70. Tagung der Vereinigung der Pflanzenzüchter und Saatgutkaufleute Österreichs
25.-27. November 2019, HBLFA Raumberg-Gumpenstein, Irdning, Österreich

Loaf volume - Is it predictable? An improved method to predict loaf volume in winter wheat by regression models

Friedrich LAIDIG[1], Hans-Peter PIEPHO[1], Alexandra HÜSKEN[2], Jens BEGEMANN[2], Dirk RENTEL[3], Thomas DROBEK[3], Uwe MEYER[3]

[1] Biostatistics Unit, Institute of Crop Science, University of Hohenheim, Fruwirthstr. 23, 70599 Stuttgart, Germany

[2] Department of Safety and Quality of Cereals, Max-Rubner-Institute, Federal Institute of Nutrition and Food (MRI), Schuetzenberg 12, 32756 Detmold, Germany

[3] Bundessortenamt, Osterfelddamm 80, 30627 Hannover, Germany

(✉) Friedrich.Laidig@uni-hohenheim.de

Crude protein content (PC) and Zeleny-sedimentation value (SV) are important indirect traits to evaluate baking quality of wheat. Loaf volume (LV), determined by the Rapid-Mix-Test (RMT), is used to describe the baking ability of wheat varieties. However, the routine application of such baking tests is time consuming and laborious. Therefore, in German official variety trials (VCU), LV is calculated by the Bolling formula based on the indirect traits PC and SV in the framework of the special harvest and quality assessment (*Besondere Ernte- und Qualitätsermittlung (BEE) des BMEL*). The formula was introduced by Bolling in 1969. Two regression equations are used: (a) LV = 410 – 10 × PC + 3 × SV for varieties of the E and A quality groups, and LV = 306 + 17 × PC + 3 × SV for B-varieties (Bolling 1969). It could be shown that during the last 50 years breeding resulted in a change of the protein quality of varieties in a way that the prediction accuracy of the Bolling formulas decreased continuously. In a comparison of the LV predicted for BEE data by the Bolling formulas with LV assessed by the RMT from VCU trials, we found that the predicted LV of varieties was considerably overestimated as compared to the realized LV in VCU trials (Laidig et al. 2017).

The analysis to find an improved method for the prediction of LV was based on two data sets (VCU trials and mill samples) including only registered varieties of quality groups E, A and B. Available were the RMT-LV, PC and SV in both data sets covering the years 1987-2016. The VCU data set (n = 9253) provides a solid basis for a robust prediction model comprising a broad spectrum of varieties and environmental conditions congruent to the BEE samples. The mill samples (n = 3931) arise from on-farm grain lots provided by mills. In a first step, we applied to both data sets a simple linear regression model with LV as dependant, and PC and SV as independent variables for each year and quality group separately. Then we plotted the estimated regression coefficients against the years. The plots indicated negative time trends for the intercept

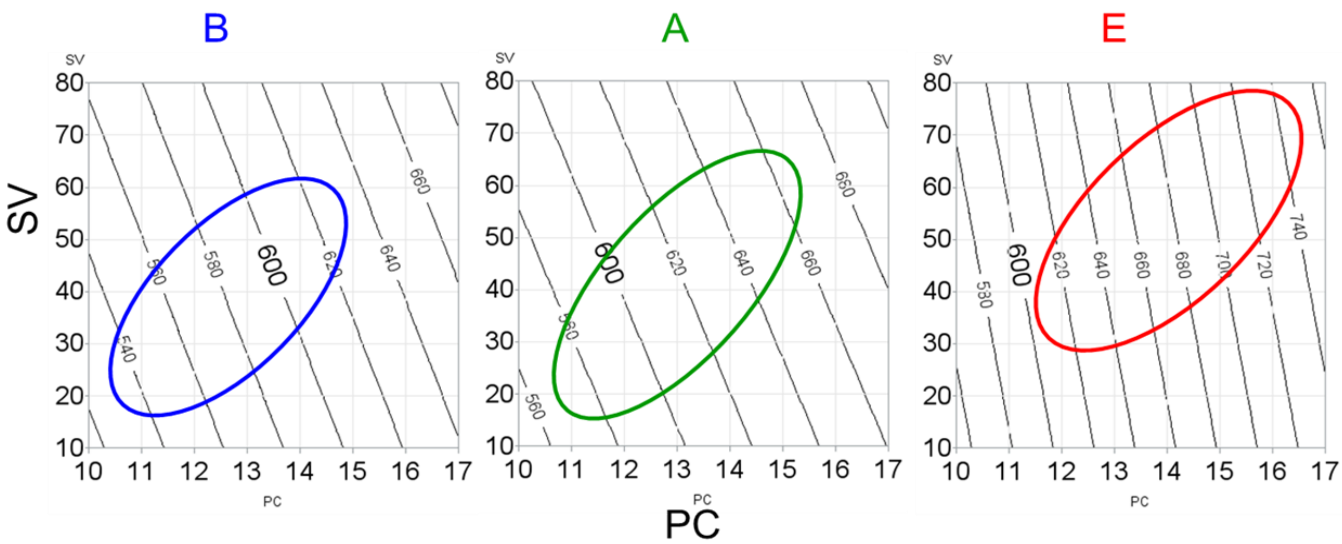

Figure 1: Contour plots predicting loaf volume (LV, mL) for year 2018 with regression equation Opt2q estimated from VCU data 1998-2017 for varieties of B, A and E quality. A contour line indicates the combination of protein content (PC) and Zeleny sedimentation value (SV) predicting the same LV. The confidence ellipses include about 95% of the samples.

Laidig F, Piepho HP, Hüsken A, Begemann J, Rentel D, Drobek T, Meyer U (2020) Loaf volume - Is it predictable? An improved method to predict loaf volume in winter wheat by regression models. In: Vereinigung der Pflanzenzüchter und Saatgutkaufleute Österreichs (Ed), 70. Jahrestagung 2019, 25-27 November, Raumberg-Gumpenstein, pp 25-26. BOKU-University of Natural Resources and Life Sciences, Vienna, Austria. ISBN-13: 978-3-900932-73-2

and the regression coefficient for SV, but a positive trend for PC. The plots further showed a considerable deviation of the year-wise coefficients from the corresponding coefficients of the Bolling formulas. We considered the time trends when generating regression models to be tested. The predictive power of four model options was compared with the Bolling formulas (Opt0q).

For option 1 (Optx1) a single regression was calculated not separately for quality groups and no time trends were taken into account. For option 2 (Opt2x) regression was not calculated separately, but time trend was taken into account. For option 3 (Opt1q) regression was calculated separately for quality groups, but time trend was not taken into account. Finally, for option 4 (Opt2q) regression was calculated separately and time trend was taken into account.

The regression equations obtained by the four model options and given by Bolling's formulas were subject to four independent evaluation schemes. For two direct prediction schemes, samples from VCU/mills were predicted by equations derived from VCU/mill data set, and for two cross prediction schemes, samples mill/VCU were predicted by equations derived from VCU/mill data set. For each evaluation scheme 10 cycles were carried out by calculating the regression equations from data including 20 years to predict LV for the 21st year. For the 10th cycle, for example, the regression equations were estimated from years 1996-2015 to predict LV for samples from year 2016. The prediction accuracy was measured by the root of the mean squared deviations of the observed vs. the predicted LV (sMSD). The average sMSD across the four evaluation schemes was 76.0 for the Opt0q (Bolling formulas), 46.7 for Opt1x, 43.5 for Opt2x, 44.5 for Opt1q and 40.3 for Opt2q. The result indicates that Opt2q predicts LV with the highest accuracy, hence, it is best to take the time trend in the regression coefficients into account and predict the samples separately for each quality group. The results reported by Laidig et al. (2018) further showed that Opt2q provided the most robust predictions across quality groups. When the regression estimates from Opt2q of the most recent 20 years of VCU data are used, then two to three new years of BEE samples can be predicted.

In Figure 1, contour plots are shown for predicting LV for year 2018 with regression equation Opt2q estimated from VCU data 1998-2017. A contour line indicates the combination of PC (x-axis) and SV (y-axis) which predict the same LV. Within the confidence ellipses, about 95% of the samples are located. The three plots reveal that LV is increasing and that the contour lines are getting steeper from B- to E-varieties. For E-varieties, LV is mainly determined by PC, whereas for B-varieties both PC and SV are determining LV.

This study has shown that the model Opt2q improves the prediction accuracy of LV considerably, and that the procedure can be updated by using VCU results and can easily be applied. In the future, this procedure will be applied to predict BEE samples and it can also be used by the complete processing chain for baking wheat.

Keywords

Baking quality · protein content · regression analysis · *Triticum aestivum* · Zeleny sedimentation

References

Bolling H (1969) Wissenschaftliche Grundlage für eine Gradierung von Weizen. Mühle+Mischfuttertechnik 106: 799-801.

Laidig F, Piepho HP, Rentel D, Drobek T, Meyer U, Huesken A (2017) Breeding progress, environmental variation and correlation of winter wheat yield and quality traits in German official variety trials and on-farm during 1983-2014. Theor Appl Genet 130: 223-245. DOI: 10.1007/s00122-016-2810-3

Laidig F, Piepho HP, Hüsken A, Begemann J, Rentel D, *et al*. (2018) Predicting loaf volume for winter wheat by linear regression models based on protein concentration and sedimentation value using samples from VCU trials and mills. J Cereal Sci 84: 132-141. DOI: 10.1016/j.jcs.2018.07.015