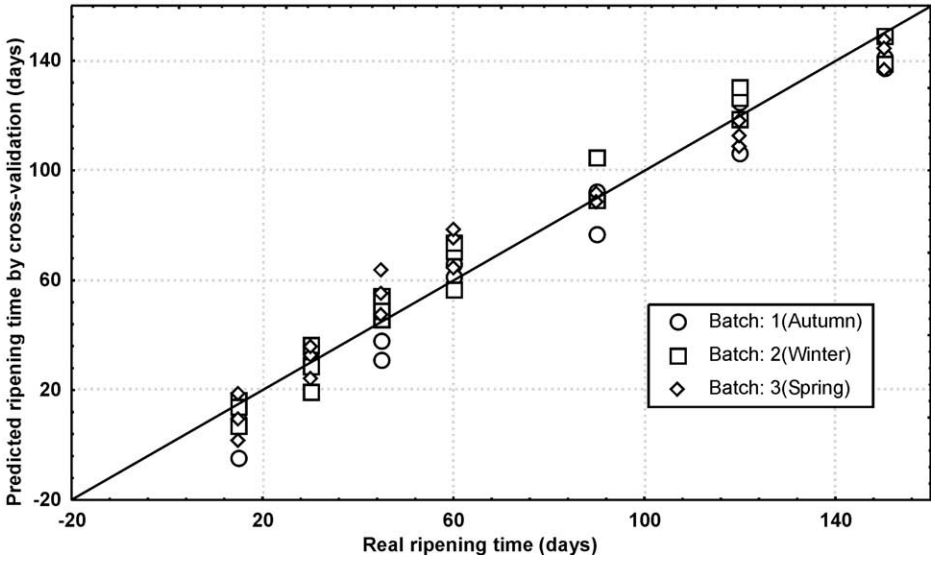

Fig. 1. Correlation between the predicted ripening times ($\hat{t}_{(i)}$), obtained by cross-validation, and the real ripening times (t_i), in the standard cheeses.

3.2. Prediction of the ripening times of commercial cheeses

PLS regression was applied to the 63 calibration samples using the variables showed in Table 1. The regression coefficients obtained for standardised variables showed that the variables that most contributed to predict the ripening time of these commercial cheeses were a_w, pH and DM. Thus, with the aim of reducing the number of variables that take part in the prediction model, only these three parameters, besides their squares and their double products, were considered as predictor variables in the model.

The PLS predictive equation obtained was the following:

$$\begin{aligned}\text{time (days)} = {} & 75.168 - 364.400a_w + 41.902\ \text{pH} \\ & + 1.344\ \text{DM} - 22.275a_w\ \text{pH} \\ & + 1.804a_w\ \text{DM} + 0.252\ \text{pH DM} \\ & - 192.406a_w^2 + 4.179\ \text{pH}^2 + 0.011\ \text{DM}^2,\end{aligned}$$

with two components selected by cross-validation, R^2 (determination coefficient) = 0.969 and s (residual standard deviation) = 7.88 days. The $RMSEP_{cv}$ obtained was 8.48 days. Fig. 1 represents the correlation between the predicted values, obtained by cross-validation ($\hat{t}_{(i)}$), for the ripening times of the calibration samples and the real ripening times of these cheese samples. The prediction of the ripening times of the standard cheeses by the PLS model can be considered good.

This model of prediction was applied to 12 commercial cheeses (evaluation samples), manufactured in the same factory as the standard cheeses and ripened for up to 150 days. Table 2 shows the values for a_w, pH and DM of these commercial cheeses. The values were similar for those obtained for the standard Manchego cheeses. The obtained results for the predicted ripening times, included in Table 2, were good, as shown by the small differences between the predicted ripening times and the real ripening times. The root mean square error of prediction obtained for these 12 samples was 11.9 days.

Table 2
a_w, pH and DM values and mean predicted ripening times for the 12 commercial Manchego cheeses, manufactured in the same factory as the standard cheeses, in the evaluation set

Sample	Real time (days)	a_w	pH	DM (%)	Mean predicted time (days)
1	45	0.951	5.025	56.98	41.1
2		0.968	5.121	56.91	
3		0.955	5.016	55.75	
4	60	0.946	4.970	57.70	53.3
5		0.959	5.002	59.38	
6		0.946	4.966	59.10	
7	90	0.958	5.054	63.01	77.6
8		0.950	5.005	63.67	
9		0.955	4.983	62.57	
10	150	0.910	5.117	67.44	134.0
11		0.914	5.011	66.79	
12		0.937	5.108	68.97	

Table 3
a_w, pH and DM values and mean predicted ripening times for commercial Manchego cheeses manufactured in different factories

Sample	Marked category	a_w	pH	DM (%)	Predicted time (days)
13	Semi-hard (2–4 months)	0.955	4.87	63.26	66.9
14		0.952	5.01	63.36	80.6
15		0.953	5.18	63.31	93.0
16		0.957	5.32	65.88	116.4
17		0.948	5.35	63.50	111.5
18		0.958	5.28	62.58	93.2
19		0.939	5.00	66.45	106.9
20		0.947	4.95	64.07	83.6
21	Hard (>4 months)	0.949	5.40	64.41	120.2
22		0.936	5.52	67.44	157.4
23		0.931	5.07	67.71	125.5
24		0.934	5.10	68.60	131.0
25		0.949	5.25	68.06	129.3
26		0.940	5.29	66.77	131.5

The PLS predictive equation was also applied to predict the ripening time of fourteen commercial Manchego cheeses acquired in different factories. The values for a_w, pH, DM and the predicted ripening times are shown in Table 3. Again, a_w, pH, DM were in accordance to the values obtained for the standard cheeses. All the values of the predicted ripening times fell within the intervals of real ripening times indicated by the cheesemakers.

4. Conclusions

The model of prediction of the ripening time of Manchego cheeses obtained in this study improved the equation proposed by García-Ruiz et al. (1998), since it included a higher number of samples for the calibration (including cheeses manufactured in different seasons) and it also reduced the number of variables that take part in the predictive equation (pH, a_w and DM). These three parameters constitute very simple and fast analysis, so the predicted time of maturation of the cheeses is very easy to calculate.

The predictive model yielded good results for the prediction of the ripening time of Manchego cheeses manufactured in the same factory as the standard cheeses. The model also allowed calculation of the time of maturation of Manchego cheeses manufactured in different factories, which is not specified on the label.

References

AOAC. (1995). *Official methods of analysis* (13th ed.). Washington DC: AOAC.

Bütikofer, V., Rüegg, M., Ardö, Y. (1993). Determination of nitrogen fractions in cheese: evaluation of a collaborative study. *Lebensmittel Wissenschaft Technology*, 20, 271-275.

Casal, V., Martín-Alvarez, P. J., & Herraiz, T. (1996). Comparative prediction of the retention behaviour of small peptides in several RP-HPLC columns by using Partial Least Squares Regression. *Analytical Chimica Acta*, *326*, 77–84.

Fontecha, J., Peláez, C., & Juárez, M. (1994). Biochemical characteristics of a semi-hard ewés milk cheese. *Zeitschrift Lebensmittel-Untersuchung und Forschung*, *198*, 24–28.

Frister, H., Michaelis, M., Schwerdtfeger, T., Folkenberg, D. M., & Sorensen, N. K. (2000). Evaluation of bitterness in Cheddar cheese. *Milchwissenschaft*, *55*(12), 691–695.

García-Ruiz, A., Cabezas, L., Martín-Alvarez, P. J., & Cabezudo, M. D. (1998). Prediction of the ripening times of Manchego cheese using multivariate statistical analyse: a prelyminary study. *Zeitschift Lebensmittel Untersuchung Forschung*, *206*, 382–386.

González-Viñas, M. A., Poveda, J. M., & Cabezas, L. (2001). Sensory and chemical evaluation of Manchego cheese and other cheese varieties available in the Spanish market. *Journal of Food Quality*, *24*(2), 157–165.

González-Viñas, M. A., Poveda, J. M., García-Ruiz, A., & Cabezas, L. (2001). Changes in chemical, sensory and rheological characteristics of Manchego cheeses during ripening. *Journal of Sensory Studies*, *16*(4), 361–371.

Kuchroo, C. N., & Fox, P. F. (1982). Soluble nitrogen in Cheddar cheeses: comparison of extraction procedures. *Milchwissenschaft*, *37*, 331–335.

Martens, H., & Naes, T. (1993). *Multivariate calibration*. Chichester: John Wiley & Sons.

Molina, E., & Martín-Alvarez, P. J. (1996). Comparación de diferentes técnicas de regresión para la predicción de los porcentajes de leche de vaca, oveja y cabra en quesos de mezcla. *Revista Española de Lechería*, *73*(2), 33–43.

Molina, E., Martín-Alvarez, P. J., & Ramos, M. (1999). Analysis of cows, ewés and goats milk mixtures by capillary electrophoresis: quantification by multivariate regression analysis. *International Dairy Journal*, *9*, 99–105.

Pérez-Coello, M. S., Martín-Alvarez, P. J., & Cabezudo, M. D. (1999). Prediction of the storage time in bottles of Spanish white wines using multivariate statistical analysis. *Zeitschrift Lebensmittel-Untersuchung und Forschung*, *208*, 408–412.

Picón, A., Gaya, M., Medina, M., & Núñez, M. (1995). The effect of liposome-encapsulated *Bacillus subtilis* neutral proteinase on Manchego cheese ripening. *Journal of Dairy Science*, *78*, 1238–1247.

Pripp, A. H., Shakeel-Ur-Rehman, McSweeney, P. L. H., & Fox, P. F. (1999). Multivariate statistical analysis of peptide profiles and free amino acids to evaluate effects of single-strain starters on proteolysis in miniature Cheddar-type cheeses. *International Dairy Journal*, *9*, 473–479.

Pripp, AH, Shakeel-Ur-Rehman, McSweeney, P. L. H., Sorhaug, T., & Fox, P. F. (2000). Comparative study by multivariate statistical analysis of proteolysis in a sodium caseinate solution under cheese-like conditions caused by strains of Lactoccoccus. *International Dairy Journal*, *10*(1/2), 25–31.

Vakaleris, B., & Price, W. (1959). A rapid spectrophotometric method for measuring cheese ripening. *Journal of Dairy Science*, *42*, 264–276.

Virgili, R., Parolari, G., Bolozoni, L., Barbieri, G., Mangia, A., Careri, M., Spagnoli, S., Panari, G., & Zannoni, M. (1994). Sensory–chemical relationships in Parmigiano-Reggiano cheese. *Lebensmittel Wissenschaft Technology*, *27*(5), 491–495.