Prediction of Wisconsin Tester Breakage Susceptibility of Corn from Bulk Density and NIRS Measurements of Composition

Juraj Siska
Iowa State University

Charles R. Hurburgh Jr.
Iowa State University, tatry@iastate.edu

Follow this and additional works at: http://lib.dr.iastate.edu/abe_eng_pubs
Part of the Agriculture Commons, and the Bioresource and Agricultural Engineering Commons

The complete bibliographic information for this item can be found at http://lib.dr.iastate.edu/abe_eng_pubs/422. For information on how to cite this item, please visit http://lib.dr.iastate.edu/howtocite.html.
Prediction of Wisconsin Tester Breakage Susceptibility of Corn from Bulk Density and NIRS Measurements of Composition

Abstract
An equation was developed to predict corn breakage susceptibility based on the protein content, oil content, starch content, kernel density, and test weight. Reference values of breakage susceptibility were measured by Wisconsin Breakage Tester. Two statistical techniques were used to design the prediction equation, multiple linear regression (MLR) and principal factor method (PFM).

Keywords
Breakage susceptibility, Corn, Near infrared, Principal factor method, Grades and standards

Disciplines
Agriculture | Bioresource and Agricultural Engineering

Comments
This article is from Transactions of the ASAE 37 (1994): 1577–1582. Posted with permission.
ABSTRACT. An equation was developed to predict corn breakage susceptibility based on the protein content, oil content, starch content, kernel density, and test weight. Reference values of breakage susceptibility were measured by Wisconsin Breakage Tester. Two statistical techniques were used to design the prediction equation, multiple linear regression (MLR) and principal factor method (PFM).

The calibration sample set included 94 market corn samples. This set was used to predict corn breakage susceptibility from protein content, starch content, test weight, and kernel density with MLR. Two sample sets were used for PFM calculations. One sample set of 252 corn samples was used to calculate standardized scoring coefficients of a common factor, which represented protein, oil, starch, and density. The other set, the same calibration sample set from MLR analysis, was used to regress breakage susceptibility against the common factor and test weight.

The calibrations were validated on 36 samples not included in any of the sample sets described. For the MLR analysis, the $R^2$ (validation) was 65%, and the standard error of prediction (SEP) was 1.0 percentage point. The PFM equation gave $R^2$ of 64% and the SEP was 1.0% point. Keywords. Breakage susceptibility, Corn, Near infrared, Principal factor method, Grades and standards.

Breakage susceptibility is defined as the potential for kernel fragmentation or breakage when corn is subjected to impact forces (AACC, 1991a). This is an important property for the grain-handling and processing industry. Excessive broken kernels can reduce the grade of grain, lower the selling price farmers receive, decrease production from dry and wet milling processes, reduce efficiency of drying, and increase potential dust hazards.

There is a need for a simple, rapid method that would allow grain handlers to differentiate lots containing corn with different breakage susceptibility values (Eckhoff, 1989). A number of devices have been developed to measure grain breakage susceptibility. Among the most common devices are the Stein Breakage Tester (SBT) (McGinty, 1970) and the Wisconsin Breakage Tester (WBT) (Singh and Finner, 1983). However, the speed of analysis by both of these devices (ca 2 to 4 min/sample) is not appropriate to grain market requirements. In addition, these devices are not capable of measuring any other property except breakage susceptibility.

The Wisconsin breakage tester was determined to be the most precise among eight breakage susceptibility devices (Watson and Herum, 1986). However, results from this tester do not always relate well to actual handling breakage experience (Eckhoff et al., 1989). This is particularly true at low values of breakage susceptibility.

A study conducted by Kirleis and Stroshine (1990) showed that test weight and density are the best predictors of corn hardness. Corn hardness can also be determined by corn protein content, oil content, and test weight (Dorsey-Redding et al., 1991). Thus, test weight is important in determining not only hardness, but also breakage susceptibility. The relationship between breakage susceptibility and hardness is dependent on test weight (Pomeranz et al., 1986). Test weight is not only an indirect measure of hardness (a genetically intrinsic property), but also is reduced by mechanical and heat treatments, percentage of broken corn and foreign material, and other stress situations.

Jackson et al. (1988) reported stress cracks to be significantly correlated with breakage susceptibility as measured by WBT. Stress cracks are found in the horny endosperm and weaken the kernel. Corn endosperm consists of a thin outer layer of aleurome cells, containing oil and protein, and a large inner portion of storage tissue containing starch and protein. As reported by Cox et al. (1944) and Wolf et al. (1952), starch granules are embedded in the proteinaceous matrix. It seems reasonable that the amount of protein present and the form it takes affect the force-deformation properties of the kernel. When a smaller proportion of starch is embedded in a highly structured protein matrix, the kernel is probably relatively hard.

Therefore, corn proximate constituents probably have a significant impact on corn breakage susceptibility, and certainly affect hardness. Because near-infrared spectroscopy is in common use for rapid measurement of corn composition (Hurburgh, 1988) and it will measure density (Siska and Hurburgh, 1993), there is a potential for...
developing a rapid, near-infrared based method for estimating breakage susceptibility.

**OBJECTIVE**

The objective of this study was to develop a breakage susceptibility prediction equation as a function of corn protein content, oil content, starch content, density, and test weight.

**MATERIALS AND METHODS**

**SAMPLES**

A set of market corn samples from different U.S. locations was collected by the Grain Quality Laboratory at Iowa State University, Ames. The set included samples dried with different drying treatments. With respect to shelled corn which moves through marketing channels, the samples represented a wide range of composition and physical properties.

The set contained 111 samples from Continental Grain Company, Chicago, Illinois, and 19 samples from the Iowa Corn Growers Association, Des Moines. A group of validation samples (36 samples) was randomly selected. The remaining group (94) was used as a calibration sample set.

A different sample set was used for the principal factor method (PFM) to establish fundamental relationships (common factors) among the independent variables. The correlation set was also used by the Grain Quality Laboratory for moisture and constituent composition calibration of near-infrared analyzers, and includes data from 1986-1992 crop years. Coefficients calculated from such a set should be more representative of the whole corn population than a set of samples restricted to one year. For purposes of this study, the set was named GQL sample set.

**COMPOSITION**

Corn moisture, protein, oil, and starch contents of calibration, and validation sample sets were determined by using a ground grain near-infrared reflectance (NIR) instrument the DICKEY-john Instalab 800 (DICKEY-john, Inc., Auburn, Ill.). The NIR instrument was calibrated against chemical methods done of Woodson-Tenent Labs, Inc., Des Moines, Iowa. Protein, oil, and starch contents were adjusted to 15% moisture content.

Proximate composition of the GQL samples (on a 15% moisture basis) from the GQL data set was determined by Woodson-Tenent Labs, Inc., Des Moines, Iowa. Oven moisture contents were determined at Iowa State (AACC, 1991b).

**KERNEL DENSITY**

Approximately 33 g of corn were weighed to ± 0.001 g. Volume was measured with a Beckman model 930 air-comparison pycnometer (Beckman Instruments, Inc., Fullerton, Calif.) as described in Thompson and Isaacs (1967). Two replications of each corn sample were made. If the difference between the two replicates was more than twice the previously determined standard deviation of measurement (0.003 g/cm³), a third replication was made. Kernel density was determined as the ratio between weight and volume and was converted to 15% moisture content by using the equation derived by Dorsey-Redding et al. (1990):

\[ d_f = d_i - 0.00289 (M_f - M_i) \]

where
- \( M_f \) = ... final moisture content % (15%)
- \( M_i \) = ... initial moisture content %
- \( d_f \) = final kernel density (g/cm³)
- \( d_i \) = initial kernel density (g/cm³)

Kernel moisture for the density correction in the calibration and validation sets was measured with a Dickey-john GAC 2000 (DICKEY-john, Inc., Auburn, Ill.) capacitance moisture meter. Oven moisture was used for moisture adjustment of the samples from the GQL data set.

**TEST WEIGHT**

Test weight was determined by the USDA, Federal Grain Inspection Service standard method (FGIS, 1988). For the conversion to 15% moisture basis, moisture content was determined by using the DICKEY-john GAC 2000 capacitance moisture meter for samples from calibration and validation sample sets. Oven moisture was used for the GQL data set. Conversion to 15% moisture content was done using the following moisture conversion equation (Dorsey-Redding, 1990):

\[ T_f = T_i - 0.4412 (M_f - M_i) \]

where
- \( T_f \) = final test weight (kg/hL)
- \( T_i \) = initial test weight (kg/hL)

**BREAKAGE SUSCEPTIBILITY**

A Wisconsin Breakage Tester (Cargill Research Laboratory, Minneapolis, Minn.) was used for the breakage susceptibility test. Approximately 200 g of corn (weighed to ± 0.1 g) were tested in the WBT as described in Singh and Finner (1983) and Watson and Herum (1986). Moisture content was determined by the DICKEY-john meter. The Dutta (1986) moisture adjustment equation was used to convert the measured breakage susceptibility to 15% moisture content:

\[ B_f = B_i e^{0.29(M_i-M_f)} \]

where
- \( B_f \) = final breakage susceptibility (%)
- \( B_i \) = initial breakage susceptibility (%)

**STATISTICAL ANALYSIS**

Equations which predicted breakage susceptibility used composition, density, and test weight data as independent variables. Two statistical techniques were used: multiple linear regression (MLR), and principal factor method (PFM). All data were converted to 15% moisture content before analysis.

**MLR.** The following prediction equation was developed using the stepwise technique (SAS Institute Inc., 1987):

\[ B = b_0 + b_1 P + b_2 T + b_3 D + b_4 S + b_5 O \]

where
- \( B \) = breakage susceptibility (%)
The factors are frequently referred to as the "loadings". The basic factor analysis model is in the form:

$$z_j = a_{j1} F_1 + a_{j2} F_2 + \ldots + a_{jn} F_n + u_j Y_j \quad (5)$$

$$z_j = \text{variable in standardized form (variance of 1 and zero mean)}$$

$$a_{j1}, \ldots, a_{jn}, u_j = \text{factor loadings}$$

$$F_1, \ldots, F_n = \text{common factors}$$

$$Y_j = \text{unique factor}$$

The variance of the variable $z_j$ may be expressed in terms of the factors (Harman, 1976):

$$s_j^2 = 1 = a_{j1}^2 + a_{j2}^2 + \ldots + a_{jn}^2 + u_j^2 \quad (6)$$

$$s_j^2 = \text{variance of variable } z_j$$

$$a_{j1}, \ldots, a_{jn} = \text{loadings of common factors}$$

$$u_j = \text{loading of a unique factor}$$

From the composition of this total variance, two important concepts are defined: 1) communality and 2) uniqueness. Communality ($h_j^2$) is given by the sum of the squares of the common-factor coefficients:

$$h_j^2 = a_{j1}^2 + a_{j2}^2 + \ldots + a_{jn}^2 \quad (7)$$

It is the part of the total variance of variable $z_j$ described by factors $F_1, F_2, \ldots, F_n$. Uniqueness is the contribution of the unique factor. It indicates the extent to which the common factors fail to account for the total unit variance of the variable.

The procedure for calculating factor loadings in the principal factor method is described in Harman (1976). The first stage of the principal factor method involves the selection of the first-factor coefficients $a_{j1}$ so as to make the sum of the contributions of that factor to the total communality a maximum. This sum is given by:

$$V_1 = a_{j1}^2 + \ldots + a_{ni}^2 \quad (8)$$

where $a_{j1}$ is the loading of factor one for the $jth$ variable.

It has been proven that this sum is equal to the first root of the characteristic equation (Harman, 1976) and is called an eigenvalue. Factor coefficients (loadings) for the first factor are then calculated from this eigenvalue. The second largest root of this equation is used to derive the coefficients of the second factor. By the same argument, the coefficients of the successive factors are calculated. Total communality for the $n$ variables is the limiting value and determines when factorization should be stopped.

Using PROC FACTOR, (SAS Institute, Inc., 1988), the variables were used to calculate common factors from the Grain Quality Laboratory set. The five variables formed two common factors, whose contribution to the total common variance was examined. Loadings of these two factors per particular variable form the factor pattern.

The plot of factor patterns was examined to decide which of the variables should be used in the linear combination of the common factors. Variables that form clusters should be either described with different axis rotation (Oehrtman, 1970), or eliminated from further analysis.

Standardized scoring coefficients (coefficients $a_{jm}$) were obtained from the GQL data set. Scores of standardized data from the calibration set were calculated by:

$$S_{fi} = S C_1 \frac{P - \mu_p}{STD_p} + S C_2 \frac{O - \mu_o}{STD_o} + S C_3 \frac{D - \mu_d}{STD_d} + S C_4 \frac{S - \mu_s}{STD_s} \quad (9)$$

where

$$S_{fi} = \text{score for } ith \text{ sample}$$

$$\mu_p, \mu_o, \mu_s = \text{mean of protein, oil, and starch}$$

$$STD_p, STD_o, STD_s = \text{standard deviation of protein, oil, and starch}$$

$$STD_D = \text{standard deviation of density (g/cm}^3)$$

$$SC_1, SC_2, SC_3, SC_4 = \text{standardized scoring coefficients}$$

There was a statistically significant difference between means and standard deviations from the Grain Quality Laboratory data set and the calibration data set (table 1). Therefore, means and standard deviations for equation 9 were obtained from the calibration data set.

### Table 1. Chemical and physical properties* of corn samples used for breakage

<table>
<thead>
<tr>
<th>Set</th>
<th>Protein (%)</th>
<th>Oil (%)</th>
<th>Starch (%)</th>
<th>Density (g/cm$^3$)</th>
<th>Test Weight (kg/L)</th>
<th>B † (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GQL* (n = 252)</td>
<td>8.2a</td>
<td>3.5</td>
<td>60.4</td>
<td>1.276</td>
<td>73.5</td>
<td>-</td>
</tr>
<tr>
<td>STDI</td>
<td>1.2</td>
<td>0.4</td>
<td>1.6</td>
<td>0.026</td>
<td>3.3</td>
<td>-</td>
</tr>
<tr>
<td>Range</td>
<td>5.3-13.0</td>
<td>2.8-5.9</td>
<td>54.7-65.3</td>
<td>1.185-1.349</td>
<td>64.3-84.3</td>
<td>-</td>
</tr>
<tr>
<td>Calibration (n = 94)</td>
<td>Mean 8.1ab</td>
<td>3.7a</td>
<td>60.4ab</td>
<td>1.206</td>
<td>74.5</td>
<td>6.7</td>
</tr>
<tr>
<td>STD</td>
<td>0.6c</td>
<td>0.3</td>
<td>0.8c</td>
<td>0.020</td>
<td>2.8</td>
<td>1.7</td>
</tr>
<tr>
<td>Range</td>
<td>7.2-10.4</td>
<td>3.2-5.7</td>
<td>57.0-61.7</td>
<td>1.233-1.351</td>
<td>69.5-80.8</td>
<td>2.3-9.9</td>
</tr>
<tr>
<td>Validation (n = 36)</td>
<td>Mean 8.1ab</td>
<td>3.6a</td>
<td>60.5ab</td>
<td>1.291</td>
<td>73.9</td>
<td>6.6</td>
</tr>
<tr>
<td>STD</td>
<td>0.6c</td>
<td>0.2</td>
<td>0.7c</td>
<td>0.016</td>
<td>1.9</td>
<td>1.7</td>
</tr>
<tr>
<td>Range</td>
<td>7.0-10.1</td>
<td>3.2-4.0</td>
<td>58.5-61.7</td>
<td>1.264-1.340</td>
<td>70.8-77.5</td>
<td>3.4-9.5</td>
</tr>
</tbody>
</table>

* Composition and physical properties are converted to 15% moisture content.
† Breakage susceptibility.
¶ Standard deviation values not statistically different (p = 0.001) are indicated by similar letters, with vertical comparison only within a particular property for all three sets.
‖ Similar letters, only for vertical comparison with a particular property for all three sets.
Multiple linear regression was used to establish the breakage susceptibility prediction equation:

\[ B = b_0 + b_1 S_f + b_2 T \]  

(10)

where
- \( S_f \) = calculated scores by sample
- \( T \) = test weight (kg/L)

**Evaluation.** In this study, calculation of regression coefficients in both prediction techniques was referred to as calibration. Since density and test weight formed clusters, in the two factorial model, test weight was eliminated from factorial analysis, and was included only in the multiple regression model.

The prediction models were evaluated on the 36 corn samples that were not part of the calibration or the GQL data set. The correlation coefficient, the bias, and the standard error of prediction (SEP) were calculated for both of the prediction techniques (Williams, 1987).

**RESULTS AND DISCUSSION**

The range in corn composition and physical properties for the three sample sets are compiled in table 1. Because most of the measured variables had different means and standard deviations from the calibration set were used to calculate regression coefficients. If their values had not been different, then the GQL data set could have been used for the regression coefficient calculation.

Table 2 shows the eigenvalues and variances for the two-factor and one-factor models. Eigenvalues and appropriate variances are shown in the order in which they were calculated. Subsequent eigenvalues represent residual information after the information included in previous eigenvalues was removed. The eigenvalue of a factor, divided by the sum of the eigenvalues, gives the proportion of the total variance accounted for by that factor. The two large positive eigenvalues (table 2) account for 121% of total variance, which indicates overfactorization (an excessive number of factors were used). This means that the contribution of two factors to the total variance was more than 100%. In addition, the plot of the unrotated factor pattern (fig. 1) showed one tight cluster of variables, test weight and density, at the positive ends of factors 1 and 2. Therefore, the variable with the smaller loading (test weight in this instance) was removed from factor analysis. This reduced the number of common factors to one (Harman, 1976). One common factor was created from the variables protein, oil, starch, and density. Loadings for this solution are shown in table 3. This solution yielded one large eigenvalue that still accounted for more than 100% of the total variance.

Protein, oil, starch, density, and test weight were also used in the multiple linear regression to find a model for predicting breakage susceptibility. Oil was not significant (\( P = 0.05 \)). The two different prediction techniques are compared in table 3. The bias was -0.5% for the multiple linear regression and -0.6% for the principal factor method. The standard error of prediction was the same (1.0% point) for both prediction techniques and the coefficients of determination (\( R^2 \)) were nearly identical.

Predicted versus measured breakage susceptibility values for both of the prediction techniques are plotted in figures 2 and 3. Both models overestimated at low breakage susceptibility and underestimated at high breakage susceptibility. The slopes of the correction equations that would be adjusted for these differences are 0.44 for MLR and 0.43 for PFM. A large slope is indicative

**Figure 1—Factor pattern for two factorial model designed by principal factor method.**

**Table 3. Comparison of two prediction methods for calibration and validation sets**

<table>
<thead>
<tr>
<th>Method</th>
<th>( R^2 )</th>
<th>( F )</th>
<th>( P &gt; F )</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Calibration</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multiple linear regression*</td>
<td>0.79</td>
<td>81.31</td>
<td>0.0001</td>
</tr>
<tr>
<td>Principal factor method†</td>
<td>0.79</td>
<td>170.51</td>
<td>0.0001</td>
</tr>
<tr>
<td><strong>Validation</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multiple linear regression</td>
<td>0.65</td>
<td>-0.5</td>
<td>0.99</td>
</tr>
<tr>
<td>Principal factor method</td>
<td>0.64</td>
<td>-0.6</td>
<td>0.99</td>
</tr>
</tbody>
</table>

* \( B = 41.3559 - 0.5549 \) \( P + 0.2820 \) \( S - 10.1402 \) \( D - 0.4566 \) T.
† \( B = 43.0673 - 0.7881 \) \( S_f - 0.4895 \) T. The \( S_f \) was calculated from:
\[ S_f = 0.3993(P - 8.1340)/0.5728 + 0.1761(D - 3.7356)/0.3197 + 0.1630(D - 1.2955)/0.0197 - 0.3620(S - 60.3615)/0.7622. \]
‡ Bias.
§ Standard error of prediction.
of a need for a more robust calibration set. However, the linearity indicates that the concept is valid.

Presently, breakage susceptibility measurement is time-consuming. The regression equation developed by multiple linear regression provided for simple and rapid site-specific estimation of breakage susceptibility. The slightly more complicated equation derived by the principal factor method could be advantageous because it allows regression coefficients to be derived for the particular conditions of the user. However scoring coefficients can be derived from a common source, which represents a wider range of data.

The high variance of prediction error may have been caused, in part, by the error of measurement. Different instruments for moisture measurement (with different precision of measurement) were used for conversion of chemical and physical properties to the 15% moisture basis. The influence of moisture, which was measured with different instruments, may have been crucial in this case, because the moisture conversion equation for breakage susceptibility is exponential.

Near-infrared technology provides for measurement of chemical constituents (Hurburgh, 1983) and is capable of measuring corn density (Siska and Hurburgh, 1993). Near-infrared technology measures moisture content with lower variation compared to capacitance moisture instruments (Wu and Hurburgh, 1993). Values of chemical constituents and density as measured by near-infrared transmittance (NIT), and test weight measured by the conventional method could provide for precise and rapid prediction of corn breakage susceptibility.

**CONCLUSIONS**

Equations to predict corn breakage susceptibility (Wisconsin Breakage tester method ('B') from protein content ('P'), starch content ('S'), density ('D'), and test weight ('T'), were developed. The two statistical methods used, the principal factor method (PFM) and multiple linear regression method (MLR), yielded similar results. The prediction equation developed using the MLR technique was:

\[
B = 41.3559 - 0.5549 P + 0.2820 S - 10.1402 D - 0.4566 T
\]

\[
R^2 = 0.65, \text{ and } SEP = 1.0\% \text{ point (validation)}
\]

The prediction equation from the PFM was:

\[
B = 43.0673 - 0.7881 Sf_i - 0.4895 T
\]

\[
R^2 = 0.64, \text{ and } SEP = 1.0\% \text{ point (validation)}
\]

$Sf_i$ is the calculated score for the i-th sample, based on the following equation:

\[
Sf_i = 0.3993 \frac{P - \mu_p}{STD_p} + 0.1761 \frac{O - \mu_o}{STD_o} + 0.1630 \frac{D - \mu_d}{STD_d} - 0.3620 \frac{S - \mu_s}{STD_s}
\]

where

- $\mu_p$, $\mu_o$, $\mu_s$ = mean of protein, oil, and starch from calibration data set (percentage points)
- $\mu_D$ = mean of density from calibration data set (g/cm$^3$)
- $STD_p$, $STD_o$, $STD_s$ = standard deviation of protein, oil, and starch from GQL data set (percentage points)
- $STD_D$ = standard deviation of density from GQL data set (g/cm$^3$)

Each equation offers an advantage. The MLR equation has general use for rapid breakage susceptibility predictions. However, the PFM equation may give better predictions for those who have access to the source which represents wide range of data. This equation can then be adjusted for particular conditions of the user.
REFERENCES


Eckhoff, S. R. 1989. Corn breakage susceptibility testing; what it will mean to the corn marketing system? Feedstuffs 27 February:24-25, 50.


