59 **EVALUATION OF CRITICAL AGRO-METEOROLOGICAL PARAMETERS AFFECTING WHEAT QUALITY ON THE CANADIAN PRAIRIES USING PARTIAL LEAST SQUARES (PLS) REGRESSION**

Manasah S. Mkhabela[1]*, Paul R. Bullock[1] and Harry Sapirstein[2]
[1]Department of Soil Science, [2]Department of Food Science
University of Manitoba, Winnipeg, MB, R3T 2N2

1. INTRODUCTION

Spring wheat is one of the major crops grown on the Canadian Prairies, of which, about 80% is exported (Qian et al., 2009). However, wheat quality fluctuation in the Prairies is a major issue for both local and international wheat customers/buyers with expectations of consistent quality to meet the needs of the baking industry. Wheat dough quality is very important in baking industries for producing quality products according to consumers demand. Besides genetics and management, the environment, especially the growing season weather, has a significant impact on wheat quality (Finlay et al., 2007). The Canadian Prairies experience a wide range of weather conditions within and between growing seasons. Thus, the overall objective of this study was to identify agro-meteorological factors that impact spring wheat quality in western Canada using partial least squares (PLS) regression. PLS regression is particularly suited when the matrix of predictors has more variables than observations and there is multicollinearity among the predictor variables (Abdi, 2010; Liu et al., 2013).

2. MATERIALS AND METHODS

2.1 Experimental location and design

Crop and weather data collected from small-plot experiments (Figure 1) during the 2003 through 2006 growing seasons from Carman and Winnipeg in Manitoba and Regina, Swift Current and Melfort in Saskatchewan were used in the study. The experimental layout at all sites was a randomised complete block (RCB) design with three replicates (see Finlay et al., 2007 for details).

2.2 Data Analysis

Fifty three (53) agro-meteorological indices (predictor variables) categorised into (i) Water Supply, (ii) Water Demand, (iii) Water Balance and (iv) Water Use were derived from the weather data (Mkhabela et al., 2010) and used in the partial least squares (PLS) regression. Wheat quality characteristics (response variables) including (i) Grain Protein Content (GPC), (ii) Farinograph Absorption (FarAB), (iii) Dough Development Time (DDT) and (iv) Loaf-Volume (LVol) for two varieties (AC Barrie and Superb) were analysed using standard laboratory methods and used in the PLS regression.

**Corresponding author address*: Department of Soil Science, University of Manitoba, Winnipeg, Manitoba, R3T 2N2, Canada. Phone: +1 204-474-9155; Fax: +1 204-474-7642
Email: Manasah.Mkhabela@umanitoba.ca

The number of PLS latent variables (components) was selected using a cross-validation method (Mevik and Wehrens, 2007). Quality data for AC Barrie were used to develop the PLS models (training or calibration set), which were in-turn used to predict wheat quality characteristics for Superb (test or validation set). A prior t-test analysis had shown that there was no significant difference (p>0.05) between the quality characteristics of the two cultivars, except for Farinograph Absorption (Table 1). Variable importance of projection (VIP) was used to select the most influential predictor variables for each quality characteristic (Liu et al., 2013). VIP values (typically set at >0.8) describe the relative contribution of the predictor variables to the PLS latent variables. To avoid model over-fitting problems, predictor variables with VIP scores ≥1.3 were selected as the most influential in this study. Statistical Analysis and Model fitting were performed using R statistical packages (Mevik and Wehrens, 2007; Abdi, 2010).

Figure 1: Wheat small-plot experiments at different growth stages and weather recording instruments.

3. RESULTS AND DISCUSSION

Table 1: Summary statistics of quality characteristics for cultivars AC Barrie and Superb.

	Protein Content (%)	Farinograph Absorption (%)	Dough Development Time (minutes)	Loaf Volume (cc)
Average (mean):				
AC Barrie	14.54	60.69	5.13	961.44
Superb	14.06	62.48	5.29	1013.72
p-value	0.44	0.03	0.84	0.18
Standard Deviation:				
AC Barrie	1.67	2.04	2.07	123.67
Superb	1.71	2.34	2.23	76.13
Coefficient of Variation:				
AC Barrie	11.46	3.37	40.44	12.86
Superb	12.14	3.75	42.11	7.51
Range (Min & Max):				
AC Barrie	10.80-16.70	57.70-65.13	1.65-9.33	666.67-1145.83
Superb	10.47-16.83	59.20-67.80	2.20-9.33	833.33-1116.67

Three separate 3-latent-variable PLS models explained 83% of the variability in wheat GPC and 86% of the variability in the predictor variables; 80% of the variability in DDT and 85% of the variability in the predictor variables; 69% of variability in LVol and 85% of the variability in the predictor variables. Meanwhile, a 4-latent-variable PLS model explained 79% of the variability in FarAB and 89% of the variability in the predictor variables. Six, 4, 6 and 7 predictor variables were identified as the most influential factors (VIP≥1.3) affecting PC, FarAB, DDT and LVol variability, respectively.

When the developed models were used to predict quality characteristics for the variety Superb, the predicted values correlated very well with the observed values with R^2 values of 0.96 ($p<0.001$) for GPC, 0.75 ($p<0.001$) for DDT, 0.69 ($p<0.001$) for LVol and 0.84 ($p<0.001$) for FarAB. This indicates that the models explained from 69% to 96% of the variability in the four quality characteristics. Figures 2 and 3 show the linear relationship between observed and predicted grain protein content and the linear relationship between observed and predicted farinograph absorption, respectively. A t-test showed that the overall average (mean) of the predicted values were not significantly different ($p>0.05$) from the overall average (mean) of the observed values for all the wheat quality characteristics except for FarAB, which was generally under- predicted (Figure 3). The root mean square error (RMSE) values were 0.32 for GPC, 0.85 for DDT, 59.15 LVol, and 0.71 for FarAB.

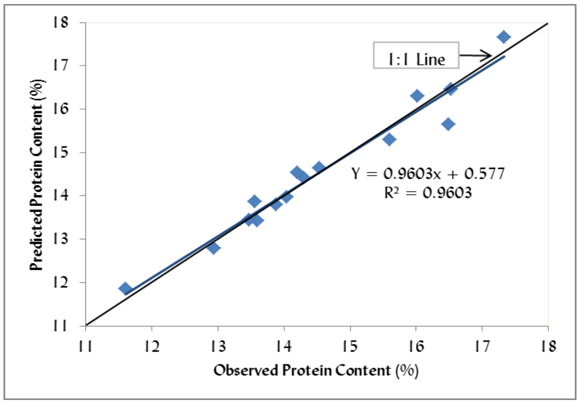

Figure 2: Linear relationship between observed grain protein content and predicted grain protein content for the cultivar Superb.

4. CONCLUSIONS

Sixty nine to 83% of the variability in the four wheat quality characteristics studied was explained by either a 3- or 4-latent variable PLS model. The developed models were able to predict all quality characteristics for the variety Superb with reasonable accuracy, except for FarAB, which was generally under predicted. Overall predicted quality characteristics values were not significantly different ($p>0.05$) from observed values. The models will be further tested and validated using data that is currently being collected across the Canadian Prairies in a project that started during the 2015 cropping season.

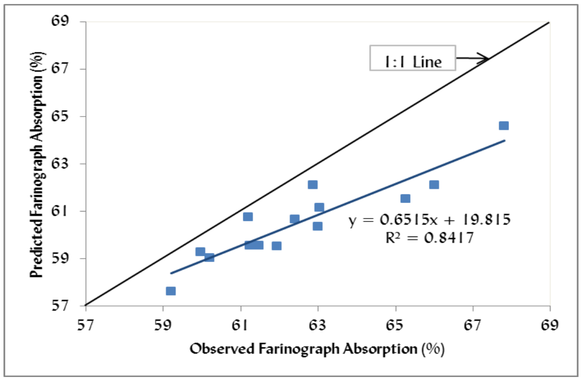

Figure 3: Linear relationship between observed farinograph absorption and predicted farinograph absorption for the cultivar Superb.

ACKNOWLEDGMENTS

Funding was provided by the Natural Science and Engineering Research Council (NSERC) Strategic Grant and a Canadian Wheat Board (CWB) graduate fellowship. Assistance in field trials and grain quality analysis was provided by staff from Agriculture and Agri-Food Canada (AAFC), Canadian Grain Commission (CGC) and the University of Manitoba.

REFERENCES

Abdi, H., 2010: Partial least squares regression and projection on latent structure regression (PLS Regression). *WIREs Comp. Stat*. 2, 97-106.

Finlay, G.J., Bullock, P.R., Sapirstein, H.D. Naeem, H.A., Hussain, A., Angadi, S.V. and DePauw, R.M., 2007: Genotypic and environmental variation in grain, flour, dough and bread-making characteristics of western Canadian spring wheat. *Can. J. Plant Sci.*, 87, 679–690.

Liu, K., Elliott, J.A., Lobb, D.A., Flaten, D.N. and Yarotski, J. 2013: Critical factors affecting field-scale losses of nitrogen and phosphorus in spring snowmelt runoff in the Canadian prairies. *J. Environ. Qual.,* 42, 484–496.

Mevik, B.H. and Wehrens, R., 2007: The pls Package: principal component and partial least squares regression in R. *J. Stat. Softw.*, 18 (2), http://www.jstatsoft.org/

Mkhabela, M., Bullock, P., Gervais, M., Finlay, G. and Sapirstein, H. 2010. Assessing indicators of agricultural drought impacts on spring wheat yield and quality on the Canadian prairies. *Agric. Forest Meteorol.*, 150, 399–410.

Qian, B., De Jong, R., Warren, R., Chipanshi, A., Hill, H., 2009: Statistical spring wheat yield forecasting for the Canadian Prairie Provinces. *Agric. Forest Meteorol.*, 149, 1022–1031.