Image Analysis of Whole Grains to Screen for Flour-Milling Yield in Wheat Breeding

M. BERMAN,[1] M. L. BASON,[2,3] F. ELLISON,[4] G. PEDEN,[1] and C. W. WRIGLEY[2]

ABSTRACT

Cereal Chem. 73(3):323–327

Image analysis of whole-grain samples was used to predict milling quality in wheat breeding to select for this aspect of quality, while preserving the seed intact for sowing. About 66% of the variation in flour yield for 38 grain samples could be explained by four factors computed from the images of 100 grains for each sample (mean of grain area, lengths of minor and major axes, and ellipsoidal volume), plus test weight. Test weight alone accounted for only 17% of the variation. The set of grain samples consisted of eight genotypes (three cultivars and five breeders' lines) grown at up to six sites. The method devised is suitable for a breeding program, being relatively low in labor requirement, not requiring time consuming positioning of the grains, and having low cost (less than $3,000 plus personal computer and software). The results of this preliminary study should provide direction for further development of noninvasive analysis of milling quality.

Noninvasive methods for analyzing grain quality attributes offer considerable advantages in breeding, if reliable predictions of quality can be provided, because the whole-grain samples are left intact to plant after testing. Image analysis (Keefe 1992, Sapirstein 1995) is one of the few techniques with the potential to fulfill this role for the wheat breeder. We examined milling quality as an attribute likely to be amenable to prediction nondestructively by image analysis, because milling quality is related to grain morphology and is an important characteristic used in the early stages of selection for quality in breeding (Marshall et al 1986, Wrigley and Morris 1995).

Much of the rationale for initiating this project was the observation that an experienced breeder could make predictions of milling quality by visual examination of whole-grain samples (see also Marshall et al 1986). If this is possible, the characteristics thus observed could be quantified more reliably and objectively by machine vision. Furthermore, we evaluated a test procedure (involving a minimum of labor and time requirement in sample handling) that would be suitable for screening the very large numbers of grain samples generated in breeding. In addition, we used a reasonably simple set of discriminators that would not require very large computing power and that would allow virtually "instant" results.

Research reports indicate the feasibility of using image analysis to distinguish between wheat-grain samples according to quality attributes that relate to milling quality, particularly kernel vitreosity, grain color, or grain hardness class (e.g., see Zayas et al 1986, Draper and Keefe 1989, Neuman et al 1989, Bason et al 1993, Sapirstein 1995). Evers (*unpublished data*) applied image analysis to whole and half grains, in specific orientations, from the U.K. wheat harvest but obtained relatively poor predictions of flour-milling yield.

Even if image analysis were demonstrated to provide useful predictions of milling quality, it would need to be shown to be superior to competing technologies. The main alternative under investigation at present would be whole-grain near infrared (NIR) spectroscopy (Williams and Sobering 1993, Allen et al 1995, Delwiche et al 1995).

In this preliminary evaluation of image analysis, considerable success was obtained in achieving the stated aims of predicting the milling quality of grain samples using a relatively simple image analysis procedure. By combining four of the imaged descriptors (means of grain area, lengths of minor and major axes, and ellipsoidal volume) with the test weight of the samples, it was possible to account for two-thirds of the variation in flour yield for a set of 38 grain samples.

MATERIALS AND METHODS

Samples

The potential of this approach was investigated by examining grain samples of eight genotypes (the advanced lines Sun200A, Sun216A, Sun217A, Sun231A, and Sun234A, and three check varieties, Hartog, Sunco, and Sunelg) grown in 1992 at six sites or sowing times (Moree, Myall Vale, North Star, Spring Ridge, and early and late crops at Narrabri) used in the variety trial system of the University of Sydney's Plant Breeding Institute at Narrabri in northern New South Wales, Australia. Only 38 of the theoretical maximum number of 48 samples were available for test milling, due to the rejection of 10 of the samples because of defects such as rain damage. The 10 samples not available were Sun217A, grown at Moree; Sunelg, Sun216A and Sun217A from Myall Vale; Sun216A and Sun217A from the early crop at Narrabri; Sun 216A, Sun 217A, and Sun 231A from North Star; and Sun 216A from Spring Ridge. All 38 samples tested were available as greater than 2 kg of grain, permitting quality testing to be performed at normal laboratory scale.

Test weights were determined for each of the samples as bulk density (kg/hl). Milling quality was determined by laboratory-scale milling of 2 kg of each sample on a Buhler MLU 202 test mill, as described by Butcher and Stenvert (1972) in the Bread Research Institute of Australia, Sydney. Milling quality was mainly considered as the yield (percentage) of white flour expressed on a total-products basis. Flour yields ranged from 72 to 79%, with a mean of 75.4% (standard deviation of 1.7%). For this type of test milling, Brown and Scanlon (1992) reported a repeatability of 0.5% in extraction rate and an operator error of 0.95%.

Image analysis

Images were captured for 37 of the 38 samples as duplicate sets (A and B) of 50 grains per sample. One hundred randomly chosen grains were randomly divided into two sample sets of 50 grains each. (For a 38th sample [Sunelg at North Star], 95 grains were divided into subsamples of 45 and 50 grains.) Each sample set of 45 or 50 grains was, in turn, arranged on corrugated cardboard

---

[1]CSIRO Division of Mathematics and Statistics, Macquarie University, NSW 2113, Australia.
[2]CSIRO Division of Plant Industry, North Ryde, NSW 2113, Australia.
[3]Newport Scientific, Warriewood, NSW 2102, Australia.
[4]University of Sydney Plant Breeding Institute, Narrabri, NSW 2390, Australia.

consisting of 9 or 10 corrugated columns with five grains end-to-end in each column, as shown in Figure 1. This arrangement of grains was chosen to minimize contact between the grains and to stop them from moving during handling. There was no attempt to impose uniform orientation of the grains; as a result, images range from dorsal to ventral view and other aspects in between. This approach to grain presentation suited our requirements well, namely, to be quick and "foolproof" (needed in the routine screening of many samples) and to facilitate automatic measurement of grain parameters.

Color images of each sample set were captured using a color CCD camera (JVC KY-F30 3-CCD). Each image was 512 × 768 pixels in size, and at the magnification used, 1 mm on the image was equivalent to 10.3 pixels. Figure 1 is a typical image. Analysis of the images (routinely the red component of the RGB image) first involved segmentation to obtain discreet outlines of the grains. This was done using a "seeded-region growing algorithm" (Adams and Bischof 1994). In this technique, the lightest pixels are used to identify "seeds" for grains, and the darkest pixels are used to identify seeds for the background. The algorithm grows out from the seeds until the natural boundaries are found. This sometimes causes grains to touch one another. The touching grains are separated using a modified version of the binary watershed algorithm of Creek (1991). This latter algorithm has a parameter that controls the amount of touching allowed. If the parameter is set too high, it sometimes separates the ends of grains from the main grain as well as separating touching grains. Where necessary, this parameter was adjusted manually to obtain satisfactory results. The result of applying these steps to the image in Figure 1 is shown in Figure 2; the identified grains are marked in black.

RESULTS

In developing a strategy for measuring desirable aspects of grain morphology, we started with the visual factors that were presumed to be observed by the trained expert in grain examination for milling quality, namely: large size and uniformity of size; plumpness and shape, well-filled short grain with narrow or flat crease; rounder rather than longer/narrower; consistency of shape and plumpness; uniform smooth surface, absence of depressions or corrugations on surface; small- to medium-sized embryo protruding, rather than sunken or depressed; large dense brush undesirable; semitranslucency desirable for hard-grained samples in the appropri-

ate protein range and in the absence of weather damage; and chalky or starchy (soft) grain are rejected if a hard grade is the objective.

The simplified image analysis process chosen could provide only satisfactory measurements of the first few features in this list, namely those related to size and overall shape. For each grain, we computed the following five size and shape measures: area, perimeter, major axis length, minor axis length, and eccentricity (a measure of plumpness) (see Table I). The last three measures are obtained by using a common and fast means of fitting an ellipse to each grain boundary, by equating the observed second-order moments of the grain to the theoretical second-order moments of an ellipse (Rosenfeld and Kak 1982). A sixth measure, which we call ellipsoidal volume, was also computed for each grain. This was chosen as a predictor of grain volume, and hence flour yield, although grain volume cannot be measured directly from the two-dimensional images. However, if we make the crude assumption that a grain is an ellipsoid with circular cross sections (i.e., with axis lengths M, m, and m), then its volume is $\pi M m^2/6$ (the ellipsoidal volume).

For each image, the mean and standard deviation (SD) of each of the six measures were calculated over all the grains in the image. The SDs were considered as measures of uniformity. We also considered the test weight of the samples, as this attribute is routinely measured during screening and it provided another nondestructive indication of grain morphology.

About 40% of the variation in flour yield was explained by either mean grain area, mean minor axis length, or mean ellipsoidal volume, used one at a time (Table I). The basis of this estimation was the estimated percentage variation (EPV) (sometimes called the adjusted R^2) obtained by performing a linear regression of flour yield separately on each of the six means, six standard deviations or on test weight. The EPV is given by:

$$\max\,[0, 100(1 - s_1^2/s_0^2)]$$

where s_0^2 is the estimated residual variance obtained from fitting a constant to flour yield, and s_1^2 is the estimated variance obtained from the linear regression.

Poorer estimates of variance were provided by using any one of the variables in Table I, other than the above three variables. The standard deviations had little predictive power. Estimates were reasonably similar using either of the two sample sets (A or B), or the combined results from both sets of 2 × 50 grains. Flour yield showed a broadly monotonic relationship with each of the six means
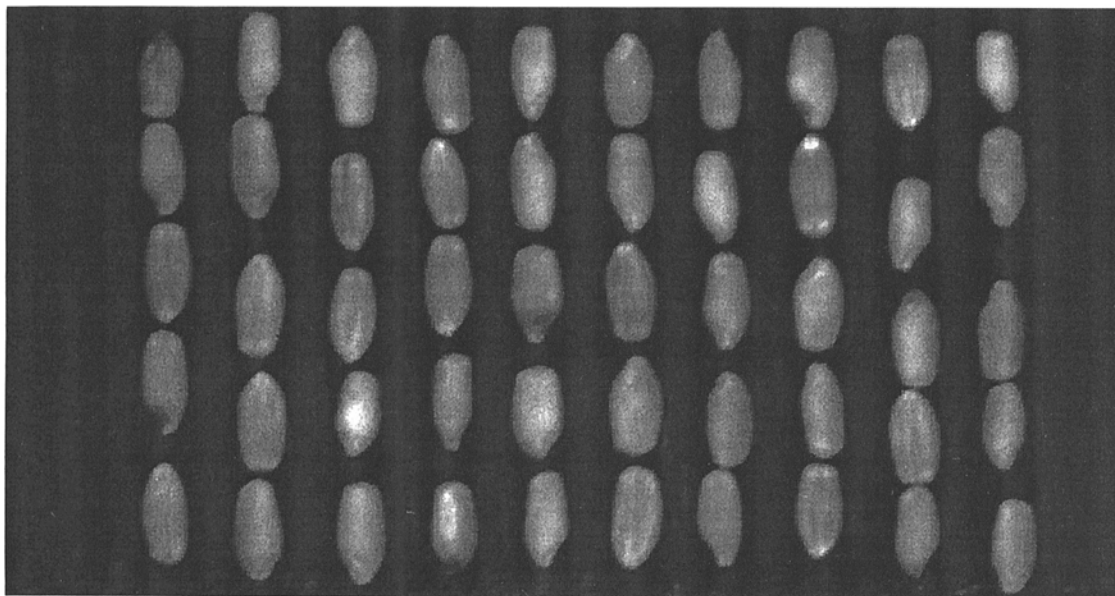

Fig. 1. Image of one of the two sample sets for Sun231A grain, grown at Moree, showing the arrangement of grains.

of image analysis variables. Test weight, used alone, explained only 17% of the variation in flour yield.

Next, the variables in Table I were taken in various combinations to thus improve the prediction of flour yield. All possible combinations were examined. The best combinations of between one and 10 variables are shown in Table II for the results from the combined A and B sets of images. Similar "best" combinations of attributes were obtained when only A or B sets of images were compared. In this case, "best" is defined as the subset minimizing s_p^2, the estimated residual variance obtained from a linear regression of flour yield on the p variables, where p varies from 1 to 10. Table II also shows the estimated explained percentage variation (EPV_p), whose general formula is:

$$EPV_p = \max [0, 100(1 - s_p^2/s_0^2)]$$

In this exercise, minimizing s_p^2 is equivalent to maximizing EPV_p. None of the standard-deviation measures was included in any of the best combinations of one to four attributes. On the other hand, test weight was included in all sets other than those based on a single attribute.

To formally test for the best combination of attributes (optimal value of p) to use for predictive purposes, a naive approach would

be to maximize EPV_p. This approach would indicate the use of nine variables, thereby explaining 71% of the variation in flour yield. However, with this choice of p, a different subset of variables was indicated depending whether the A, B, or both sets of images was examined, and intuitively, this appears to be an unnecessarily large set of attributes. Another widely used technique to resolve such a problem is that of Mallows (1973), based on the statistic C_p, defined as:

$$C_p = RSS_p/s_{13}^2 - n + 2(p + 1)$$

where n is the sample size (38), and RSS_p is the residual sum of squares obtained from fitting the best subset of size p, and s_{13}^2 is the estimated variance obtained from fitting all 13 variables (note that $s_p^2 = RSS_p/[n - p - 1]$). Assuming that there is a "true" linear model with p parameters, and if we replace the estimate s_{13}^2 in the above formula with the true residual variance, then it can be shown that the expected value of C_p, $E(C_p)$, equals $p + 1$ if all the true values are included, and is greater than $p + 1$ if not all the true variables are included. Based on this result, one rule says to

TABLE I
Percent Variation Explained by Each (Individually) of 12 Variables Determined by Image Analysis of the A and B Sets of Grains, Plus Test Weight

Variable	Variable Symbol	A Set	B Set	Combined Samples
Area mean	μ_A	41.52	35.76	39.28
Perimeter mean	μ_P	30.19	23.98	27.77
Minor axis length mean	μ_m	41.91	37.31	40.14
Major axis length mean	μ_M	14.09	6.75	10.36
Eccentricity mean	μ_E	28.34	26.50	27.76
Ellipsoidal volume mean	μ_V	42.25	38.07	40.79
Area SD	σ_A	4.62	5.96	7.19
Perimeter SD	σ_P	0	6.22	0.77
Minor axis length SD	σ_m	0	0	0
Major axis length SD	σ_M	0.71	16.73	9.12
Eccentricity SD	σ_E	0	0	0
Ellipsoidal volume SD	σ_V	12.72	9.54	13.39
Test weight	TW	17.03	17.03	17.03

TABLE II
Combinations of Multiple Variables (p = number selected) According to Their Ability to Predict Flour Yield, Based on the Combinations of A and B Images of Grains, Indicated as Estimated Percent Variation (EPV_p), Mallows' Statistic (C_p), and Mean Squared Error of Prediction ($MSEP_p$)

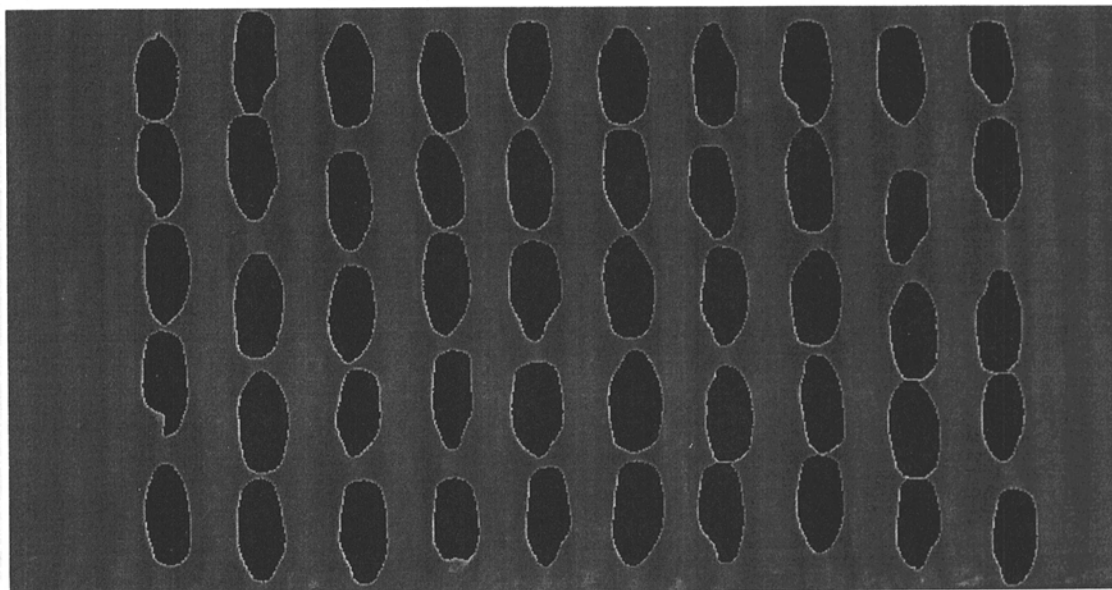| p	Chosen Subset	EPV_p	C_p	$MSEP_p$
1	μ_V	40.8	31.65	1.81
2	(μ_V, TW)	48.6	23.41	1.62
2	(μ_A, TW)	47.5	24.61	1.65
3	(μ_A, μ_P, TW)	53.0	19.18	1.52
4	(μ_A, μ_m, μ_M, TW)	58.7	13.97	1.38
4	(μ_A, μ_m, μ_E, TW)	58.2	14.44	1.39
5	(μ_A, μ_m, μ_M, μ_V, TW)	66.4	7.15	1.16
5	(μ_A, μ_P, TW, σ_P, σ_V)	60.8	12.68	1.35
5	(μ_A, μ_m, μ_E, μ_V, TW)	59.3	14.09	1.40
6	(μ_A, μ_m, μ_M, μ_E, μ_V, TW)	67.9	6.66	1.14
6	(μ_A, μ_P, μ_m, μ_M, μ_V, TW)	66.2	8.26	1.20
6	(μ_A, μ_P, TW, σ_P, σ_M, σ_V)	59.7	14.49	1.42
7	(μ_A, μ_m, μ_M, μ_V, TW, σ_A, σ_M)	69.8	5.92	1.11
8	(μ_A, μ_P, μ_m, μ_M, μ_V, TW, σ_A, σ_M)	70.2	6.60	1.13
9	(μ_A, μ_P, μ_m, μ_M, μ_E, μ_V, TW, σ_A, σ_M)	71.1	6.91	1.13
10	(μ_A, μ_P, μ_m, μ_M, μ_E, μ_V, TW, σ_A, σ_m, σ_M)	70.6	8.43	1.20

Fig. 2. Image of grains in Figure 1, processed to separate touching grains.

choose p minimizing C_p. We have included C_p for the best subset for each p in Table II. Note that it sometimes has several minima. If the smallest minimum is chosen, this rule gives $p = 5$, based on the values in Table II, or $p = 7$ or 5 based on only the A or B sets of images, respectively.

A second rule chooses p such that $E(C_p)$ is closest to $p + 1$. This rule gives $p = 6$ for the values in Table II, or 8 and 5 for the A and B sample sets, respectively.

The third widely used technique, and the one most directly related to what one would ultimately wish to do with these variables, is the *PRESS* (prediction sum of squares) statistic of Allen (1974), summarized as follows. Let Y_i denote the ith flour yield out of $n = 38$. For a given set of p predictors, remove Y_i and perform a linear regression based on the remaining $n - 1$ observations; let Y_{ip} denote the predicted value from this linear regression. Then the *PRESS* statistic for a particular subset of p predictors is:

$$PRESS_p = \sum_{i=1}^{n}(Y_{ip} - Y_i)^2.$$

In this technique, one chooses the best subset of any size minimizing $PRESS_p$. In practice, this operation need not actually be carried out, as mathematical analysis can be used to work out a good approximation (see Section 6.3 in Miller 1990). Thus, to a reasonable approximation:

$$PRESS_p = \sim RSS_p/(1 - (p + 1)/n)^2$$

Table II also shows the approximate estimated mean squared error of prediction ($MSEP_p = PRESS_p/n$), which indicates how accurately, on average, a particular subset of predictors is likely to predict flour yield. The minimum *PRESS* rule chooses $p = 7$ for the combined set of image samples (9 or 7 for A or B sets, respectively, each alone). This value tends to be higher than minimizing C_p but lower than maximizing EPV_p.

Overall, based on indications from these three estimates of "best" prediction, a combination of four image measurements (means of area in square millimeters, minor and major axes in millimeters, and ellipsoidal volume in cubed millimeters) together with test weight (kg/hl), is recommended for the prediction of flour yield as percent extraction for this set of data. The value of this relationship is shown graphically in Figure 3, in which a correlation coefficient of ($r = 0.84$) is obtained for the actual versus predicted flour yield.
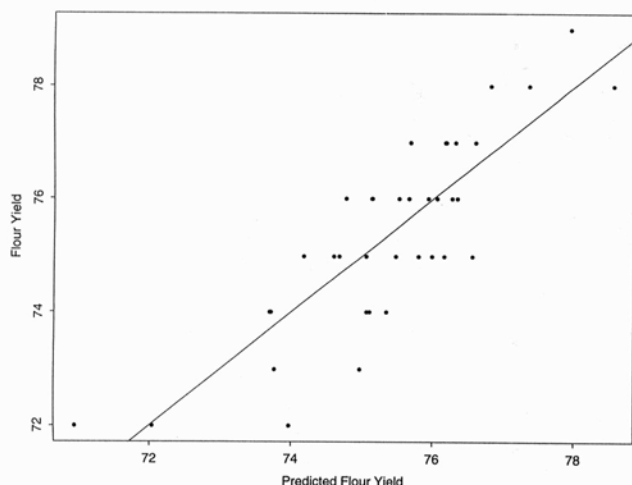

Fig. 3. Actual flour yield versus the yield predicted using the linear model based on means of area, of minor and major axes, and of ellipsoidal volume, measured automatically from the images for all sample sets (A and B), plus test weight.

The model of prediction thus resulting from this study can be expressed as:

Flour yield = 343.4 + 30.6 × mean area – 96.1 × mean minor axis length – 60.2 × mean major axis length – 2.59 × mean ellipsoidal volume + 0.294 × test weight

DISCUSSION AND CONCLUSIONS

About 66% of the variation in milling quality (as percent actual flour yield) could be predicted by a linear equation, based on four of the measurements from the images, namely, means of area, minor and major axes, and ellipsoidal volume, together with test weight (Table II). Test weight alone, on the other hand, accounted for only 17% of the variation, consistent with the results of Marshall et al (1986) of a weak positive correlation between test weight and flour yield. As flour yield has its own intrinsic variability (due to lack of reproducibility), the percentage variation in the "true" flour yield explained by these five measurements is likely to be somewhat higher that the estimated 66%.

The linear regression in Figure 3 does not necessarily provide the best indication of the value of the prediction as it might be applied in breeding. A better evaluation is to apply the predictive conclusions to the actual decision that the breeder must make, namely, to determine which samples to discard or retain for future propagation. In the case of the prediction shown in Figure 3, the breeder might decide to discard those samples with a predicted flour yield of 74% and below. This decision would identify all but four of those with actual flour yields of 74% or less. On the other hand, a decision based on over 76% predicted yield would include all the best lines (with actual yields of over 76%), together with a few samples having medium milling quality (actual flour yields of 75% and 76%), but no poor samples. These classification errors could probably be reduced a little by applying a discriminant analysis rather than a linear regression. We have not, however, pursued this possibility.

This extent of predictability would be valuable in a breeding program, given that the procedure is nondestructive and not labor-intensive, and if the equipment can be obtained at reasonable cost. Because a relatively simple and small range of grain characteristics was selected for use in this study, low-cost equipment would be sufficient for its routine application in breeding, namely a camera, with frame grabber (costing under about $3,000), plus personal computer and software.

As advanced lines and registered cultivars were the test samples in this study, the range of variability in milling quality was probably narrower than it would be for a more typical set of samples encountered at early generation in a breeding program, because poorest quality lines would not have been culled out. Thus, the application of this approach to early-generation lines could be expected to be more effective than was obtained in this experiment.

There would probably be advantages in combining image analysis with NIR analysis, which, as mentioned earlier, also shows some promise for nondestructive evaluation of milling quality (Allen et al 1995). As NIR might be expected to provide predictions on a chemical basis, it should complement and extend the predictive value of physical measurement by whole-grain image analysis.

It is surprising that none of the best predictive measures included any of the standard deviation measures, which were included on the assumption that the indication of uniformity that they would provide would be relevant to milling quality. On the other hand, the results for this set of grain samples, (expressed in the last equation) suggest that better flour yield might be expected from larger grains, but the prediction equation does not necessarily serve to indicate what aspects of grain shape should provide better milling quality because the component measures are highly positively correlated.

However, any such deductions are severely limited by the preliminary nature of this exercise. It is important to point out that even the success of this predictive model is confined in its implications because of the limited extent of samples examined. The approach needs to be extended and tested on a much wider basis, covering more seasons, sites, and genotypes, within this general quality type (breeding directed toward Australian Prime Hard type) and beyond. Nevertheless, the results provide an indication of research directions for further such endeavors.

ACKNOWLEDGMENT

LITERATURE CITED

ADAMS, R., and BISCHOF, L. 1994. Seeded region growing. IEEE Trans. Pattern Anal. Machine Intelligence 16:641-647.

ALLEN, D. M. 1974. The relationship between variable selection and data augmentation and a method of prediction. Technometrics 16:125-127.

ALLEN, D. M., BLAKENEY, A. B., and OLIVER, J. R. 1995. Early generation quality selection using NIT. Pages 340-342 in: Proc. 44th Aust. Cereal Chem. Conf. J. F. Panozzo and P. G. Downie, eds. Royal Australian Chemical Institute: Melbourne.

BASON, M. L., PEDEN, G. M., ZOUNIS, S., WRIGLEY, C. W., and BERMAN, M. 1993. Detection of red-grained wheat by tristimulus colorimetry and digital image analysis. Pages 29-34 in: Proc. 43rd Aust. Cereal Chem. Conf. C. W. Wrigley, ed. Royal Australian Chemical Institute: Melbourne.

BROWN, G. L., and SCANLON, M. G. 1992. Ensuring accuracy of experimental milling through NAMAS accreditation. Assoc. Op. Millers Nov. 1992: 6122-6126.

BUTCHER, J., and STENVERT, N. L. 1972. An entoleter for the Buhler laboratory mill. Milling 154(7):27-29.

CREEK, R. C. 1991. Convex segmentation for natural objects. Pages 466-474 in: Conference Proceeding of the Australian Pattern Recognition Society. DICTA-91.

DELWICHE, S. R., CHEN, Y.-R., and HRUSCHKA, W. R. 1995. Differentiation of hard red wheat by near-infrared analysis of bulk samples. Cereal Chem. 72:243-247.

DRAPER, S., and KEEFE, P. D. 1989. Machine vision for the characterisation and identification of cereals. Plant Varieties Seeds 2:53-62.

KEEFE, P. D. 1992. A dedicated wheat grain image analyzer. Plant Varieties Seeds 5:27-33.

MALLOWS, C. L. 1973. Some comments on c_p. Technometrics 15:661-675.

MARSHALL, D. R., MARES, D. J., MOSS, H. J., and ELLISON, F. W. 1986. Effects of grain shape and size on milling yields in wheat. II. Experimental studies. Aust. J. Agric. Res. 37:331-342.

MILLER, A. J. 1990. Subset Selection in Regression. Chapman and Hall: London.

NEUMAN, M. R., SAPIRSTEIN, H. D., SHWEDYKE, E., and BUSHUK. W. 1989. Wheat grain colour analysis by digital image processing. II Wheat class discrimination. J. Cereal Sci. 10:183-188.

ROSENFELD, A., and KAK, A. 1982. Digital Picture Processing, Vol. 2. Academic Press: San Diego.

SAPIRSTEIN, H. D. 1995. Variety identification by digital image analysis. Pages 91-130 in: Identification of Food-Grain Varieties. C. W. Wrigley, ed. Amer. Assoc. Cereal Chem.: St. Paul, MN.

WILLIAMS, P. C., and SOBERING, D. C. 1993. Comparison of commercial near infrared transmittance and reflectance instruments for analysis of whole grains and seeds. J. Near Infrared Spectrosc. 1:25-32.

WRIGLEY, C. W., and MORRIS, C. F. In press. Breeding cereals for quality improvement. In: Cereal Grain Quality. R. J. Henry and P. S. Kettlewell, eds. Chapman and Hall: London.

ZAYAS, I., LAI, F. S., and POMERANZ, Y. 1986. Discrimination between wheat classes and varieties by image analysis. Cereal Chem. 63:52-56.