Validation of Near Infrared Spectroscopy (NIRS) Milling Index Method

Dr Marena Manley
Department of Food Science
Faculty of AgriSciences
Stellenbosch University
Private Bag X1
Matieland 7602

Tel: 021 808 3511
Fax: 021 808 3510
Email: mman@sun.ac.za

June 2007

Introduction and methods used

The previously developed Roff Milling Index (MI) has been used as reference method to develop a near infrared spectroscopy (NIRS) prediction model for the rapid prediction of the MI of whole maize kernels. The MI of an independent validation set (n = 87) has been determined using the reference method as well as the NIRS prediction models. The consequent results, as supplied by Jonathan Wong, will be evaluated and reported on in this report.

Statistical and data analysis

Analysis of variance (ANOVA) has been performed to compare measurements as performed by the Roff and NIRS MI methods, respectively, to determine absolute differences. Additionally the intra-class correlation coefficients (ICC) have been determined as the ICC agreement that correlates measurements with each other while also taking into account the differences in absolute values of the respective measurements and the ICC consistency that correlates measurements with each other, but ignores absolute differences. The standard error of prediction (SEP) and bias has been calculated to evaluate the NIRS MI calibration.

Results and Discussion

The MI results obtained by means of the reference method as well as NIRS are listed in the Addendum. The correlation plot of the Roff MI measurements vs the NIRS MI measurements are shown in Figure 1A and Figure 1B depicted the plot showing the significant difference (p = 0.00) between the Roff MI and NIRS MI measurements. From the ICC consistency (0.847) it is clear that there is a correlation between the obtained results, but the ICC agreement (0.345) indicates that there are differences in the absolute values of the two respective measurements. This is also evident from the clear bias observed in Figure 1A.

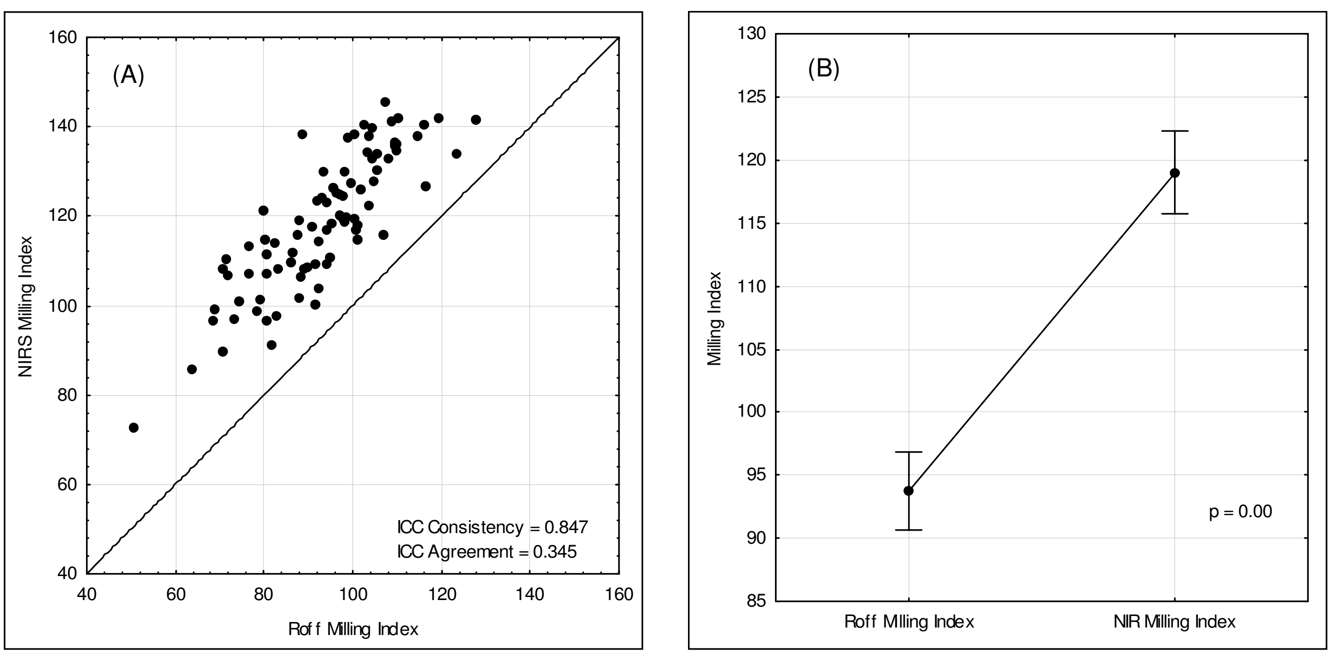

Figure 1 (A) Intra-class correlation plot of the Roff Milling Index values vs the NIRS Milling Index values and (B) plot showing significant difference between the Roff MI and NIRS MI measurements.

The residiual (difference between reference and NIRS MI results) can be used to calculate the standard error of prediction (SEP) according to the equations as listed in the Addendum (Table 2). This resulted in a SEP MI value of 26.545. The SEP can statistically be interpreted as the expected error, with a probability of 68%, to be within an interval of ±1SEP and with a probability of 95% within an interval of ±2SEP. This means that 68% of all the results can be expected to have an error of 8.196 or 95% of all the results, an error of 16.392 (if bias correction has been applied). This value should be as small as possible or as close as possible to the error of laboratory (SEL). The SEL can be calculated if duplicate reference measurements have been made. There is thus no rule as to how small the SEP should be and it is for the user to determine the acceptable error for a specific application. Subsequently, the bias was calculated as being -25.263. After correcting for bias (-25.263), as observed in Figure 1A, the SEP could be reduced to 8.196. This is only to illustrate that if the bias can be removed the accuracy of the method can be improved. The source of the bias error should therefore first be investigated and removed.

Guidelines for interpretation of the R^2 in terms of NIR calibrations are given in the Addendum (Table 3). The R^2 obtained for this validation (0.85) is therefore suitable, but should be treated with caution.

The obtained results can also be expressed in terms of the RPD, as suggested by Williams (2001) which is the ratio between the SEP to the standard deviation (SD) of the validation set. According to the interpretation guidelines in Table 4 the RPD of 1.75 do not suggest this calibration to be suitable. However, due to the complexity of the Roff Milling Index determination I would not use the RPD value as the deciding factor in this case.

Conclusion

An acceptable R^2 has been obtained during the validation of the current NIRS method for predicting the milling index in whole maize samples, however, the large observed bias resulted in a low ICC agreement and the subsequent low accuracy (high SEP). This is a cause for concern as due to this large bias the current NIR measurements are not useful. It is therefore advised that the source of the bias error is investigated and understood and a method being developed that would greatly reduce this error. If all else fails, it will be necessary to require a small set of data to be measured in advance to determine the bias correction before any future, unknown samples are measured. To improve the robustness of the calibration the 87 samples used for this validation could be included in the previous calibration set and the calibration recalculated.

References

Williams, P.C. (2001). Implementation of near-infrared technology. In: *Near-infrared Technology in the Agricultural and Food Industries*, 2nd ed. (edited by P. Williams & K. Norris). Pp. 145-169. St. Paul, USA: American Association of Cereal Chemists.

Addendum A

Table 1 Roff Milling Index (reference) and NIRS Milling Index data

Sample number	Roff Milling Index	NIRS Milling Index	Residual
1	89.71	108.67	-18.97
2	80.70	96.46	-15.76
3	87.90	119.04	-31.15
4	91.75	109.10	-17.35
5	68.80	99.17	-30.38
6	68.66	96.59	-27.92
7	105.49	133.99	-28.51
8	98.41	119.62	-21.21
9	89.09	108.01	-18.91
10	73.42	96.87	-23.45
11	71.37	110.40	-39.02
12	63.68	85.81	-22.13
13	96.39	125.32	-28.93
14	94.25	122.96	-28.71
15	78.59	98.65	-20.07
16	108.24	132.76	-24.52
17	98.33	129.99	-31.66
18	105.61	130.23	-24.63
19	116.40	126.60	-10.20
20	82.63	114.03	-31.40
21	95.20	118.14	-22.93
22	101.00	114.77	-13.77
23	74.37	100.92	-26.55
24	91.48	100.28	-8.80
25	81.67	91.21	-9.55
26	87.70	115.85	-28.15
27	97.70	124.57	-26.87
28	70.87	108.01	-37.15
29	100.97	117.89	-16.92
30	76.44	107.02	-30.58
31	93.57	129.94	-36.38
32	85.95	109.58	-23.62
33	92.30	103.99	-11.69
34	83.23	108.29	-25.06
36	97.02	120.15	-23.13
38	123.44	133.93	-10.49
39	100.92	116.97	-16.05
40	109.44	135.59	-26.15
41	95.80	126.29	-30.48
42	88.76	138.16	-49.40
43	109.74	136.17	-26.43
44	109.90	134.59	-24.69
45	109.85	136.07	-26.22
46	110.08	141.98	-31.90
47	102.71	140.36	-37.65
48	107.29	145.47	-38.19
49	100.46	138.14	-37.68
50	104.39	132.72	-28.34
51	99.02	137.42	-38.40
52	127.99	141.67	-13.68
53	104.85	127.82	-22.97

Table 1 …/ continue

Sample number	Roff Milling Index	NIRS Milling Index	Residual
54	108.69	141.03	-32.34
55	119.23	142.04	-22.81
56	101.94	125.94	-24.00
57	103.60	137.96	-34.36
58	116.23	140.28	-24.04
59	114.48	137.95	-23.47
72	95.05	110.60	-15.55
73	50.48	72.70	-22.22
74	80.52	111.50	-30.98
75	80.42	114.56	-34.14
76	70.91	89.66	-18.75
77	94.20	117.04	-22.84
78	92.28	114.16	-21.88
79	93.23	124.12	-30.89
80	100.43	119.27	-18.84
81	71.73	106.84	-35.10
82	86.43	111.84	-25.41
83	104.42	139.87	-35.46
84	87.92	101.70	-13.78
85	106.86	115.81	-8.95
86	88.18	106.28	-18.10
87	97.85	119.52	-21.67
88	97.18	124.80	-27.62
89	79.33	101.37	-22.04
90	80.05	121.07	-41.02
91	82.72	97.68	-14.95
92	103.28	134.36	-31.08
93	103.62	122.30	-18.68
94	98.06	118.85	-20.79
95	90.77	117.71	-26.94
96	109.58	136.30	-26.72
97	94.00	109.19	-15.20
98	99.71	127.39	-27.67
99	80.45	107.07	-26.61
100	76.54	113.33	-36.78
101	91.94	123.39	-31.45
Mean	**93.75**	**119.01**	
SD	**14.30**	**15.29**	
Minimum	**50.48**	**72.70**	
Maximum	**127.99**	**145.47**	

Table 2 Equations for statistical calculations.

Statistic	Equation	Recommendations
SEP[a]	$\sqrt{\frac{\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2}{n}}$	As small as possible or close as possible to SEL[b] value
SEP(C)[c]	$\sqrt{\frac{\sum_{i=1}^{n}\left(y_i - \hat{y}_i - BIAS\right)^2}{n-1}}$	As small as possible or close as possible to SEL value
BIAS[d]	$\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)$	As close to zero as possible
RPD[e]	$\frac{SD_{\hat{y}}}{SEP}$	See Table 9
SEL[b]	$\sqrt{\frac{\sum\left(y_1 - y_2\right)^2}{2n}}$	As small as possible

[a] Standard error of prediction

[b] Standard error of laboratory

[c] Standard error of prediction (corrected for bias)

[d] Bias of the validation set

[e] Ratio of standard error of prediction to standard deviation

y = reference value

$\hat{y}$ = predicted value

y_i = reference value for the i^{th} sample

$\hat{y}_i$ = NIR predicted values for the i^{th} sample

y_1 and y_2 = duplicate reference values

n = number of samples

t = number of terms in the model

Table 3 Guidelines for interpretation of R^2 (adapted from Williams, 2001).

R^2 value	Interpretation
Up to 0.25	Cannot use in NIRS calibration
0.26-0.49	Poor correlation. Investigation is necessarily
0.50-0.64	Can be used for rough screening. More than 50% of variance in *y* (NIR data) accounted for by *x* (reference data)
0.65-0.81	Can be used for screening and some approximate calibrations
0.82-0.90	Can be used in most applications but with caution. More research is necessary
0.91-0.96	Can be used in most applications, including quality assurance
0.97+	Can be used in any applications

Table 4 Guidelines for interpretation of RPD (Williams, 2001).

RPD value	Classification	Application
0.0-2.3	Very poor	Not recommended
2.4-3.0	Poor	Very rough screening
3.1-4.9	Fair	Screening
5.0-6.4	Good	Quality control
6.5-6.8	Very good	Process control
8.1+	Excellent	Any application