Application of solvent retention capacity tests for prediction of rheological parameters of wheat flour mill streams

Milan S. Vukić[1,2], Elizabet P. Janić Hajnal[3], Jasna S. Mastilović[3], Dragan P. Vujadinović[2], Marko M. Ivanović[2] and Dragana M. Šoronja-Simović[1]

[1]*University of Novi Sad, Faculty of Technology, Novi Sad, Serbia*
[2]*University of East Sarajevo, Faculty of Technology Zvornik, Zvornik, Republic of Srpska, Bosnia and Herzegovina*
[3]*University of Novi Sad, Institute for Food Technology, Novi Sad, Serbia*

Abstract

This paper presents relationship between the rheological properties of dough and individual polymer swelling properties in wheat flour mill streams. The swelling properties were measured by applying the Solvent Retention Capacities (SRC) tests. Significant correlation coefficients were determined for certain rheological parameters. In an effort to extract additional insights from the properties measured, a multivariate analysis was used to develop relationships between the studied parameters. To determine relevant relationships among the parameters, the data exploration step by the Principal Component Analysis was performed. Then, multivariate Partial Least Squares Regression (PLSR) models were developed, to predict certain empirical rheology parameters based on the SRC parameters. The processing of experimental data indicated the possibility of using SRC parameters for predicting rheological properties in conjunction with a suitable mathematical model. The presented approach may be useful for rapid prediction of wheat flour mill streams characteristics and for optimization of the end-flour performances.

Keywords: partial least squares regression; modelling; polymer swelling; rheology

Available on-line at the Journal web address: http://www.ache.org.rs/HI/

SCIENTIFIC PAPER

UDC: 664.641.12+ 664.71.05: 664.653: 665.7.035.6

Hem. Ind. **74 (1)** 37-49 (2020)

1. INTRODUCTION

Wheat (*Triticum aestivum* L.) is one of the most important crops which is used in a range of food products. Processing of wheat to wheat flour involves grain milling, either fully or partially, by separating the bran and germ from the endosperm. This is achieved by a series of size-reduction operations producing wheat flour mill streams. Wide range of wheat flours are produced that results from different combinations of wheat flour mill streams. Not every wheat flour mill stream is equally suitable for producing specialty wheat flour. The variations in characteristics of wheat flour mill streams results in complexity of optimization of flour production. Irrespective of the end-use, it is necessary to maintain the highest accuracy in determination or prediction of wheat flour mill streams quality. Hence, an optimal merge of wheat flour mill streams into the desired end-flour is of great importance for subsequent baking processes [1]. Accordingly, millers and bakers have to agree concerning the methodology used for quality characterization of flour. A multitude of analyses is available, such as physico-chemical analyses, rheological tests, and baking tests [2]. Tests requiring small amounts of the sample and short time, and easily performed in daily production, would be preferable. However, technological properties of flour are the result of complex interactions between all constituting polymers and are not only related to the protein and gluten contents. It is known that the protein quality, as well as that of other flour polymers, influences the bread-making quality of wheat flour and they should be taken into consideration when characterizing wheat flour mill streams [3]. To this day, flour classification is the most commonly performed using

Corresponding author: Milan S. Vukić, University of Novi Sad, Faculty of Technology; University of East Sarajevo, Faculty of Technology Zvornik, *Republic of Srpska, Bosnia and Herzegovina*
E-mail: milan.vukic@tfzv.ues.rs.ba
Paper received: 25 June 2019
Paper accepted: 06 January 2020
https://doi.org/10.2298/HEMIND190625001V

several parameters, such as the flour protein content or rheological behaviour, which represent quantitative determination of dough mechanical properties [4]. Predominantly, the milling industry relays on empirical rheological tests as these are still considered as the most accurate way to assess the flour quality [5]. Accordingly, to produce high-quality products, tailored to specific requirements, the dough must have the “optimum rheology” for the specific purpose. Empirical rheological tests are purely descriptive, and the judgment is made taking into consideration several important factors and their interwoven interactions in each specific case. Most of the empirical rheology tests exhibit the following disadvantages: results are dependent on imposed testing conditions (amount of the analysed flour, geometry of a mixer unit, operating parameters of the device, *etc.*) as defined by the specific equipment used, so that experienced staff and considerable time are required. Moreover, empirical rheological tests are poorly suited for quick routine analyses. Among the most widely accepted empirical rheology techniques are based on the use of farinographs and extensographs. Their application for quality evaluation of wheat flour is promoted by the existence of high correlations of the parameters obtained from these techniques and indices of the quality of the end-use food products [6]. Another method for predicting the functionality of wheat flour is the Solvent Retention Capacity (SRC) test, which is increasingly used by wheat breeders, millers, and bakers. The Solvent Retention Capacity test methodology is based on quantifying the enhanced swelling behaviour of flour polymer networks in diagnostic solvents [7]. Each flour polymer network is associated with the corresponding diagnostic solvent so that damaged starch is associated with the sodium carbonate SRC (SRCSo), flour arabinoxylan with the sucrose SRC (SRCSu), glutenin characteristics and gluten strength with the lactic acid SRC (SRCLa) while the water retention capacity SRC (SRCw) is an indicator of the overall water holding capacity of all polymeric constituents [1,7]. SRC tests produce a practical functionality profile of wheat flour, which was first utilized for prediction of soft wheat flour baking characteristics [2]. Most of the published papers consider only wheat breeding aspects. However, having in mind first principles on which the SRC methodology is based, there are no obstacles for testing properties of wheat flour mill streams as other authors proved [8]. Besides abovementioned, there is a global trend toward time reduction and introduction of data analyses, modelling and automation of processing, from which the milling business is not excluded. The challenging situation is how to predict properties of wheat flour mill streams by using a relatively fast and simple method instead of slow, traditional empirical rheology techniques.

Wheat flour mill streams, commercially milled, were considered in the present study and characterized by SRC tests, followed by farinograph and extensograph rheological analysis methods. Firstly, a SRC test was used to evaluate its ability to differentiate between the samples. Subsequently, we have investigated the possibility for developing regression models to predict wheat flour mill streams quality based on the results of SRC tests and to obtain relationships between SRC parameters and farinograph and extensograph rheological parameters. Up to the authors' knowledge, studies considering application of regression modelling of SRC test results to assess the wheat flour mill streams rheological quality do not exist in literature. Predictor variables included SRC parameters as the authors hypothesized that these indices indicate the flour mill streams features and would allow forecasting the outcome of empirical rheology measurements with a sufficient degree of accuracy. Partial least-squares regression (PLSR) was chosen as a modelling technique. PLSR demonstrated the utility to analyse data with the so-called large p small n problem that is, many variables and few samples [9]. As a supervised method, PLSR is specifically suited to overcome noisy, collinear, and even incomplete variables and produce good predictions in multivariate problems. This resilience allows PLSR to be utilized in situations where the use of conventional methods is particularly limited.

This study was performed aiming to: i) correlate SRC parameters with empirical rheology parameters for all investigated wheat flour mill streams, and ii) explore the applicability of SRC tests and PLSR modelling to predict rheological parameters of wheat flour mill streams.

2. MATERIALS AND METHODS

For SRC tests, deionized water and solutions of sucrose (50 %), sodium carbonate (5 %) and lactic acid (5 %) in deionized water were used, expressed as weight concentration (w/w). All chemicals and solvents used were of at least ACS grade (Sigma - Aldrich, St. Louis, MO).

2. 1. Milling of wheat

The cleaned and conditioned wheat, procured on the local market, was milled in a commercial-scale plant (Molaris d.o.o., Republic of Srpska, Bosnia and Herzegovina). The commercial-scale plant consists of five break rolls (B1 to B5), six reduction rolls (C1 to C6), one purifier rolls (D1) and a bran finisher. The break, reduction and purifying rolls adjustments were set as in regular commercial wheat milling operations. After each grinding passage, the mixture of endosperm, bran and germ, in released middling, was purified in the purifier. In total, nineteen wheat flour mill stream samples were collected from all break, reductions and purifier passages. The simplified mill flow diagram is shown in Figure. 1.

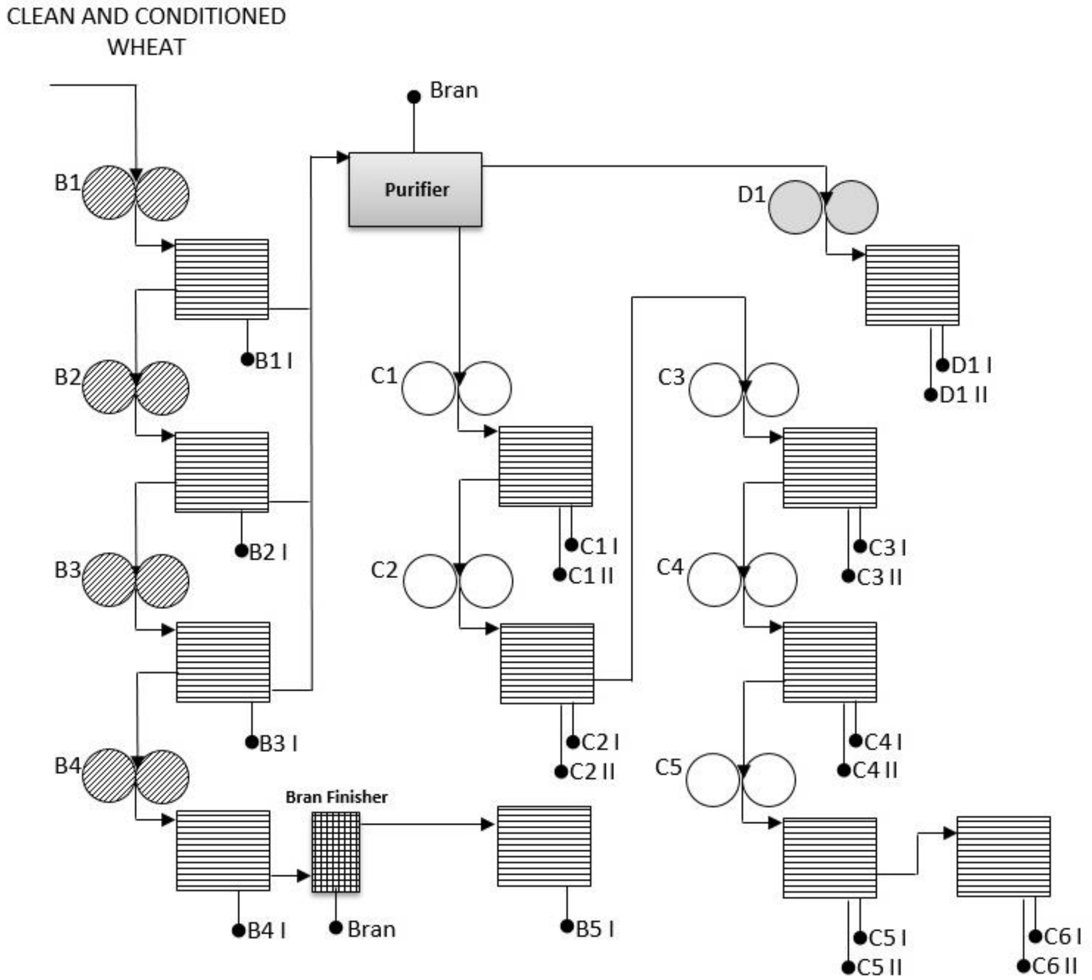

Figure 1. The simplified mill flow diagram. B1 I - B5 I: break flour streams; C1 I, C1II, C2 I, C2II, C3 I, C3II, C4 I, C4II, C5 I, C5II, C6 I, C6II: reduction flour streams; D1 I – D1II: purifying flour streams

2. 2. Rheological dough methods

Farinograph rheological measurements Farinograph®-E, Brabender GmbH & Co. KG, Germany) were carried out as specified by the ICC 115/1 method [10]. Extensograph rheological measurements (Extensograph®-E, Brabender GmbH & Co. KG, Germany) were carried out as specified by the ICC 114/1 method [11].

2. 3. Solvent retention capacity tests

Solvent retention capacities (SRC) of the investigated samples were determined according to the modified AACC Standard Method 56-11 [12]. The modification refers to the reduced mass of the sample (1 g instead of 5 g) as previously proposed by Bettge *et al.* [13]. SRC is the weight of solvent held by flour after centrifugation (SRCw - water retention capacity; SRCSo - sodium carbonate solvent retention capacity; SRCLa - lactic acid solvent retention capacity; SRCSu – sucrose solvent retention capacity). Results (n = 3) are expressed as a percent of the flour weight, on the 14 % moisture basis. Values were calculated by using the equation:

$$\text{SRC, \%} = \left(\frac{\text{gel weight}}{\text{flour weight}}\right)\left(\frac{86}{100 - \text{flour moisture}}\right)100 \tag{1}$$

where the gel weight is the weight of wet pellet calculated by subtracting the weight of tube and cap from the total weight of the tube, cap, and gel after sample centrifugation, and supernatant decantation. ...

A derived SRC parameter, the glutenin performance index (GPI), was calculated by the equation [14]:

$$\text{GPI} = \frac{\text{SRCLa}}{\left(\text{SRCSo+SRCSu}\right)} \tag{2}$$

2. 4. Data analysis

Significant correlations between the measured parameters were analysed adopting the Pearson correlation analysis procedure. All variables were centred and scaled to unit variance prior to the multivariate analyses. Principal Component Analysis (PCA) was used as an unsupervised explorative technique to represent the variation present in the dataset and reduce the dimensionality using a small number of Principal Components (PC) [15]. These PCs are independent variables and explain variability of the data in a decreasing order. The PCA analysis was conducted by using R packages "FactoMineR" and "factoextra" [16,17]. Statistically significant differences among the data in this study were analysed by one-way ANOVA followed by the Tukey's test and a p-value < 0.05 was considered statistically significant. All data analyses were performed in R for Windows, an open-source language and environment for statistical computing R-3.3.2. [18].

2. 5. PLSR modelling

All PLSR models were formulated using five predictor variables: SRCSo, SRCSu, SRCLa, SRCw and GPI. The PLSR method was chosen as an alternative method to the ordinary regression, which is impaired by limitations of the sample size or by highly co-linear predictor variables [19]. This regression method avoids inflation of errors in such circumstances which hinder classical multiple regression analyses and is based on projection of the predictor and response variables on latent structures or latent variables and corresponding scores. By minimizing the dimensionality present in the data and projecting the predicted variables and the observable variables to a new space, the method finds a linear regression model within that smaller space. In contrast to the principal component analysis, in which dimension reduction ignores the predicted variable, the PLSR procedure aims to maximally explain the predicted variable. PLSR produces a linear model, in the form of a general equation, presented below (eq. 3), where Y is an n cases by m variables response matrix, X is an n cases by p variables predictor (design) matrix, B is a p by m regression coefficient matrix, and Є is a noise term for the model which has the same dimensions as Y. Detailed mathematical descriptions can be found in literature [18,20].

$$Y = XB + \epsilon \tag{3}$$

Before implementing modelling in R with the package "pls", we min-max normalized the independent variables space to give each variable the same importance in the analysis [21]. Separate PLSR models were fitted by using a kernel algorithm for every rheological parameter to identify the main parameters suitable for multiple responses modelling. To cope with the problem of overfitting, cross-validation was used. In this way, an optimal balance between explained variation in the response, and the predictive ability of the model was achieved. To validate the predictive ability of resulting models the Leave-one-out cross-validation (LOOCV) was used [22,23]. Primary fitted models were rebuilt with the optimal number of LV (Latent variables). The "best.dims" function of the R package was used to determine how many LVs are needed to find an adequate PLSR model by minimizing the root mean square error of prediction (RMSEP). Multiple responses PLSR modelling was carried out with the above-described procedure. For the response variables, the farinograph water absorption (FWA), energy (E), extensibility (Ex) and the maximum of resistance (R_{max}) were chosen. Then, the standardized regression coefficients were normalized so that their absolute sum equals 100 and the results are sorted.

2. 5. 1. Accuracy of the models

Performance verification of all developed models was performed by using the coefficient of determination (r^2) referred to the calibration of the training set, coefficient of prediction (r^2pred) referred to the cross-validation results, root mean square error of prediction (RMSEP) and the cross-validated prediction (*i.e.* the PRESS), which reports the error of the PLS model in the same units as the Y variable. These commonly used parameters are described in detail in literature [21].

3. RESULTS AND DISCUSSION

3. 1. Farinograph, extensograph and solvent retention capacity tests

Rheological properties of wheat flour mill streams are presented in Supplementary material. The obtained samples demonstrated a wide range of rheological properties. According to the parameters, samples used in this study had varying rheological properties and bread making potentials as it was expected, and which is desirable for modelling.

SRC parameters provide a measure of solvent compatibility for the three polymeric components of wheat flour: gluten, damaged starch, and arabinoxylan, *i.e.* SRCLa, SRCSo and SRCSu [8]. Flour for bread production typically requires considerable farinograph water absorption (FWA), high energy (E), and relatively high damaged starch and arabinoxylan contents. Flour for cookie production typically necessitates low FWA, minimal E and low damaged starch and arabinoxylan contents [8]. Large variations in SRC values were found among all studied what flour mill streams. SRCSo varied from 70.28 to 99.77 % (av. 83.17 %), SRCSu from 93.15 to 120.17 % (av. 104.07 %), SRCLa from 104.3 to 146.01 % (av. 134.75 %) and SRCw from 60.79 to 77.79 % (av. 68.89 %). GPI varied from 0.6 to 0.83 with the average of 0.72. The SRC values presented in Figures 2, and 3, showed that the break flour streams had the lowest values of SRCSo, SRCSu, SRCw, and GPI, while the last reduction flour streams had the highest values of SRCSu, SRCw and GPI. On the other hand, the highest SRCLa value was found in the fourth break stream, while the lowest estimates of SRCLa and SRCSo were found for the fifth break stream. Data clearly indicated the distribution in wheat flour mill streams quality as expected and the average of all SRC values were found in the reduction flour streams. The SRCLa values are related to the amount of protein in a flour sample [1,8]. A more recent study observed that the SRCLa is related specifically to the glutenin content, not to the total protein content [23]. Hence, the mill stream samples with higher SRCLa values specifically indicate higher glutenin content. The SRCSo values are related to the damaged starch content of the flour while the SRCSu values refer to the arabinoxylan content of the flour sample [8]. The break flour streams had lower SRCSu values than the other samples and this is probably due to the low arabinoxylan content. Break rolls reduce endosperm farthest from aleurone and cause lower arabinoxylan content. Compared to break flour streams, SRCSu values gradually increased in reduction flour streams, while there was a sharp decrease in purifying flour streams. Water is used as a control solvent, hence, SRCw is not specifically connected to a certain polymer. SRCw values describe the overall water holding capacity and are lower compared to the other SRC values. SRCw is modulated by all polymers in flour, which slightly increase its water absorption [8]. Similarly, as with the SRCSu values, SRCw values were lower in break flour streams, then values gradually increased in reduction flour streams till purified flours that showed values similar to those of the break flour streams. It has been previously documented that other flour polymers such as arabinoxylan and damaged starch influence the SRCLa value [1,25]. Glutenin performance index (GPI value), sometimes referred to as corrected SRCLa value, is obtained by the Eq. (2), as described in literature [26]. The break flour streams had the highest GPI, while the last reduction flour streams had the lowest GPI values. In another study, the GPI values ranged between 0.63 and 0.85, which is similar to the GPI values determined in this study, although it was expected that the GPI values would be higher due to the modification of the method used [27]. From Figure 2 a steady increase in SRCSo values, with the number of passes, can be observed. The last wheat flour mill streams showed higher SRCSo values than the first ones. As SRCSo values relate to the damaged starch content, it can be concluded that the content of this component is increasing as well. These results indicate that the SRC values had a strong diagnostic potential for wheat flour mill streams quality due to different distribution of grains constituents in different mill streams and can be used to differentiate between individual wheat flour mill streams.

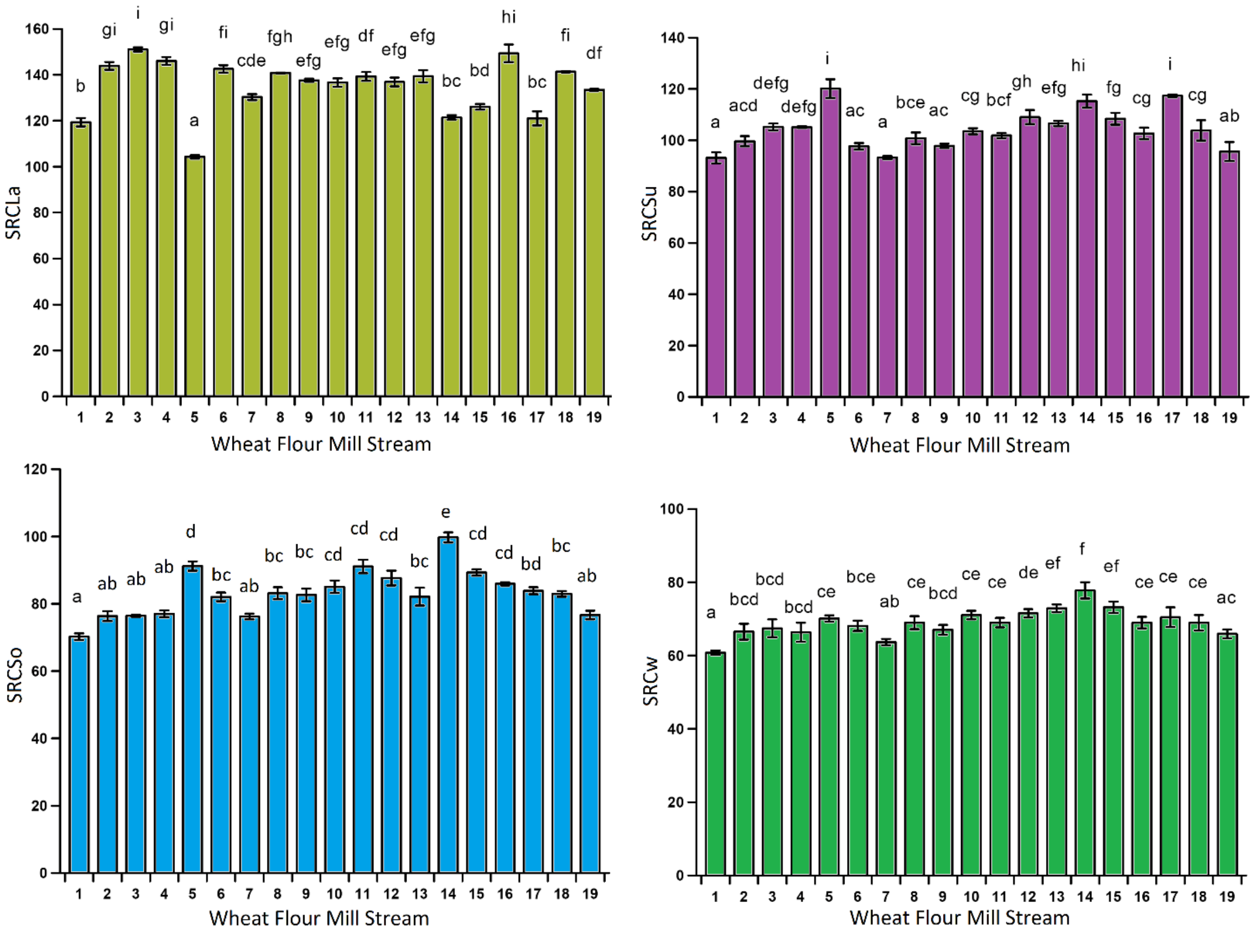

Figure 2. Solvent retention capacities of wheat flour mill streams; 1-5: break flour streams; 6-17: reduction flour streams; 18-19: purifying flour streams. Different letters of bars denote significant differences (Tukey's test p>.05)

Actually, the SRC test presented in Figure 2 reveals that the properties of wheat flour mill streams could be differentiated with statistically significant differences ($p < 0.05$). Presented differentiation could be expected in line with prior rheology knowledge. Therefore, the importance of this rapid test in differentiating functionality profiles of wheat flour mill streams was confirmed.

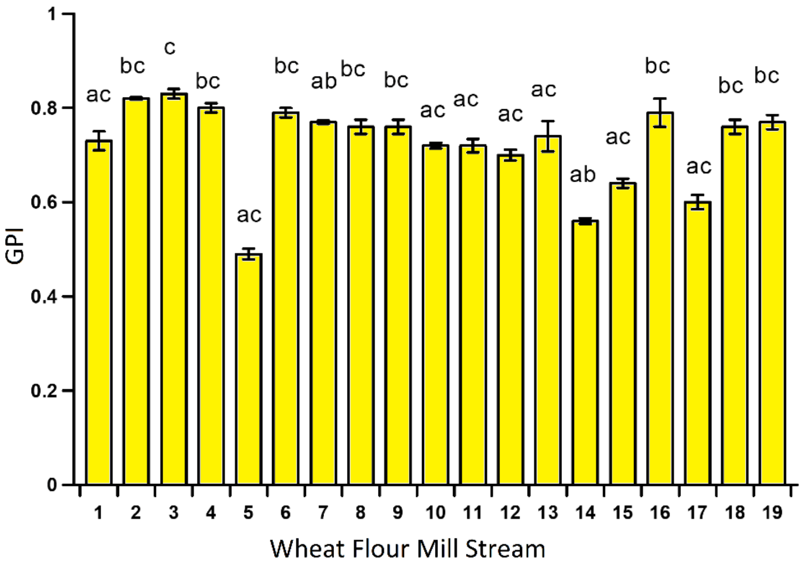

Figure 3. Glutenin performance index of wheat flour mill streams; 1-5: break flour streams; 6-17: reduction flour streams; 18-19: purifying flour streams. Different letters of bars denote significant differences (Tukey's test p>.05)

3. 2. Relationships between SRC parameters and rheological parameters

To acquire a unifying view of the relationship between the SRC and rheological parameters, a multivariate technique PCA was employed and Pearson's correlation coefficients between different parameters were calculated. PCA allowed us to summarize the systematic patterns of variations in the data and to reduce a "complex" data set to a lower dimension. PCA revealed dominant types of variations in both the observations and the variables. The importance of the first 10 principal components (PCs) is visualized in a scree plot (Fig. 4). It shows that 74.2 % of the information (variances) incorporated in the data is explained by the first two principal components. Therefore, influence of other principal components is not expected to be considerable. In this way, the raw data set was reduced to a lower dimension to reveal the structures or dominant types of variations in both the SRC values and the rheological parameters.

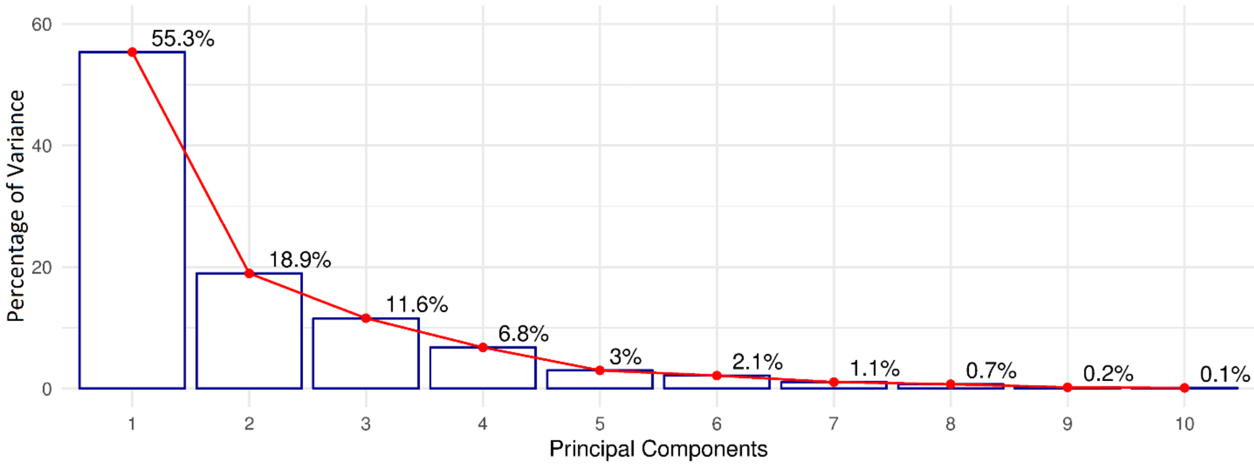

Figure 4. Principal component analysis (PCA) scree plot

The first PC, PC1, explained 55.3 % of the present variance while the second PC, PC2, explained 18.9 % of variance present in data. To explore relationships between all variables, and to elucidate the underlying nature of a particular PC, the original 13 variables are projected onto a 2-dimensional circle of correlation. The correlations between each variable and the respective PC are used as coordinates of the variable for projection. With the information obtained, we can interpret the "key variables" behind the PCs. The squared correlation coefficients for variables are called squared cosine ($\cos^2$) [28]. Analogous to the Pearson's r, the squared cosine is the percent of the variance in that variable accounted by the PC and is used to estimate the quality of the variable representation. In the case when a variable is perfectly represented by only two components, the sum of the $\cos^2$ is equal to one and the variable will be positioned on the circle of correlations. Figure 5 shows the circle of correlation on the PCs. Variables are coloured according to the values of the squared cosine. The percent of the variance in red coloured variables (R, R_{max}, GPI, SRCSo, SRCw, FWA and *Ex*) accounted by the PCs is in the range between 0.75 and 1 while violet coloured variables variance (SRCLa, E, SRCSu, FDT and FDS) is accounted in the range between 0.50 and 0.75. Green coloured FDS is unsatisfactorily explained with PCs. PC1 was strongly positively determined by the GPI, FST, R_{max}, SRCLa, E, FDT, and R while it was strongly negatively determined by SRCw, SRCSo, SRCSu and FWA.

The second PC, PC2, was strongly positively determined by the values of SRCw, SRCSo, and E, while it was negatively determined by the Ex value. From Figure 5, two groups of variables can be distinguished: variables sensitive for dough strength (right side) and variables that are more sensitive for water absorption (left side). Positively correlated variables *e.g.* FWA and SRCSu on the left side and SRCLa and Rmax on the right side are grouped together and pointing in the same direction. Orthogonal variables are unrelated, and variables positioned on opposed quadrants *e.g.* SRCSu and Rmax of the plot origin are negatively correlated. The distance between a variable and the plot origin measures the quality of the representation of the variable. From Figure 5, it can be seen that 12 of 13 variables are away from the origin and thus are satisfactory represented, all except the FDS. This means that for FDS, more than 2 components are required to satisfactorily represent the data and as such it is less important for the first two PCs.

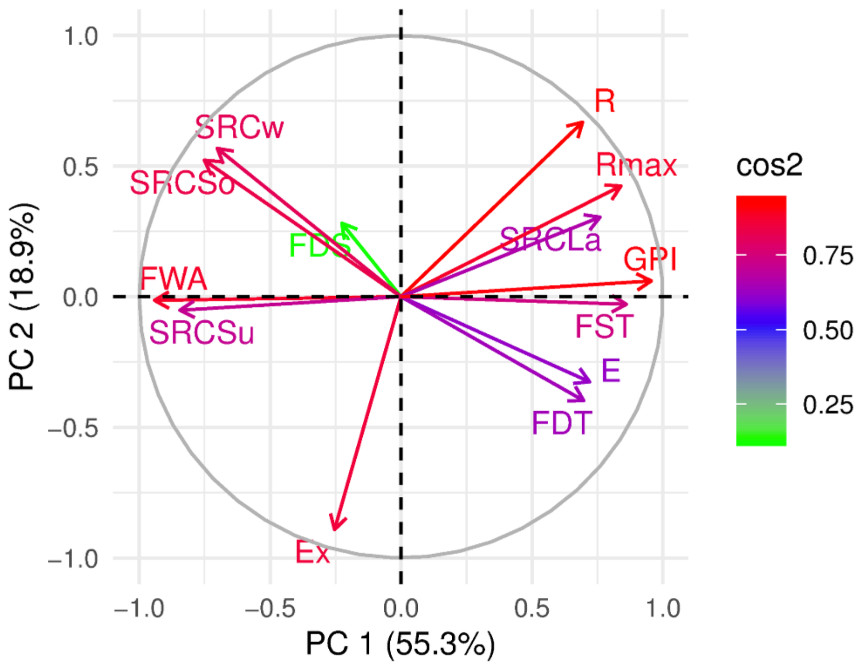

Figure 5. Principal component analysis (PCA) of measured variables, circle of correlation: cos^2, cosine squared; FDS, degree of softening; FDT, dough development time; FST, dough stability; E*, energy;* Ex*, extensibility; FWA, farinograph water absorption; GPI, glutenin performance index;* R$_{max}$*, maximum of resistance; PC, principle component;* R*, resistance at 5 cm; SRCLa, lactic acid solvent retention capacity SRCSo, sodium carbonate solvent retention capacity; SRCSu, sucrose solvent retention capacity; SRCw, water retention capacity.*

Along with the PCA analysis, correlation analysis was conducted relating SRC values, and farinograph and exstensograph parameters. Pearson's correlation coefficients between the different parameters are presented in Table 1.

Table 1. Correlation coefficients of SRC values and conventional rheological parameters

Parameters	SRCLa	SRCSu	SRCSo	SRCw	GPI
FWA, %	-0.57 **	0.93 ***	0.76 ***	0.76 ***	-0.86 ***
FDT, min	0.51 *	-0.46 *	-0.65 **	-0.60 **	0.66 **
FST, min	0.64 **	-0.68 **	-0.55 *	-0.58 **	0.76 ***
FDS, BU	0.04	0.26	0.26	0.4	-0.09
E / cm^2	0.68 **	-0.31	-0.63 **	-0.52 *	0.73 ***
Ex / min	-0.29	0.42	-0.2	-0.21	-0.25
R / BU	0.65 **	-0.58 **	-0.17	-0.13	0.64 **
R_{max} / BU	0.79 ***	-0.61 **	-0.39	-0.32	0.81 ***

*Correlation is significant at the levels * p < .05, ** p < .01, *** p < .001. FDS, degree of softening; FDT, dough development time; FST, dough stability;* E*, dough energy;* Ex*, extensibility; FWA, farinograph water absorption; Rmax, maximum of resistance;* R*, resistance at 5 cm; BU, Brabender unit.*

For the flour samples, a strong negative linear relation was observed between the flour GPI and FWA ($r = -0.86$, $p < 0.001$), and intermediate between FWA and SRCLa ($r = -0.57$, $p < 0.01$). Reported relationships are in accordance with the findings reported in a similar study [8]. Positive linear relations were found between GPI and Rmax ($r = 0.81$, $p < 0.001$), between GPI and E ($r = 0.73$, $p < 0.001$), and between GPI and FST ($r = 0.76$, $p < 0.001$). The parameter FWA exhibited positive linear relations with SRC values. In specific, FWA was positively related with the SRCSu ($r = 0.93$, $p < 0.001$) as well as with the SRCSo and SRCw ($r = 0.76$, $p < 0.001$). These results are in line with those of earlier studies [28]. Also, the linear positive relation of E values with the SRCLa ($r = 0.68$, $p < 0.01$) was expected because both of these parameters are influenced by the protein properties in the flour. A similar relation was earlier observed between the SRCLa values and dough strength [8,30]. SRCw, SRCSo and SRCSu were grouped close to each other in

Figure 5, indicating a positive correlation between these three SRC variables. These results are in line with those of earlier studies that also observed strong correlations between those three SRC values [31 - 33]. The reason for abovementioned correlations is that all these variables indicate the flour water holding capacity. SRCLa was orthogonal to the other SRC values, displaying weaker correlations, except to the GPI value. This behaviour is to be expected as the GPI value is directly derived from the SRCLa value.

3. 2. PLSR modelling to predict rheological parameters of wheat flour mill streams

In the PLSR modelling, dimension of the data matrix was reduced to a small set of informative super axes, called latent variables (LVs). Then PLSR was used to predict rheological parameters from the SRC values. Partial least squares regression models fitted by using a kernel algorithm and cross-validated were developed for all rheological parameters. PLSR allowed modelling while dealing with multicollinearity and a limited number of samples [19]. In Table 2, summary of PLSR models fit statistics are presented.

Table 2. Summary of PLSR models fit statistics

Parameters	LV	r^2	r^2pred	RMSEP	PRESS
FWA, %	2	0.93	0.89	0.096	0.177
FDT, min	2	0.60	0.10	0.312	1.850
FST, min	5	0.85	0.60	0.192	0.703
FDS, BU	5	0.71	0.37	0.198	0.747
E / cm^2	4	0.92	0.88	0.078	0.116
Ex / min	4	0.81	0.70	0.176	0.588
R / BU	2	0.58	0.39	0.223	0.291
R_{max} / BU	2	0.70	0.55	0.196	0.736
FWA + E + Ex + R_{max}	4	0.89	0.85	0.184	0.648

FDS, degree of softening; FDT, dough development time; FST, dough stability; E*, dough energy;* Ex*, extensibility; FWA, farinograph water absorption; Rmax, maximum of resistance;* R*, resistance at 5 cm; BU, Brabender unit; FWA +* E *+* Ex *+* R_{max}*, model with multiple responses; LV, Latent variables;* r^2*, coefficient of determination;* r^2*pred coefficient of prediction; RMSEP, root mean square error of prediction; PRESS, cross-validated prediction.*

Summary of the eight PLSR models constructed separately for prediction of each rheological parameter separately and for simultaneous prediction of multiple responses (FWA + E + Ex + R_{max}) are as indicated in Table 2. The optimum model performances for FWA, FST, R and R_{max} were obtained using two LVs, whilst the remaining rheological parameters were better modelled by using four or five LVs. Regression models were recognized based on the minimum error in the root mean square error of prediction (RMSEP) by the number of LVs that provided the best strength of models (r^2), models ability to predict new samples (r^2pred) while minimizing cross-validated predictions errors (PRESS). On the whole, considering the r^2pred values, acceptable to good results (r^2pred ≥ 0.6) were obtained for 5 out of the 8 considered parameters. However, for the FDT, FDS and Rmax parameters, the models, were not satisfactory, as signified by the r^2pred values, indicating poor model stability, or inability to identify any causality related to these rheological parameters from the SRC values. For FST and Ex values, the performance of developed models can be categorized as good. All developed models have RMSEP and PRESS error values, which can be considered satisfactory. The optimal number of LVs for the PLSR of FWA and E models is 2 and 4 LVs, respectively. As expected, these results indicate that SRC analyses can be used to predict these two rheological parameters with a very satisfactory predictive power (r^2pred = 0.89 and r^2pred = 0.88, respectively). Additionally, PLSR allows modelling of multiple responses, simultaneously. Among the examined parameters, four parameters that provide the most useful information for comparison of different samples were chosen for model development. Moreover, their fit statistics implied that it was reasonable to employ them. The SRC readings, namely SRCLa, SRCSu, SRCSo and SRCw, together with the derived value of GPI were used as X-variables and the FWA, E, Ex, and R_{max} as multiple responses variables (Y-variables). The overall performance of the prediction (r^2pred = 0.85) and standardized regression coefficients are presented in Figure 6. The values of the RMSEP and PRESS, obtained by the PLSR model, are comparable to the values of the single response models.

One of the attractive features of developed PLSR models is that the relationships between predictors variables (in our case the SRC values) and the response variables (rheological parameters) can be induced from standardized regression coefficients of predictor variables in the most explanatory LVs. In this way, standardized regression coefficients provide the direction (one-dimensional) of the influence of predictor variables. In the multi-response PLSR model, standardized regression coefficients were utilized to explore the influence of relevant predictor's variables on the model. In Figure 6, the regression coefficients are normalized so that their absolute sum is 100 and the results are sorted according to the sign.

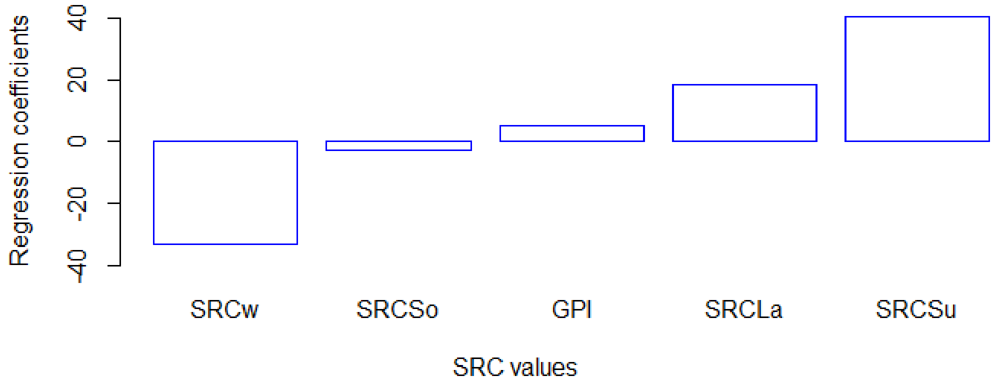

Figure 6. The standardized regression coefficients of predictive variables for the multi-response PLSR model

Standardized coefficients of SRC values in the PLSR model signify the mean change of dependent variables given a one standard deviation shift in the independent variable. From these values, we can relate the contribution of each variable to the regression model. It is obvious that variables SRCw, SRCLa and SRCSu provide significant contributions to the predictive power of the model, while variables SRCSo and GPI contribute to a much lower extent. An explanation for the observed significance of variables can be found in the nature of samples and in their properties as SRC test allows the separation of effects of different flour components. As shown in Figure 6 variations in SRCSo values had a low influence on the PLSR model. The importance of SRCw and SRCSu values can be likely related to the fact that both indices are strongly associated with the water holding capacity, as one of the most important properties of flour. The SRCw value provides the total water holding capacity across all flour polymers and is therefore modulated by all of them, while the SRCSu value is specifically related to arabinoxylans [14]. Different directions of SRCw and SRCSu values could be attributed to different examination orientations of these values. Directions of GPI and SRCLa values can be likely related to the fact that both the indices are strongly related to the protein content in flour and strongly determinate the dough strength. The dough strength, together with the water holding capacity defines the dough rheological properties [34,35].

Finally, when standardized regression coefficients of models for predicting E and FWA individually are presented (Figures 7, and 8), the contribution of each variable to the regression models can be better understood.

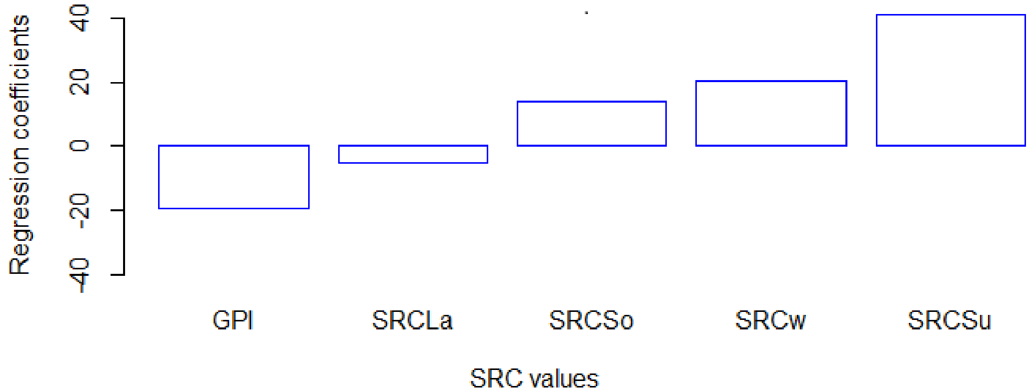

Figure 7. The standardized regression coefficients of predictive variables for the FWA predictive PLSR model

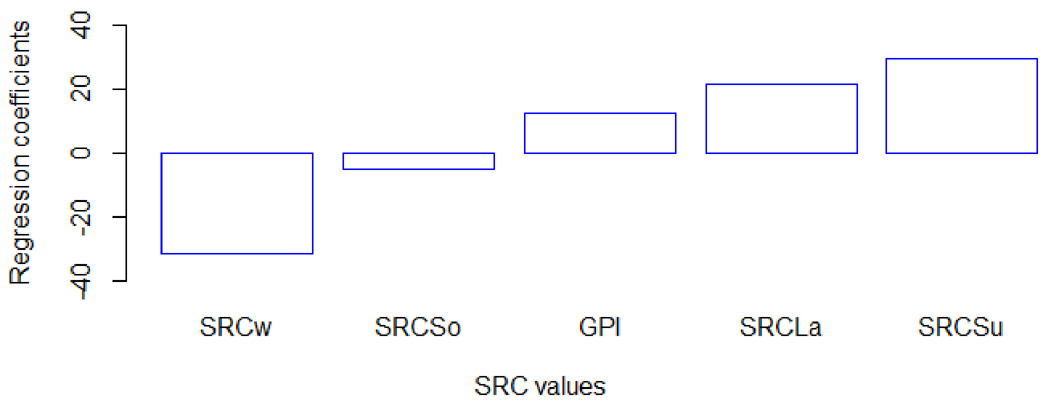

Figure 8. The standardized regression coefficients of predictive variables for the E predictive PLSR model

In the case of the FWA parameter, the same direction of SRCSo, SRCw, and SRCSu, on one side, and of SRCLa and GPI on the other, define the direction of the contribution of each variable to the model behaviour in accordance with their explanatory orientations. The contribution of each variable is as expected; the SRCSu value has the highest contribution, given that arabinoxylans molecules exert the largest water holding capacity of all wheat flour polymers. Likewise, directions of the contribution of each variable on predicting the E parameter can be explained. Here, the SRCSu value shares the direction with GPI and SRCLa values, which is in line with results of other authors that have shown that the arabinoxylans influence on dough strength should not be underestimated. Soluble arabinoxylans can strengthen the protein structure and influence to some extent rheological parameters, and combined with their high water-holding capacity strongly influence the end-use product quality [14]. In our study, SRCSu values are of great importance for all developed PLSR models. The explanation for the observed influence of SRCSu on developed PLSR models is in the wheat milling process itself. Each wheat flour mill stream differed from the others in terms of the arabinoxylan content, due to the distribution of pericarp, aleurone and endosperm layers between individual streams [36]. It is well known that the quantity of total protein significantly differs among the wheat flour mill streams [37]. Furthermore, the structure of glutenin polymers makes a significant contribution to differences in wheat flour mill stream quality [37,38]. This explains the observed importance of determining SRCLa and GPI values. The damaged starch content increased with the number of grinding steps as expected [39]. SRCSo values showed a moderate significance for the FWA parameter model, in agreement with the results of other authors [39]. The developed PLSR models were able to interpret correctly contribution of effects of different flour components in wheat flour mill streams to certain rheological parameters.

4. CONCLUSION

Results obtained by the SRC test in the present study showed differences in the quality of wheat flour mill streams. The multivariate analysis provided valuable information regarding the relationships between SRC values and rheological parameters. Additionally, the results have shown the possibility to model certain rheological parameters using SRC values with a satisfactory degree of accuracy. On the basis of SRCLa, SRCSu, SRCSo, SRCw and GPI values, models were developed for predicting FWA and E parameters. Also, the model for prediction of multiple rheological parameters exhibited good accuracy. Certainly, there are also limitations in the presented approach, since the total sample size was limited by the milling process. On the other hand, the developed PLSR models allowed general insights into the relationships between SRC and modelled rheological parameters, and provided a ground for future studies, in which a more powerful machine learning approach can be employed.

REFERENCES

[1] Duyvejonck AE, Lagrain B, Pareyt B, Courtin CM, Delcour JA. Relative contribution of wheat flour constituents to Solvent Retention Capacity profiles of European wheats. *J Cereal Sci.* 2011; 53(3): 312-8.

[2] Duyvejonck AE, Lagrain B, Dornez E, Delcour JA, Courtin CM. Suitability of solvent retention capacity tests to assess the cookie and bread making quality of European wheat flours. *LWT - Food Sci Technol*. 2012; 47(1): 56-63.

[3] Caballero PA, Gómez M, Rosell CM. Improvement of dough rheology, bread quality and bread shelf-life by enzymes combination. *J Food Eng*. 2007; 81(1): 42-53.

[4] Dobraszczyk BJ, Morgenstern MP. Rheology and the breadmaking process. *J Cereal Sci*. 2003; 38(3): 229-45.

[5] Khatkar BS, Fido RJ, Tatham AS, Schofield JD. Functional properties of wheat gliadins. I. Effects on mixing characteristics and bread making quality. *J Cereal Sci*. 2002; 35(3): 299-306.

[6] Huen J, Börsmann J, Matullat I, Böhm L, Stukenborg F, Heitmann M, Zannini E, Arendt EK. Wheat flour quality evaluation from the baker's perspective: comparative assessment of 18 analytical methods. *Eur Food Res Technol*. 2018; 244(3): 535-45.

[7] Slade L, Levine H. Structure-function relationships of cookie and cracker ingredients. In: Faridi H, ed. *The Science of Cookie and Cracker Production*. New York, NY: Chapman & Hall; 1994: 23-141.

[8] Lindgren A, Simsek S. Evaluation of Hard Red Spring Wheat Mill Stream Fractions Using Solvent Retention Capacity Test. *J Food Process Preserv*. 2016; 40(2): 131-9.

[9] Chung D, Keles S. Sparse partial least squares classification for high dimensional data. *Stat Appl Genet Mol Biol*. 2010; 9(1): 1-32.

[10] ICC: ICC-Standard No. 115/1. Method for using the Brabender Farinograph, ICC. 1996.

[11] ICC: ICC-Standard No. 114/1. Method for using the Brabender Extensograph, ICC. 1996.

[12] AACC: Approved methods of the American Association of Cereal Chemists, Edn. 10, Methods 56-11. AACC, St. Paul, MN, 2000.

[13] Bettge AD, Morris CF, DeMacon VL, Kidwell KK. Adaptation of AACC Method 56-11, solvent retention capacity, for use as an early generation selection tool for cultivar development. *Cereal Chem.* 2002; 79(5): 670-4.

[14] Kweon M, Slade L, Levine H. Solvent retention capacity (SRC) testing of wheat flour: Principles and value in predicting flour functionality in different wheat-based food processes and in wheat breeding-A review. *Cereal Chem*. 2011; 88(6): 537-52.

[15] Granato D, Santos JS, Escher GB, Ferreira BL, Maggio RM. Use of principal component analysis (PCA) and hierarchical cluster analysis (HCA) for multivariate association between bioactive compounds and functional properties in foods: A critical perspective. *Trends Food Sci Technol*. 2018; 72: 83-90.

[16] Lê S, Pagès J, Husson F. FactoMineR: An R Package for Multivariate Analysis. *J Stat Softw*. 2008; 25(1): 1-8.

[17] Kassambara A, Mundt F. Package "factoextra for R: Extract and Visualize the Results of Multivariate Data Analyses". *R Package version 1.0.4*. https://cran.r-project.org/web/packages/factoextra/index.html/. 2017.

[18] R Core Team 2019. A language and environment for statistical computing. *R Found Stat Comput Vienna, Austria URL http://www.R-project.org/*. 2019.

[19] Carrascal LM, Galván I, Gordo O. Partial least squares regression as an alternative to current regression methods used in ecology. *Oikos*. 2009; 118(5): 681-90.

[20] Martens H. Reliable and relevant modelling of real world data: A personal account of the development of PLS Regression. *Chemometrics and Intelligent Laboratory Systems*. 2001; 58(2): 85-95.

[21] Mevik B-H, Wehrens R. The pls Package: Principal Component and Partial Least Squares Regression in R. *J Stat Softw*. 2007; 18: 1-24.

[22] Andrew YN. Preventing Overfitting of Cross-Validation Data. In: *ICML '97 Proceedings of the Fourteenth International Conference on Machine Learning*. Nashville, Tennessee, USA, 1997, pp. 245-253.

[23] Lachenbruch PA, Mickey MR. Estimation of Error Rates in Discriminant Analysis. *Technometrics*. 1968; 10(1): 1-11.

[24] Kaur A, Singh N, Kaur S, Ahlawat AK, Singh AM. Relationships of flour solvent retention capacity, secondary structure and rheological properties with the cookie making characteristics of wheat cultivars. *Food Chem*. 2014; 158: 48-55.

[25] Barrera GN, Pérez GT, Ribotta PD, León AE. Influence of damaged starch on cookie and bread-making quality. *Eur Food Res Technol*. 2007; 225(1): 1-7.

[26] Kweon M, Slade L, Levine H, Martin R, Andrews L, Souza E. Effects of extent of chlorination, extraction rate, and particle size reduction on flour and gluten functionality explored by solvent retention capacity (SRC) and mixograph. *Cereal Chem*. 2009; 86(2): 221-4.

[27] Hammed AM, Ozsisli B, Ohm JB, Simsek S. Relationship between solvent retention capacity and protein molecular weight distribution, quality characteristics, and breadmaking functionality of hard red spring wheat flour. *Cereal Chem*. 2015; 92(5): 466-74.

[28] Abdi H, Williams LJ. Principal component analysis. *Wiley Interdiscip Rev Comput Stat*. 2010; 2(4): 433-59.

[29] Ram S, Dawar V, Singh RP, Shoran J. Application of solvent retention capacity tests for the prediction of mixing properties of wheat flour. *J Cereal Sci*. 2005; 42(2): 261-6.

[30] Guttieri MJ, Bowen D, Gannon D, O'Brien K, Souza E. Solvent retention capacities of irrigated soft white spring wheat flours. *Crop Sci*. 2001; 41(4): 1054-1061.

[31] Colombo A, Pérez GT, Ribotta PD, León AE. A comparative study of physicochemical tests for quality prediction of Argentine wheat flours used as corrector flours and for cookie production. *J Cereal Sci*. 2008; 48(3): 775-80.

[32] Gaines CS. Collaborative study of methods for solvent retention capacity profiles (AACC method 56-11). *Cereal Foods World*. 2000; 45(7): 303-6.

[33] Xiao ZS, Park SH, Chung OK, Caley MS, Seib PA. Solvent retention capacity values in relation to hard winter wheat and flour properties and straight-dough breadmaking quality. *Cereal Chem*. 2006; 83(5): 465-71.
[34] Mastilović J, Kevrešan Ž, Vukić M, Ivanović M, Radovanović J, Džinić D. Possibilities for utilization of dietary fiber-rich supplement from pepper (*Capsicum annum L.)* processing waste in bakery products. *J Eng Process Manag*. 2018; 10(1): 28-33.
[35] Goesaert H, Brijs K, Veraverbeke WS, Courtin CM, Gebruers K, Delcour JA. Wheat flour constituents: How they impact bread quality, and how to impact their functionality. *Trends in Food Science and Technology*. 2005; 16(1-3): 12-30.
[36] Delcour JA, Van Win H, Grobet PJ. Distribution and structural variation of arabinoxylans in common wheat mill streams. *J Agric Food Chem*. 1999; 47(1): 271-5.
[37] Wang YG, Khan K, Hareland G, Nygard G. Distribution of protein composition in bread wheat flour mill streams and relationship to breadmaking quality. *Cereal Chem*. 2007; 84(3): 271-5.
[38] Wang YG, Khan K, Hareland G, Nygard G. Quantitative glutenin composition from gel electrophoresis of flour mill streams and relationship to breadmaking quality. *Cereal Chem*. 2006; 83(3): 293-9.
[39] Iuliana B, Georgeta S, Violeta I, Iuliana A. Physicochemical and rheological analysis of flour mill streams. *Cereal Chem*. 2010; 87(2): 112-7.

САЖЕТАК

Примена тестова задржавања растварача за предвиђање реолошких параметара пасажних брашна млевења пшенице

Милан С. Вукић[1,2], Елизабет П. Јањић Хајнал[3], Јасна С. Мастиловић[3], Драган П. Вујадиновић[2], Марко М. Ивановић[2] и Драгана М. Шороња-Симовић[1]

[1]*Универзитет у Новом Саду, Технолошки факултет, Нови Сад, Србија*
[2]*Универзитет у Источном Сарајеву, Технолошки факултет Зворник, Зворник, Република Српска, Босна и Херцеговина*
[3]*Универзитет у НовомСаду, Институт за прехрамбене технологије, Нови Сад, Србија*

(Научни рад)

У раду је приказана веза између реолошких својстава теста и особина бубрења појединих полимера пшеничног брашна. Својства бубрења су мерена применом тестова способности апсорпције растварача (енгл. Solvent Retention Capacity, SRC). За одређене реолошке параметре утврђени су значајни коефицијенти корелације. У настојању да се оствари додатни увид у мерене особине, коришћена је мултиваријатна анализа како би се испитали односи између параметара апсорпције растварача (SRC) и параметара добијених реолошким тестовима. Да би се открили релевантни односи између параметара, извршен је корак истраживања података кроз анализу главних компоненти. Затим су развијени модели мултиваријатне регресије методом парцијалних најмањих квадрата (енгл. Partial Least Squares Regression, PLSR), за предвиђање одабраних емпиријских реолошких параметара из SRC параметара. Обрада експерименталних података указује на могућност параметара теста апсорпције растварача за предвиђање реолошких својстава у вези са одговарајућим математичким моделом. Представљени приступ могао би бити користан за брзо предвиђање карактеристика пасажних брашна и за оптимизацију квалитета крајњег брашна.

Кључне речи: регресија методом парцијалних најмањих квадрата, моделовање, бубрење полимера, реологија