Calibration and field standardization of Tecator Infratec analyzers for corn and soybeans

Glen R. Rippke, Connie L. Hardy and Charles R. Hurburgh, Jr

Iowa State University, Ames, IA 50011, USA.

Thomas J. Brumm

MBS Inc., Story City, IA, USA.

Introduction

Whole-grain near infrared (NIR) transmittance or reflectance analyzers are a relatively new use of a technology developed in the early 1970s. The rapid, no-sample-preparation analysis is popular with the grain industry and supporting government agencies. The Tecator Infratec, a monochromator-based NIR transmittance analyzer is one such instrument (Perstorp Analytical Inc. Höganäs, Sweden).

Infratec units are approved for official USDA analysis of wheat protein and soybean protein and oil. Infratec accuracy, as measured by standard error of prediction (*SEP*, in percentage points) is reported by Perstorp at 0.19, 0.38 and 0.30 for wheat protein,[1] soybean protein and oil,[2] respectively. Perstorp has also reported *SEP*s in corn of 0.47 for moisture, 0.24 for protein and 0.21 for oil in their MA4 corn calibration.[3]

Infratec calibrations are created on a master instrument and then transferred to like units (slaves). Transferring of calibrations is necessary because not all users have the resources (time, samples and budget) to make their own calibrations but still have the need of the analyzer. Calibration transfer (standardization) is done with a slope and intercept, or bias only, applied to moisture-basis corrected predicted values.[4] Neither the calibration nor the standardization method had been optimized for the diverse grain, varying temperatures and numerous remote test sites found in the United States coarse grain (corn and soybean) markets.

Objective

Our objective was to develop calibration and standardization protocols and derive robust calibrations that:

- can accurately predict corn and soybean composition, and
- are relatively insensitive to sample temperature, and
- are transferable to like instruments.

A further objective was to test these calibrations under actual elevator conditions. The calibrations predicted moisture, protein, oil, starch and density (specific gravity) for corn, and moisture, protein, oil and fiber for soybeans.

Materials and methods

Samples and equipment

The Grain Quality lab has maintained a sample library consisting of a wide variety of corn samples collected each year from 1986 to 1994, and of soybean samples collected each year from 1988 to 1994. The samples were collected over a wide growing region under diverse conditions and with different physical characteristics. The samples were stored in glass jars kept at 3ºC. All the samples from years before 1992 were scanned during the summer of 1992 and had less than 16% moisture content. All these samples had wet chemistry analyses. Fresh harvest samples, with moistures up to 40%, were used in 1992–1994. Each sample was scanned on the master unit (Tecator 1225 Grain Analyzer serial number 65). Ten subsample spectra were averaged to a single sample scan. Each time a calibration scan was collected, a corresponding whole grain oven moisture analysis was done in our oven: 72 hours at 103ºC for corn,[5] and 3 hours at 130ºC for soybeans.[6]

After the 1993 harvest, we pooled all samples with wet chemistry data to further identify those that would be in the actual calibration. The samples scanned in the summer of 1992 (1991 crop and earlier) had previously been used in the calibration of ground-grain reflectance analyzers. The 1992 harvest samples were selected for constituent chemistry by predicted constituent values from the ground grain analyzers. Approximately four samples were scanned and received oven moistures for every one sample retained for wet chemical analysis. In 1993 and thereafter, principal component analysis (PCA) in the calibration program Infratec Calibration Maker[7] was used to identify samples for chemistry, based on their spectral significance.

The samples chosen for chemistry analysis were sent to an independent laboratory (Woodson–Tenent, Des Moines, IA). In corn, the samples were analyzed for moisture,[8] crude protein,[9] crude oil[10] and bulk starch.[11] The soybeans were analyzed for moisture,[6] crude protein,[12] crude oil[13] and crude fiber.[10]

Calibration development

Our calibration procedure consists of two steps: developing a base calibration and adding samples to the base calibration for instrument stabilization and temperature stabilization. The base calibration contains scans collected on the master unit at room temperature and includes the majority of the calibration library. Instrument stabilization scans were collected in two ways—from the master unit with deliberately altered monochromator constants and from other Infratec instruments. The instrument stabilization scans add noise to the calibration and enhance transferability of finished calibrations. Temperature stabilization scans, collected at grain temperatures from –15°C to 45°C, compensate for the varying grain temperatures during the marketing year.

Samples were chosen for the base calibration set by two methods applied consecutively: (i) PCA of the individual spectra and (ii) completion of the combination of constituents matrix (CCM). The CCM is a factorial combination of the constituents that need to be represented in the calibration. Each constituent is scored as high, medium or low for a particular sample. For example, a corn sample could have high moisture, medium protein, medium oil, and high starch, while another sample could have medium moisture, low protein, medium oil and high starch. The respective scores for these would be high, medium, medium, high (HMMH) and medium, low, medium, high (MLMH). The three ranges were initially determined by the formula:

$$< \overline{X} - 0.5\sigma = \text{LOW}$$
$$\text{between} \rightarrow \overline{X} \pm 0.5\sigma = \text{MEDIUM}$$
$$> \overline{X} + 0.5\sigma = \text{HIGH}$$

Table 1. Combination of constituents matrices for corn and soybeans.

	Factor	Range		
		Low	Medium	High
Corn	Moisture	<12.0%	12.0–22.0%	>22.0%
	Protein	<7.0%	7.0–9.0%	>9.0%
	Oil	<3.5%	3.5–5.0%	>5.0%
	Starch	<58.0%	58.0–62.0%	>62.0%
	Density (g cm^{-3})	<1.240	1.240–1.300	>1.300
Soybeans	Moisture	<9.0%	9.0–13.0%	>13.0%
	Protein	<33.0%	33.0–37.0%	>37.0%
	Oil	<16.5%	16.5–19.5%	>19.5%
	Fiber	<4.0%	4.0–5.5%	>5.5%

Corn protein, oil, starch and density were selected on a 15% moisture basis.
Soybean protein, oil and fiber were selected on a 13.0% moisture basis.

where $\overline{X}$ is the constituent mean by reference chemistry on a constant moisture basis and σ is the long term standard deviation of the constituent values.

The number of cells (n_c) in the CCM will be $n_c = 3^f$; where f = the number of constituents measured.

The CCM ranges were refined after the 1993 calibration to reflect values indicative of high, medium and low in actual market practice, without regard for the variance in the particular sample set evaluated. Table 1 gives our current CCM.

Some combinations (e.g. HHHH) are biologically or practically impossible. Normally 70–80% of the cells had one or more calibration sample(s) in them after the PCA. PCA selection, by sample spectra, will fit most, but not all, possible cells. A CCM evaluation after spectral sample selection was necessary to fill as many cells in the CCM as possible. We filled about 90% of the CCM cells in the base calibration. Because most regression-based calibration algorithms (e.g. partial least squares) reference the spectral or chemical values to their respective means, it is important that the reference mean be close to that expected in practice.

The first phase of the calibration process, creation of a base calibration for each constituent, used only "normal scans" (for 1993, 103 soybean and 120 corn) collected at room temperature, and with the proper "O and P" monochromator specific spectral settings determined at the factory. All constituent calibrations were created on an as-is moisture basis (at the moisture at the time of scan collection). The results were examined for spectra with large "influence" on the calculated calibration constants and for samples with chemical residual outliers (poor chemistry). Spectra with large "influence" have unusual PCA scores and in a calibration can overpower normal samples to the point where they distort the predictions. Spectra with large "influence" were deleted, unless they had a pre-determined importance, as representing material that would be tested frequently. The samples with chemical outliers were re-analyzed, or deleted, if the sample was

Table 2. Distribution of samples in the 1994 Infratec calibrations.

	Corn				Soybeans		
	High (%)	Medium (%)	Low (%)		High (%)	Medium (%)	Low (%)
Moisture	30	32	38	Moisture	36	16	48
Protein	25	52	23	Protein	25	65	10
Oil	27	27	36	Oil	10	64	26
Starch	11	56	33	Fiber	21	75	4

Table 3. Statistics for the 1993 Infratec corn calibrations.

Statistic		Constituent			
		Moisture	Protein (as-is)	Oil (as-is)	Starch (as-is)
Calibration: CN000093					
	n	120	120	120	120
	R^2 (%)	99.6	98.0	96.8	97.4
Std error of calibration (% pts)		0.53	0.33	0.31	1.20
	Maximum (%)	40.0	12.5	8.6	67.0
	Minimum (%)	7.0	5.0	2.5	43.0
Validation:					
	n	71	71	71	71
	R^2 (%)	99.1	98.0	85.1	92.5
Std error of prediction (% pts)		0.50	0.31	0.37	1.27
	Maximum (%)	28.0	13.0	6.3	67.0
	Minimum (%)	8.0	6.0	2.7	50.0

Corn classes included: yellow dent, flint and white.

Table 4. Statistics for the 1993 Infratec soybean calibrations.

Statistic		Constituent			
		Moisture	Protein (as-is)	Oil (as-is)	Fiber (as-is)
Calibration: SB000093					
	n	103	85	88	80
	R^2 (%)	99.0	98.4	97.9	93.9
Std error of calibration (% pts)		0.46	0.54	0.34	0.20
	Maximum (%)	19.0	46.0	23.0	6.5
	Minimum (%)	5.6	28.0	14.0	4.0
Validation:					
	n	282	203	203	194
	R^2 (%)	98.4	97.0	93.0	71.0
Std error of prediction (% pts)		0.47	0.53	0.56	0.30
	Maximum (%)	18.0	43.0	23.0	6.4
	Minimum (%)	5.0	32.0	14.0	4.5

not unique enough to warrant repeat chemistry. The base calibration was then validated on the remaining samples with known constituent values (wet chemistry analyses) that were not selected for the base calibration. The critical evaluation statistics were the standard error of prediction on the validation set and the bias and/or slope on that set.

The instrument stabilization set contained samples with a wide variety of constituent values. According to the manufacturer's recommendation, the master instruments' optical settings (mono-chromator constants) are slightly changed to add noise to the calibration. In 1994, we also added scans from two other Infratec units, to complement the simulated noise from the master. Master-predicted values from the base calibration were used as reference data. These samples were not a part of our base calibration chemistry library.

Sample scans were collected on selected samples at three temperatures (5°C, 25°C and 45°C) for use in the temperature stabilization and validation sets. In 1994, we included samples at –15°C. However, corn at –15°C and over 20% moisture was not used in either the calibration or validation sets because they distorted the calibrations and gave unreliable validation statistics. Master-predicted values from the base calibration at room temperature were used as reference values. These samples were not part of the base calibration chemistry library.

The instrument stabilization samples and temperature stabilization samples were scanned under "normal" conditions. The base calibration was used to predict the constituent values of the "normal" stabilization and temperature samples. These predicted values were used as the reference

chemistry values to replace laboratory error (an unknown) for these samples with the calibration error (known). This also prevented the stabilization sample sets from overwhelming the calibration. The base calibration is the main predictor in the finished calibrations, while the stabilization sets are to enhance the transferability and desensitize the calibration to temperature influences. We do not want the stabilization sets to do the predicting of samples. The instrument and temperature files were combined with the base calibration file and the final calibration was developed. The finished calibration was validated with the original validation file and in 1994 against a temperature validation sample set also.

The finished calibrations were installed in the master and in five slave units. The slave instruments were adjusted to the master unit so all units give consistent accurate results by a process called standardization. Slave-predicted values for the standardization sample sets (30 samples per grain) were regressed against the master predicted values. The standardization set is chosen to cover the constituent ranges and to be good predictor samples with replicate chemistry values. A slope and intercept, or a bias only, was applied to the slaves' predicted values for each constituent, as needed. Slope was not used unless the slope coefficient was significant at the 95% probability level. Moisture calibrations were standardized first, with the corrected moisture then used to make the moisture-basis adjustment for the constituents. In the Infratec operating system, instrument slope and intercept is the last calculation made after the moisture correction.

Field test and update

The five slave units were operated at various elevators and seed companies in the fall of 1994. Samples were exchanged to measure their bias and variability relative to the master unit under field conditions. Elevator operators analyzed all the comparison samples themselves without supervision.

Table 5. Transferability statistics for the 1993 Infratec corn and soybean calibrations.(5 units).

Corn				
	Moisture	Protein	Oil	Starch
Average bias[a] (% pts)	–0.01	0.01	0.01	–0.28
Average *SDD*[b] (% pts)	0.19	0.16	0.20	0.40
Range of biases (% pts)	–0.02 – 0.00	–0.02 – 0.08	–0.07 – 0.10	–0.82 – 0.21
Soybeans				
	Moisture	Protein	Oil	Fiber
Average bias[a] (% pts)	–0.02	–0.05	0.09	–0.14
Average *SDD*[b] (% pts)	0.20	0.23	0.19	0.26
Range of biases (% pts)	–0.04 – 0.01	–0.16 – 0.06	–0.16 – 0.33	–0.31 – –0.02

[a]Slave minus master, $n = 10$ or more samples per unit.
[b]Standard deviation of differences, slave to master.

Table 6. Statistics for the 1994 Infratec corn calibration.

Statistic		Constituent				
		Moisture[a]	Protein (as-is)	Oil (as-is)	Starch (as-is)	Density (at 15% M)
Calibration: CN000940						
	n	187	153	205	157	162
	R^2 (%)	99.7	97.6	98.2	98.3	71.4
Std error of calibration (% pts)		0.50	0.33	0.29	1.30	0.032 g cm^{-3}
	Maximum (%)	46.0	13.0	9.0	69.0	1.350 g cm^{-3}
	Minimum (%)	6.0	4.5	2.5	35.0	1.131 g cm^{-3}
Validation:						
	n	71	71	71	61	139
	R^2 (%)	99.6	94.5	86.6	84.6	78.7
Std error of prediction (% pts)		0.54	0.32	0.23	1.13	0.020 g cm^{-3}
	Maximum (%)	38.0	12.0	6.5	70.0	1.340 g cm^{-3}
	Minimum (%)	9.0	5.0	2.5	53.0	1.200 g cm^{-3}
Validation - Temperature:						
	n	183	126	126	126	
	R^2 (%)	99.7	91.0	97.9	98.8	N/A
Std error of prediction (% pts)		0.47	0.43	0.24	0.84	
	Maximum (%)	40.0	12.0	9.0	66.0	
	Minimum (%)	9.0	6.0	2.5	36.0	

Classes included: Yellow Dent, Flint, White, Waxy and Red.
[a]Approved for trade by National Institute for Standards and Technology.
N/A—Not available.

In the fall of 1994, 468 new corn and 223 new soybean samples were scanned in the master unit. Many of these were deliberately chosen to be beyond the limits of the 1993 calibration, but those that were within limits provided a fully independent validation.

Table 7. Statistics for the 1994 Infratec soybean calibration.

Statistic		Constituent			
		Moisture[a]	Protein (as-is)	Oil (as-is)	Fiber (as-is)
Calibration: SB000940					
	n	143	177	187	177
	R^2 (%)	99.0	98.9	97.4	93.0
Std error of calibration (% pts)		0.34	0.33	0.34	0.16
	Maximum (%)	19.5	45.0	23.0	6.5
	Minimum (%)	5.1	33.0	13.0	4.0
Validation:					
	n	272	283	250	232
	R^2 (%)	98.9	96.6	95.0	72.4
Std error of prediction (% pts)		0.40	0.57	0.48	0.36
	Maximum (%)	20.0	44.0	23.0	6.5
	Minimum (%)	4.5	32.0	14.0	4.0
Validation - Temperature:					
	n	181	157	157	157
	R^2 (%)	98.8	98.0	98.3	86.0
Std error of prediction (% pts)		0.29	0.42	0.32	0.22
	Maximum (%)	14.5	44.0	23.0	6.5
	Minimum (%)	7.0	33.0	14.0	4.4

[a]Approved for trade by National Institute for Standards and Technology.

In 1994, the temperature range was extended down to –15°C, and several new CCM cells were filled. Specialty corns (e.g. cornnuts, waxy, caribbean flints and reds) were added. A calibration for density (g cm^{-3})[14] was developed and a temperature validation for all constituents was added.

The calibration process was repeated with the new samples included in the pool. The 1994 moisture calibrations were subsequently submitted by Perstorp for National Type Approval Program evaluation.[15]

Results and discussion

Calibration statistics for the 1993 final calibrations (CN93 for corn and SB93 for soybeans) are given in Tables 3 and 4. The standard error of prediction on the validation set is an important evaluation statistic. Our results were similar to, or better than, previous reports while expanding most of the constituent ranges. Infratec moisture calibrations were also significantly more precise at higher moistures than those of the commonly used capacitance moisture meters, as reported by Hurburgh *et al.*[16] and Wu *et al.*[17]

In the process of standardizing the slave units, slopes of 0.90 to 1.10 existed. This demonstrated a clear need for the standardization process. Slopes and intercepts were generally larger (farther from 1.0 and 0.00, respectively) for harder-to-measure factors (oil and starch).

Table 5 gives the results of check testing the slave units in the field. Standard deviations, relative to the master, were approximately 50% of the standard errors of prediction, relative to chemistry. In our opinion, this is acceptable performance for grain trading situations.

The updated 1994 final calibration statistics are given in Tables 6 and 7. These are the calibrations that are approved in the National Type Approval Program (moisture) and licensed for commercial use by MBS, Composition Systems (MBS Inc., Story City, IA). They are available as original equipment corn and soybean calibrations by Perstorp Analytical Inc.

Conclusions

- The step-by-step calibration protocol produced accurate and transferable calibrations over a wide range of constituent values and operational conditions.
- Infratec achieved standard errors of prediction for corn moisture, protein, oil, starch, and density of 0.54%pts, 0.32%pts, 0.23%pts, 1.13%pts and 0.02 g cm^{-3}, respectively. Infratec achieved standard errors of prediction for soybean moisture, protein, oil, and fiber of 0.40%pts, 0.57%pts, 0.48%pts and 0.36%pts.
- The small biases and standard errors of transfer, in both corn and soybeans, demonstrated the transferability of the calibrations.
- The 1994 calibration update extended the ranges for several constituents, while the *SEP*s remained the same or were improved.
- Infratec outperformed capacitance moisture meters over a wider range of moisture for both grains.

References

1. *Infratec calibrations—Wheat, WH3,* Application note AN 95/88. Perstorp Analytical, Sliver Spring, MD (1990a).
2. *Infratec calibrations—Soybeans, US22,* Application note AN 110-20/89. Perstorp Analytical, Sliver Spring, MD (1989).
3. *Infratec calibrations—Maize, MA4,* Application note AN 110-62/90. Perstorp Analytical, Sliver Spring, MD (1990b).
4. *Protein handbook*. Federal Grain Inspection Service, USDA, Washington, DC (1990).
5. "Oven moisture determination", in *Moisture handbook.* Federal Grain Inspection Service, USDA, Washington, DC, Chapter 4 (1986).
6. "Air-oven moisture", AOCS Methods Ac 2-41, *Official Methods of the American Oil Chemists Society,* 4th Ed. AOCS, Champaign, IL (1987 rev.).
7. NIR Systems, Infratec Calibration Maker, V2.3c. Produced for Tecator AB by Camo, Norway, Tecator AB, Höganäs, Sweden (1989).
8. "Either extraction", AOCS Methods Ac 3-44, *Official Methods of the American Oil Chemists Society*, 4th Ed. AOCS, Champaign, IL (1989 rev.).

9. "Kjeldahl nitrogen", AOCS Methods Ba 4d-90, *Official Methods of the American Oil Chemists Society*, 4th Ed. AOCS, Champaign, IL (1992 rev.).

10. "Ether extraction", AOCS Methods Ba 3-38, *Official Methods of the American Oil Chemists Society,* 4th Ed. AOCS, Champaign, IL (1993 rev.).

11. "Feedstuffs analysis—starch", Method CRA G-28, *Official Methods of the Member Companies,* 6th Ed. Corn Refiners Association, Washington, DC (1986 rev.).

12. "Kjeldahl nitrogen", AOCS Methods Ba 4-91. *Official Methods of the American Oil Chemists Society,* 4th Ed. AOCS, Champaign, IL (1992 rev.).

13. "Acid-base digestion", AOCS Methods Ba 6-84. *Official Methods of the American Oil Chemists Society*, 4th Ed. AOCS, Champaign, IL (1989 rev.).

14. *AccuPyc 1330 Pyconometer Operator's Manual*, V2.01, Micromeritics, Norcross, GA (1992).

15. "Grain Moisture Meters", Section 5.56 in *Handbook 44*. National Institute of Standards and Technology, Gaihersburg, MD (1995).

16. C.R. Hurburgh Jr, T.E. Hazen and C.J. Bern, Transactions of the ASAE 28(2), 634 (1985).

17. Y. Wu and C.R. Hurburgh Jr, New technologies for grain moisture measurement", ASAE paper 936065 (1993).