

## Factor Analysis and Modelling for Rapid Quality Assessment of Croatian Wheat Cultivars with Different Gluten Characteristics

Želimir Kurtanjek<sup>1</sup>, Daniela Horvat<sup>2\*</sup>, Damir Magdić<sup>3</sup> and Georg Drezner<sup>2</sup>

<sup>1</sup>Faculty of Food Technology and Biotechnology, University of Zagreb, Pierottijeva 6, HR-10000 Zagreb, Croatia

<sup>2</sup>Agricultural Institute Osijek, Južno predgrađe 17, HR-31103 Osijek, Croatia

<sup>3</sup>Faculty of Food Technology, University J.J. Strossmayer of Osijek, F. Kuhača 18, HR-31000 Osijek, Croatia

Received: June 27, 2007

Accepted: March 10, 2008

### Summary

Factor analysis and multivariate chemometric modelling for rapid assessment of baking quality of wheat cultivars from Slavonia region, Croatia, have been applied. The cultivars Žitarka, Kata, Monika, Ana, Demetra, Divana and Sana were grown under controlled conditions at the experimental field of Agricultural Institute Osijek during three years (2000–2002). Their quality properties were evaluated by 45 different chemical, physical and biochemical variables. The measured variables were grouped as: indirect quality parameters (6), farinographic parameters (7), extensographic parameters (5), baking test parameters (2) and reversed phase-high performance liquid chromatography (RP-HPLC) of gluten proteins (25). The aim of this study is to establish minimal number (three), *i.e.* principal factors, among the 45 variables and to derive multivariate linear regression models for their use in simple and fast prediction of wheat properties. Selection of the principal factors based on the principal component analysis (PCA) has been applied. The first three main factors of the analysis include: total glutenins (TGT), total  $\omega$ -gliadins (T $\omega$ -) and the ratio of dough resistance/extensibility (R/Ext). These factors account for 76.45 % of the total variance. Linear regression models gave average regression coefficients (*R*) evaluated for the parameter groups: indirect quality *R*=0.91, baking test *R*=0.63, farinographic *R*=0.78, extensographic *R*=0.95 and RP-HPLC of gluten data *R*=0.90. Errors in the model predictions were evaluated by the 95 % significance intervals of the calibration lines. Practical applications of the models for rapid quality assessment and laboratory experiment planning were emphasized.

**Key words:** principal component analysis, wheat, technological quality, gluten proteins, RP-HPLC

### Introduction

The wheat quality is essentially determined by the gluten protein composition and concentration. Gliadins and glutenins are the main protein fractions present in

wheat gluten and they are responsible for technological and nutritional quality of wheat-based products. In particular, polymeric glutenins are mainly responsible for dough elasticity, whereas gliadins, as monomeric proteins, confer viscous flow and extensibility to the gluten

\*Corresponding author; Phone: ++385 31 515 517; E-mail: [daniela.horvat@poljin.hr](mailto:daniela.horvat@poljin.hr)

complex (1–7). The specific composition of wheat grain proteins and the ratios of particular gluten groups are genetically determined, while the environment influences protein concentration, amount of different gluten groups and size distribution of polymeric proteins (8–12).

A better understanding of the biochemical background on a protein level for the wheat quality variation determined by genotype and environment is very important, therefore wheat breeders of the Agricultural Institute Osijek have given emphasis on creating bread-making cultivars with higher gluten strength and with quality stability in different cultivation years and locations (13).

The aim of this study is to investigate functional relationship among wheat technological properties and quantity of gluten proteins of Croatian wheat cultivars with different gluten characteristics by multivariate statistical models based on principal component analysis (PCA).

## Materials and Methods

### *Wheat samples*

The pre-basic seed of winter wheat cultivars grown during the three-year period (2000–2002) at the experimental field of the Agricultural Institute Osijek at Osijek location was analyzed. Studied cultivars Žitarka, Kata, Monika, Ana and Demetra were created at the Agricultural Institute Osijek, Croatia. Cultivar Sana was developed by the Zagreb Bc Institute for Breeding and Production of Field Crops, Zagreb, Croatia and Divana by the Jošt-Seed, Križevci, Croatia.

### *Wheat and flour quality assessments*

The crude protein content of sample grains was determined on a dry mass basis by near-infrared transmission (NIT) spectroscopy (Infratec 1241, Foss Tecator, Sweden). Zeleny sedimentation value was determined according to the International Associations for Cereal Science and Technology (ICC) Standard No. 116/1 (14) and Hagberg falling number according to the ICC Standard No. 107/1 (15). Gluten strength of flour (ash content 0.55 %, Brabender Quadrumat Sr. Mill, Germany) was measured by Gluten Index Method (ICC Standard No. 155) (16) on a Glutomatic 2200 (Perten Instruments, Sweden) and by determining dough rheological properties on a Brabender farinograph (ICC Standard No. 115/1) (17) and extensograph (ICC Standard No. 114/1) (18). Baking test was done according to the ICC Standard No. 131 (19).

### *RP-HPLC analysis of gluten proteins*

Gluten proteins were analyzed by using reversed phase-high performance liquid chromatography (RP-HPLC) system (Integral 4000, PerkinElmer, USA), following the quantitative extraction procedure of Wieser *et al.* (20). The gluten proteins as well as crude protein content were analyzed in grain samples. For the quantification of wheat proteins, whole meal flour (100 mg) was extracted stepwise with 0.4 M NaCl (albumins and globulins), 50 % 1-PrOH (gliadins) and 50 % 1-PrOH+2

M urea+0.05 M of Tris-HCl (pH=7.5)+1 % DTT (glutenins). The separation of gluten components was carried out on Supelcosil LC 318 column (25×0.46 cm) at 50 °C. A linear elution gradient (0 min 28 % ACN/0.1 % TFA, 30 min 56 % ACN/0.1 % TFA) was applied to separate gliadin and glutenin components. Eluted proteins were monitored at 210 nm. Gluten proteins were eluted according to different surface hydrophobicity in the series  $\omega$ 5-,  $\omega$ 1,2-,  $\alpha$ - and  $\gamma$ -gliadins (gliadin extract) and  $\omega$ b-gliadins, high-molecular-mass glutenin subunits (HMM-GS) and low-molecular-mass glutenin subunits (LMM-GS) (glutenin extract) (14). The repeatability of the extraction procedure and RP-HPLC analysis was within  $\pm 5$  %, except for the minor  $\omega$ -gliadins. The areas under the RP-HPLC chromatograms, expressed as arbitrary units (AU), were calculated per milligram of flour and used as a direct measure for the amount of particular fraction of wheat proteins. Proportions (%) of protein fractions were calculated using the total protein area under chromatographic curves of albumins and globulins, gliadins and glutenins.

### *Multivariate chemometric analysis*

The data matrix  $X$  with the dimension of 21 rows corresponding to 7 wheat cultivars and 3 years of production, and 45 columns corresponding to individual experimentally determined parameters was analyzed. The columns of the  $X$  matrix were structured in the following order: genotype (cultivar index), year of production, indirect quality parameters, farinographic and extensographic properties, baking parameters and RP-HPLC data. The groups of the experimental parameters with corresponding measurement units are listed in Tables 1–3. Each parameter was associated with a corresponding index, ranging from 1 to 45. The indices were used for graphic presentation of the variable projections as the loadings into the space of principal components.

Together with the original data matrix  $X$ , which contains average values of 2 parallel experiments for each parameter, a double sized matrix of dimension 42×45 was also analysed, including each of the two parallel experimental data for validation of the model predictions based on average values.

For statistical analysis, the original data matrix  $X$  was transformed by autoscaling process into the new data matrix with zero average and normalized standard deviation for each parameter.

## Results and Discussion

Chemical, physical and biochemical properties of gluten are known to be strongly correlated. However, theoretical nature of the interrelationships is unknown, but their quantitative relationships can be deduced in form of linear and nonlinear statistical models. Degree of linear relationships between samples (cultivars) and variables (quality properties) is investigated by their covariance matrix. The resulting covariance matrix has 45×45 partial correlation coefficients and for the purpose of analysis and modelling, it is decomposed into the space of principal, *i.e.* statistically significant components.

Table 1. Indirect quality parameters and baking test results with abbreviated names and the corresponding indices (*N*)

Indirect quality parameters	Variable	<i>N</i>	Baking results	Variable	<i>N</i>
Year 2000–2002	Year	1	<i>V</i>	Bread volume/cm <sup>3</sup>	7
P	Protein on dry mass basis/%	2	<i>h/d</i>	Ratio of height/ diameter of loaf	8
SED	Sedimentation value/cm <sup>3</sup>	3			
WG	Wet gluten/%	4			
GI	Gluten index	5			
FN	Falling number/s	6			

Table 2. Farinographic and extensographic properties with abbreviated names and the corresponding indices (*N*)

Farinographic properties	Variable	<i>N</i>	Extensographic properties	Variable	<i>N</i>
WA	Water absorption/%	9	E	Dough energy/cm <sup>2</sup>	16
DDT	Dough development time/min	10	R	Dough resistance after 5 min/EU	17
STAB	Dough stability/min	11	R <sub>max</sub>	Dough resistance at curve maximum/EU	18
R	Dough resistance/min	12	Ext	Dough extensibility/mm	19
DS	Dough degree of softening/FU	13	R/Ext	Ratio of resistance/ extensibility	20
FQN	Farinograph quality number	14			
QG	Quality group	15			

FU=farinographic units, EU=extensographic units

Table 3. RP-HPLC data with abbreviated names and the corresponding indices (*N*)

Variable	<i>N</i>	Variable	<i>N</i>		
TAG	Total albumins and globulins/AU	21	TGT	Total glutenins/AU	34
TGLU	Total gluten/AU	22	T $\omega$ b-	Total $\omega$ b-gliadins/AU	35
TGLI	Total gliadins/AU	23	w( $\omega$ b-)/%	Fraction of $\omega$ b-gliadins	36
T $\omega$ 5-	Total $\omega$ 5-gliadins/AU	24	THMM	Total HMM/AU	37
w( $\omega$ 5-)/%	Fraction of $\omega$ 5-gliadins	25	w(HMM)/%	Fraction of HMM	38
T $\omega$ 1,2-	Total $\omega$ 1,2-gliadins/AU	26	TLMM	Total LMM/AU	39
w( $\omega$ 1,2-)/%	Fraction of $\omega$ 1,2-gliadins	27	w(LMM)/%	Fraction of LMM	40
T $\omega$ -	Total of all $\omega$ -gliadins/AU	28	GLI/GLU	Ratio of gliadins/glutenins	41
w( $\omega$ -)/%	Fraction of all $\omega$ -gliadins	29	w(AG)/%	Fraction of albumins and globulins in total protein	42
T $\alpha$ -	Total $\alpha$ -gliadins/AU	30	w(GLU)/%	Fraction of gluten in total protein	43
w( $\alpha$ -)/%	Fraction of $\alpha$ -gliadins	31	w(GLI)/%	Fraction of gliadins in total protein	44
T $\gamma$ -	Total $\gamma$ -gliadins/AU	32	w(GT)/%	Fraction of glutenins in total protein	45
w( $\gamma$ -)/%	Fraction of $\gamma$ -gliadins	33			

AU=arbitrary units

The principal components are determined by the singular value decomposition of the covariance matrix  $X^T X$  with the dimension of 45×45 of the autoscaled data. The numerical procedure provided by STATISTICA (21) software was applied. The total variability of the samples is decomposed into the decreasing order of the variances of the sample projections into subspaces of the principal components. The result of decomposition is

graphically depicted in the form of a scree plot presented in Fig. 1. The scree curve can be approximated with two tangent lines, drawn from the origin and end points. The lines intersect at the knee point which indicates the breaking point between the deterministic and random dispersion of all data given in  $X$  matrix. From the position of the knee point of the scree plot it can be concluded that the deterministic dispersion of the data

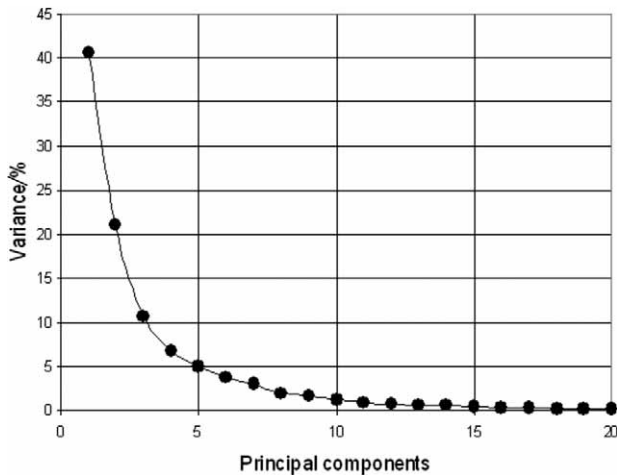


Fig. 1. Contributions of the total variance accounted by the principal components

can be explained by the first 3 or 4 components, while the rest of the data dispersion could be considered as random.

The first four principal components account correspondingly for the following percentage of the total variance: 43.47, 22.18, 10.80 and 6.23 %, or the cumulative effect of the first three components is 76.45 % and the first four components account for 82.68 %.

In Figs. 2a–c, bi-plots of the loadings (variable contributions to the principal components) and the scores (cultivar samples) for the first three principal components are presented. The loadings are depicted in the spectral form for the individual property group (indirect quality, farinographic and extensographic, baking and RP-HPLC). Each property is indexed and the loading spectra are represented as a continuous curve obtained by interpolation over the set of indices from  $N=1-45$  correspond to each measured variable as listed in Tables 1–3. The scores are also denoted for each cultivar and are grouped correspondingly to their year of production (2000–2002).

Clustering of the scores (cultivars) is depicted in Fig. 3 in a two-dimensional plane of the first two principal components. Clustering of the scores (wheat cultivars) can be interpreted as an indication of similarity between interrelated chemical, physical and biochemical gluten properties of cultivars and effects of different atmospheric conditions during three consecutive years of production.

**Factor analysis**

Principal factors were determined as the variables with the highest projection scores on the principal components (Table 4). The data matrix of average values for 21 samples (cultivars and harvests) for all 45 variables was analysed by the cluster analysis. Clusters with Euclidean non-weighted distances were applied. The clusters with the indication of the principal factors are presented in Fig. 4. It can be observed that the dominant clusters are defined by the RP-HPLC data, while the least pronounced cluster corresponds to the indirect quality and baking test parameters. The first two princi-

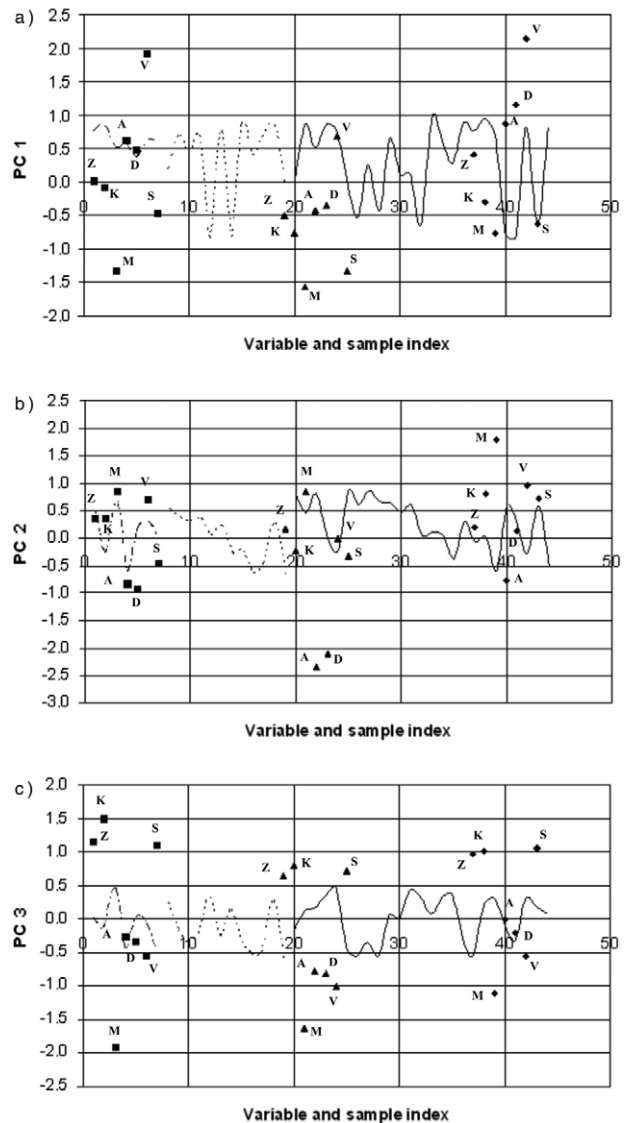


Fig. 2. Bi-plots of the loadings and scores on the principal components: a) bi-plot on the first principle component PC 1 accounting for 43.47 % of the total variance; b) bi-plot on the second principle component PC 2 accounting for 22.18 % of the total variance; c) bi-plot on the third principle component PC 3 accounting for 10.8 % of the total variance

Loadings of the variables are grouped and represented as follows: indirect quality and baking parameters (---) with corresponding indices 1–8, farinographic and extensographic properties (---) with indices 9–20, RP-HPLC data (—) with indices from 21–45

Scores of the wheat cultivars are marked with letters (Z – Žitarka, K – Kata, M – Monika, A – Ana, D – Demetra, V – Divna, S – Sana) and the harvest years are marked as follows: (■) 2000, (▲) 2001, (◆) 2002

pal factors are associated with RP-HPLC clusters, while the third factor belongs to the cluster associated with extensographic properties. The selected principal factors are: total glutenins (TGT), total  $\omega$ -gliadins ( $T\omega$ -) and the ratio between dough resistance and extensibility (R/Ext). It can be observed from Fig. 4 that the principal factors correspond to the main clusters. Chemometric analysis has revealed very strong functional relationship between

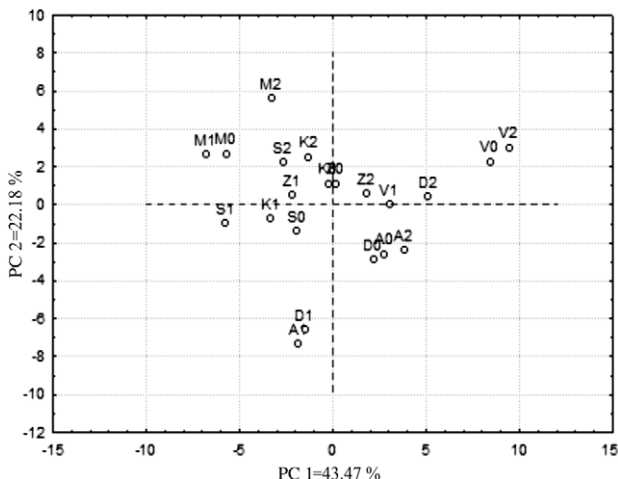


Fig. 3. Scores of the wheat cultivars projected onto the plane of the first and second principal components. The cultivars are marked with letters (Z – Žitarka, K – Kata, M – Monika, A – Ana, D – Demetra, V – Divna, S – Sana) and the harvest years are marked with numbers (0 for 2000, 1 for 2001 and 2 for 2002)

Table 4. Projections of the first three principal factors (TGT, Tω-, R/Ext) on the first three principal components (PC 1–3)

Factors	PC 1	PC 2	PC 3
TGT (34)	0.97738		
Tω- (28)		0.864972	
R/Ext (20)			-0.61087
Variance/%	43.47	22.18	10.80

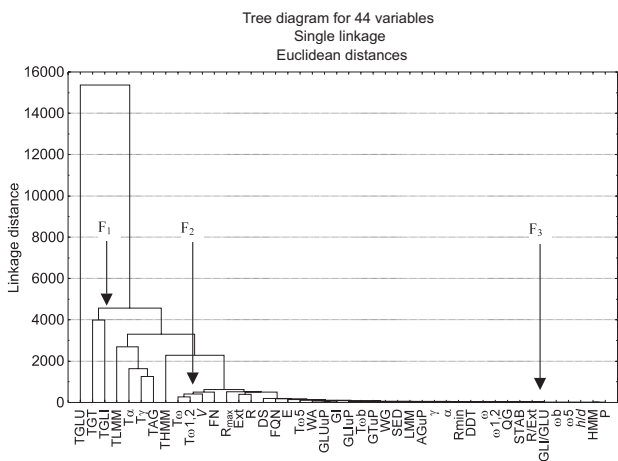


Fig. 4. Cluster analysis of all the set of 45 experimental variables. The positions of the principal factors F<sub>1</sub>, F<sub>2</sub> and F<sub>3</sub> are denoted with arrows

gluten protein fractions, wheat and baking quality, which was in accordance with our previous findings (22).

Regression models

Based on the principal factors, linear multivariate regression models were determined for all measured variables. The models are given by the common linear multivariate form:

$$y_i = b_0 + b_1 \cdot (R/Ext) + b_2 \cdot (TGT) + b_3 \cdot (T\omega-); i=1, 2, \dots, 45 \quad /1/$$

where y<sub>i</sub> are measured variables, b<sub>k</sub> (k=0, 1, 2, 3) are the model parameters, and index i denotes 42 ordered models, excluding the 3 main factors, which are given in accordance with the property indexing listed in Tables 1–3.

The model was evaluated in the scaled form in which each variable was transformed into a new set of standard and dimensionless variables (zero mean, and standard deviation equal to one) by the following relations:

$$z_i = \frac{y_i - \bar{y}_i}{\sigma(y_i)}; \quad u = \frac{R/Ext - \bar{R/Ext}}{\sigma(R/Ext)}; \quad /2/$$

$$v = \frac{TGT - \bar{TGT}}{\sigma(TGT)}; \quad t = \frac{T\omega - \bar{T\omega}}{\sigma(T\omega)}$$

The autoscaled model was given by:

$$z_i = \beta_1 \cdot u + \beta_2 \cdot v + \beta_3 \cdot t \quad /3/$$

The model parameters β<sub>i</sub> were estimated by minimization of the unweighed sum of error squares by the least square (LS) method (21). Accuracies of the model predictions were evaluated by the multiple regression coefficients R and the corresponding relative standard errors δ. The model parameters, correlations and errors are given in Tables 5a–e. The average correlation R=0.92 and relative standard error δ=7.9 % were obtained for the indirect quality parameters. The exception was the falling number, for which the model was inadequate. The lowest correlation R=0.63 and error δ=8.2 % were obtained for the baking test data. The models were also adequate for the group of farinographic properties, with the average correlation of R=0.80 and error δ=8 %, with

Table 5. Models of the properties based on the three principal factors: TGT, Tω- and R/Ext ratio. Evaluation of the models is given by standard errors of the corresponding parameters e, multiple regression coefficients R and relative standard error δ for the property prediction. Statistically significant model parameters at the level of significance 0.05 are given in bold

Table 5a. Models of indirect quality parameters based on the three principal factors

Variable	Indirect quality parameters				
	R/Ext	TGT	Tω	R	δ/%
	β <sub>1</sub>	β <sub>2</sub>	β <sub>3</sub>		
P	-0.0779	<b>0.3855</b>	<b>0.6919</b>	<b>0.9230</b>	4.1
SED	<b>0.3895</b>	-0.0441	<b>0.8426</b>	<b>0.8690</b>	9.6
WG	-0.6124	0.1746	<b>0.4958</b>	<b>0.9364</b>	6.5
GI	<b>0.7454</b>	-0.1578	<b>0.6160</b>	<b>0.9202</b>	7.4
FN	0.1068	0.2062	0.2717	0.3754	11.0
	0.2418	0.2541	0.2389		

Table 5b. Models of baking test results based on the principal factors

Variable	Baking test				
	R/Ext	TGT	T $\omega$	R	$\delta$ /%
	$\beta_1$	$\beta_2$	$\beta_3$		
	e( $\beta_1$ )	e( $\beta_2$ )	e( $\beta_3$ )		
V	-0.0362 0.1983	0.1625 0.2085	<b>0.5670</b> <b>0.1960</b>	<b>0.6494</b>	9.2
<i>h/d</i>	0.2889 0.2035	0.1857 0.2140	<b>0.5388</b> <b>0.2011</b>	<b>0.6253</b>	7.4

Table 5c. Models of farinographic properties based on the principal factors

Variable	Farinographic properties				
	R/Ext	TGT	T $\omega$	R	$\delta$ /%
	$\beta_1$	$\beta_2$	$\beta_3$		
	e( $\beta_1$ )	e( $\beta_2$ )	e( $\beta_3$ )		
WA	<b>-0.5426</b> <b>0.2050</b>	0.0336 0.2155	0.1955 0.2025	<b>0.6185</b>	2.9
DDT	-0.0639 0.1689	0.2292 0.1776	<b>0.6347</b> <b>0.1669</b>	<b>0.7619</b>	14.0
STAB	0.2608 0.2127	0.4296 0.2236	0.3125 0.2102	0.5787	37.0
R	-0.0178 0.1726	0.2530 0.1815	<b>0.6204</b> <b>0.1706</b>	<b>0.7496</b>	1.4
DS	<b>-0.4625</b> <b>0.1176</b>	-0.1411 0.1237	<b>-0.8014</b> <b>0.1162</b>	<b>0.8925</b>	11.0
FQN	0.1232 0.2768	0.1610 0.1692	<b>0.6716</b> <b>0.1590</b>	<b>0.7869</b>	18.0
QG	-0.0782 0.1402	-0.2553 0.1474	<b>-0.7375</b> <b>0.1385</b>	<b>0.8432</b>	0.9

Table 5d. Models of extensographic properties based on the principal factors

Variable	Extensographic properties				
	R/Ext	TGT	T $\omega$	R	$\delta$ /%
	$\beta_1$	$\beta_2$	$\beta_3$		
	e( $\beta_1$ )	e( $\beta_2$ )	e( $\beta_3$ )		
E	<b>0.5632</b> <b>0.0802</b>	0.1112 0.0843	<b>0.8343</b> <b>0.0792</b>	<b>0.9515</b>	10.0
R	<b>0.9378</b> <b>0.0644</b>	-0.0190 0.0677	<b>0.4009</b> <b>0.0636</b>	<b>0.9690</b>	7.1
R <sub>max</sub>	<b>0.8205</b> <b>0.0705</b>	0.0706 0.0742	<b>0.6439</b> <b>0.0697</b>	<b>0.9627</b>	9.1
Ext	<b>-0.2302</b> <b>0.0922</b>	0.0117 0.0969	<b>0.8686</b> <b>0.0912</b>	<b>0.9353</b>	9.5

Table 5e. Models of RP-HPLC results based on the principal factors

Variable	RP-HPLC data				
	R/Ext	TGT	T $\omega$	R	$\delta$ /%
	$\beta_1$	$\beta_2$	$\beta_3$		
	e( $\beta_1$ )	e( $\beta_2$ )	e( $\beta_3$ )		
TAG	-0.1168 0.1387	<b>0.8451</b> <b>0.1458</b>	-0.2092 0.1370	<b>0.8470</b>	9.1
TGLU	<b>-0.1651</b> <b>0.0408</b>	<b>0.2549</b> <b>0.0428</b>	<b>0.8194</b> <b>0.0403</b>	<b>0.9876</b>	2.7
TGLI	<b>-0.3559</b> <b>0.0880</b>	<b>0.5501</b> <b>0.0925</b>	<b>0.3700</b> <b>0.0869</b>	<b>0.9413</b>	5.1
T $\omega$ 5-	-0.0980 0.1253	-0.2198 0.1318	<b>0.9148</b> <b>0.1239</b>	<b>0.8769</b>	8.7
<i>w</i> ( $\omega$ 5-)/%	-0.0954 0.1504	<b>-0.4572</b> <b>0.1581</b>	<b>0.8486</b> <b>0.1487</b>	<b>0.8710</b>	10.9
T $\omega$ 1,2-	0.0295 0.0377	<b>1.0583</b> <b>0.0396</b>	<b>-0.2753</b> <b>0.0373</b>	<b>0.9890</b>	2.5
<i>w</i> ( $\omega$ 1,2-)/%	0.1379 0.0751	<b>0.9197</b> <b>0.0789</b>	<b>-0.7704</b> <b>0.0742</b>	0.4950	8.4
<i>w</i> ( $\omega$ -)/%	<b>0.1408</b> <b>0.0607</b>	<b>0.9971</b> <b>0.0639</b>	<b>-0.6843</b> <b>0.0600</b>	<b>0.9725</b>	1.7
T $\alpha$ -	-0.2016 0.1114	<b>0.4799</b> <b>0.1172</b>	<b>0.5239</b> <b>0.1101</b>	<b>0.9041</b>	3.8
<i>w</i> ( $\alpha$ -)/%	-0.1951 0.2266	0.3954 0.2382	-0.0198 0.2240	0.4950	6.4
T $\gamma$ -	<b>-0.5399</b> <b>0.2008</b>	0.1669 0.2111	0.0650 0.1984	0.6383	8.9
<i>w</i> ( $\gamma$ -)/%	-0.3722 0.1857	-0.1927 0.1952	<b>-0.5978</b> <b>0.1835</b>	<b>0.7023</b>	7.9
T $\omega$ b-	0.0041 0.1989	0.0198 0.2091	<b>0.6405</b> <b>0.1965</b>	0.6468	16.5
<i>w</i> ( $\omega$ b-)/%	0.1321 0.2349	-0.2849 0.2470	0.3908 0.2322	0.4343	14.3
THMM	<b>0.1849</b> <b>0.0698</b>	<b>0.3571</b> <b>0.0734</b>	<b>0.8139</b> <b>0.0690</b>	<b>0.9635</b>	6.8
<i>w</i> (HMM)/%	<b>0.5126</b> <b>0.1313</b>	0.2452 0.1381	<b>0.7100</b> <b>0.1298</b>	<b>0.8639</b>	9.8
TLMM	<b>-0.0899</b> <b>0.0312</b>	<b>-0.1743</b> <b>0.0328</b>	<b>1.0261</b> <b>0.0308</b>	<b>0.9928</b>	2.5
<i>w</i> (LMM)/%	0.0462 0.0699	<b>-0.7256</b> <b>0.0735</b>	<b>0.9141</b> <b>0.0692</b>	<b>0.9633</b>	3.4
GLI/GLU	<b>-0.2616</b> <b>0.0809</b>	<b>0.5255</b> <b>0.0851</b>	<b>-0.9262</b> <b>0.0799</b>	<b>0.9506</b>	5.0
<i>w</i> (AG)/%	0.0837 0.1164	<b>0.4937</b> <b>0.1224</b>	<b>-0.9322</b> <b>0.1151</b>	<b>0.8949</b>	5.5
<i>w</i> (GLI)/%	<b>-0.4522</b> <b>0.1221</b>	<b>0.3797</b> <b>0.1283</b>	<b>-0.7807</b> <b>0.1206</b>	<b>0.8837</b>	2.3
<i>w</i> (GT)/%	<b>0.2144</b> <b>0.0619</b>	<b>-0.4998</b> <b>0.0650</b>	<b>0.9809</b> <b>0.0611</b>	<b>0.9714</b>	2.8

the exception of the dough stability for which the model was not applicable ( $\delta=35\%$ ). The best model predictions were determined for the groups of extensographic and RP-HPLC gluten properties. The correlations were about  $R=0.90$  and relative standard errors  $\delta=8.3\%$ . The models were inadequate for estimation of fractions of  $\alpha$ -gliadins and  $\omega$ b-gliadins.

To illustrate the obtained models, a model for protein content (P) with 95% significance intervals around the calibration line is presented in Fig. 5. Measured and predicted protein content values for all cultivars in the

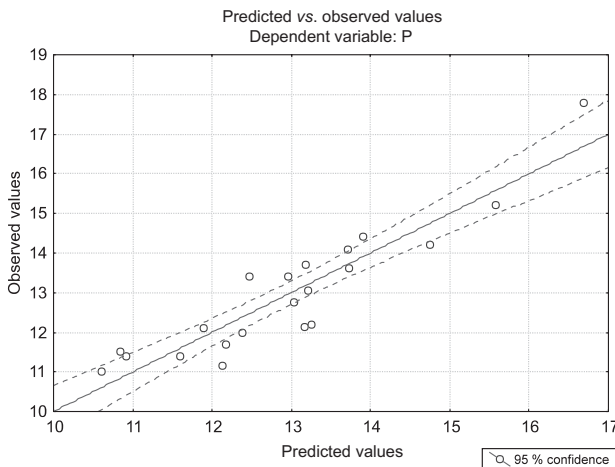


Fig. 5. Predicted *vs.* measured values of protein content (P)

period of three years of production are shown. As expected, the models show less accuracy at the extreme values, minimal and maximal protein content and most of the predicted data are in the range of average values and within the confidence boundaries. These are the same general features that have been observed in most models for other properties. The exceptions are the cases of inappropriate models for indirect quality parameter FN (falling number), farinographic property STAB (dough stability) and fraction of  $\alpha$ -gliadins.

## Conclusions

In conclusion, application of the principal components resulted in linear mathematical models with 3 principal factors by which a very effective reduction in the number of measured variables was achieved. The models proved to be applicable for 7 Croatian wheat cultivars in 3 years of production under significantly different climate conditions. The model prediction errors for the wide range of physical, chemical and quality properties were at the level of 8%. The presented methodology of principal component analysis accompanied by the graphical presentation of bi-plots and regression models, based on the three principals, provides several important advantages over classical methods. The most important advantage is the possibility to provide scenarios for fingerprinting patterns of cultivar (genomic) potentials and technological quality properties. Fast screening by the three principal factors greatly reduces load on laboratory work, since most of the properties can be pre-

dicted by the developed models at acceptable level of error instead of being measured. Furthermore, the graphic presentation by bi-plots provides intuitive and quantitative classification of multidimensional data, which greatly enhances ability to understand data patterns of cultivar (genomic) potentials under various agrotechnical treatments and meteorological conditions.

## References

1. Z. Jurković, R. Sudar, G. Drezner, D. Horvat, The HMW glutenin subunit composition of OS wheat cultivars and their relationship with bread-making quality, *Cereal Res. Commun.* 28 (2000) 271–277.
2. S. Uthayakumaran, F.L. Stoddard, P.W. Gras, F. Bekes, Effect of incorporated glutenins on functional properties of wheat dough, *Cereal Chem.* 77 (2000) 737–743.
3. S. Jood, J.D. Schofield, A.A. Tsiami, S. Bollecker, Effect of glutenin subfractions on bread-making quality of wheat, *Int. J. Food Sci. Technol.* 36 (2001) 573–584.
4. E. Johansson, H. Nilsson, H. Mazhar, J. Skeritt, F. MacRitchie, G. Svensson, Seasonal effects on storage proteins and gluten strength in four Swedish wheat cultivars, *J. Sci. Food Agric.* 82 (2002) 1305–1311.
5. H. Goesaert, K. Brijs, W.S. Veraverbeke, C.M. Courtin, K. Gebruers, J.A. Delcour, Wheat flour constituents: How they impact bread quality, and how to impact their functionality, *Trends Food Sci. Technol.* 16 (2005) 12–30.
6. D. Horvat, Z. Jurković, R. Sudar, D. Pavlinić, G. Šimić, The relative amounts of HMW glutenin subunits of OS wheat cultivars in relation to bread-making quality, *Cereal Res. Commun.* 30 (2002) 415–422.
7. D. Horvat, Z. Jurković, G. Drezner, G. Šimić, D. Novoselović, K. Dvojković, Influence of gluten proteins on technological properties of Croatian wheat cultivars, *Cereal Res. Commun.* 34 (2006) 1177–1184.
8. R. Kuktaite, E. Johansson, G. Juodeikiene, Composition and concentration of proteins in Lithuanian wheat cultivars: Relationships with bread-making quality, *Cereal Res. Commun.* 28 (2000) 195–202.
9. H. Wieser, R. Kieffer, Correlations of the amount of gluten protein types to the technological properties of wheat flours determined on a micro-scale, *J. Cereal Sci.* 34 (2001) 19–27.
10. R. Lásztity, Prediction of wheat quality – Success and doubts, *Periodica Polytechnica Ser. Chem. Eng.* 46 (2003) 39–49.
11. E. Johansson, M.L. Prieto-Linde, G. Svensson, Influence of nitrogen application rate and timing on grain protein composition and gluten strength in Swedish wheat cultivars, *J. Plant Nutr. Soil Sci.* 167 (2004) 345–350.
12. B. Lagrain, K. Brijs, J.A. Delcour, Impact of redox agents on the physico-chemistry of wheat gluten proteins during hydrothermal treatment, *J. Cereal Sci.* 44 (2006) 49–53.
13. G. Drezner, K. Dvojković, D. Horvat, D. Novoselović, A. Lalić, D. Babić, J. Kovačević, Grain yield and quality of winter wheat genotypes in different environments, *Cereal Res. Commun.* 34 (2006) 457–460.
14. Determination of sedimentation value (ac. to Zeleny) as an approximate measure of baking quality, ICC Standard No. 116/1, International Association for Cereal Science and Technology, Vienna, Austria (1994).
15. Determination of the 'Falling Number' according to Hagberg-Perten as a measure of the degree of alpha-amylase activity in grain and flour, ICC Standard No. 107/1, International Association for Cereal Science and Technology, Vienna, Austria (1995).
16. Determination of wet gluten quantity and quality (Gluten Index ac. to Perten) of whole wheat meal and wheat flour

- (*Triticum aestivum*), ICC Standard No. 155, International Association for Cereal Science and Technology, Vienna, Austria (1994).
17. Method for using the Brabender Farinograph, ICC Standard No. 115/1, International Association for Cereal Science and Technology, Vienna, Austria (1992).
  18. Method for using the Brabender Extensograph, ICC Standard No. 114/1, International Association for Cereal Science and Technology, Vienna, Austria (1992).
  19. Method for test baking of wheat flours, ICC Standard No. 131, International Association for Cereal Science and Technology, Vienna, Austria (1980).
  20. H. Wieser, S. Antes, W. Selmeier, Quantitative determination of gluten protein types in wheat flour by reversed-phase high-performance liquid chromatography, *Cereal Chem.* 75 (1998) 644–650.
  21. STATISTICA (Data Analysis Software System), v.7.1., StatSoft, Inc., USA (2006) ([www.statsoft.com](http://www.statsoft.com)).
  22. D. Magdić, D. Horvat, Z. Jurković, R. Sudar, Ž. Kurtanjek, Chemometric analysis of high molecular mass glutenin subunits and image data of bread crumb structure from Croatian wheat cultivars, *Food Technol. Biotechnol.* 40 (2002) 331–341.