A COMPUTERIZED METHOD FOR EVALUATING DURUM WHEAT QUALITY[1]

J. W. DICK and W. C. SHUEY[2]

Cereal Chemistry 53(6): 910-915

ABSTRACT

A computerized technique is described which included 11 variables for quality evaluation of durum wheat. Percentage deviation from the mean of selected standard varieties is the determining factor for most of the variables in calculating the degree of variance from acceptable performance. For 132 randomly selected samples, the correlation coefficient between the calculated and the subjective evaluations was +0.92. The computer printout included all data in one table.

Until the last decade, methods used by cereal chemists for quality evaluation of wheats were time-consuming and cumbersome. The evaluations for varietal selections were subjective, reflecting uncontrollable rater bias and more rater inconsistency than desirable. Gilles *et al.* (1) in 1965 and Niernberger and Johnson (2) in 1970 described two possible methods of quality evaluation by use of computer programming. A more recent computerized evaluation method was described by Shuey *et al.* (3). Notably, these methods were primarily concerned with evaluation of bread wheat, although Gilles *et al.* (1) mentioned a program for durum wheat.

Laboratory data for an individual variety at one location are of little consequence, because variety × environment interaction affects wheat varieties differently. The location and year in which the sample was grown must be considered when comparing varieties. Therefore, when one considers the necessary combinations of varieties and growing conditions required to make a sound evaluation, the large number of samples that would be involved becomes apparent. A computer program employing the Statistical Analysis System (SAS)[3] was developed to approximate, as nearly as possible, our subjective rating laboratory technique for evaluating durum wheat varieties and selections. This program had the marked advantages of eliminating the bias of the subjective rating technique and optimizing consistency, as well as minimizing data computation time.

SAMPLES AND METHODS

We selected 923 nursery samples of durum wheat previously evaluated subjectively to develop the quality evaluation program. The nursery samples

[1]Cooperative investigations, North Central Region, Agricultural Research Service, U.S. Department of Agriculture, and the Department of Cereal Chemistry and Technology, North Dakota State University, Fargo. Published with the approval of the Director of the Agricultural Experiment Station, North Dakota State University, Fargo, as Journal Series No. 672.

Mention of a trademark name or proprietary product does not constitute a guarantee or warranty of the product by the U.S. Department of Agriculture, and does not imply its approval to the exclusion of other products that may also be suitable.

[2]Respectively: Food Technologist and Research Food Technologist, North Central Region, Agricultural Research Service, U.S. Department of Agriculture, North Dakota State University, Fargo, ND 58102.

[3]Statistical Analysis System (SAS)—designed and implemented by A. J. Barr and J. H. Goodnight at the Department of Statistics, North Carolina State University.

from one station and year constituted a series. The samples, which made a total of 92 series, were randomly selected from 5 crop years (1967–71), seven states (Calif., Idaho, Minn., N. Dak., Oreg., S. Dak., and Wash.), and 25 stations. Leeds was the standard variety for most of the series. However, Leeds and Wells were used as standards when both were grown in a series.

All quality evaluations were made according to year and station. Originally, the means of eight variables were computed for the standard varieties; and the percentage deviations of the test sample from the corresponding standard means were determined. The eight independent variables included test weight, 1000-kernel weight, percentage large kernels, percentage small kernels, wheat protein, percentage semolina extraction, semolina dust color, and spaghetti color. Correlations between the deviation of each independent variable and the dependent variable (VAL, the final evaluation) were computed, and plots for each independent variable were constructed. Also, the differences between the means were subjected to a "Maximum R^2 Improvement" technique (4) to gain insight into the relative strength of the relation between the deviation of each of the eight independent variables and the VAL in the eight-variable model.

Three independent variables (auxiliary variables)—semolina bran specks, cooked spaghetti firmness, and cooked spaghetti residue—were not included in the original calculations because of an insufficient number of samples. However, they were included in the final evaluation equation if they were determined on the test sample. The eight-variable model was used as the basis for computing the correlations for the eight original independent variables and indicated their respective weighted contributions in past evaluations. The auxiliary variables were assigned arbitrary significance based on their past relative importance in determining pasta quality.

The total of 11 independent variables was selectively incorporated into weighted rating equations. Each of the 11 was rated by arbitrary faulting limits on a scale of 1 to 4, with 1 = "no promise"; 2 = "little promise"; 3 = "some promise"; and 4 = "good promise." The rating corresponded to the final evaluation. For all but the testing factor, wheat protein, absolute limits were established which automatically rated the final evaluation as 1, or "no promise." If there was not an automatic rating of 1, the weighted means of the independent variables in the respective weighting equations were used to derive the final evaluation.

Because of the large number of durum samples we have received in recent years, it has become prohibitive to perform all the evaluation tests on each sample. In addition, the size of the durum sample can limit the number of tests which can be completed. Such limitations prompted us to formulate 12 separate weighting equations, each representing a different combination of variables for the final evaluation of the sample. Thus, it was possible to distribute the weight of each variable in proportion to only those variables included in each weighting equation. The 12 equations were identified, or labeled, and incorporated into a single computer program. By using these 12 equations, anywhere from 7 to 11 variables in various combinations can be evaluated. The resulting program sorted the raw laboratory data, calculated the final evaluation, and printed out the report in tabular form. Percentage medium kernels was included in the output, but had no direct bearing on the outcome of the final evaluation. Each standard variety (as well as the experimental varieties) within a given station was evaluated against the average of the standard varieties.

The assigned coefficients for the one weighting equation, which included all the independent variables in the program, were as follows: 2.0% for test weight, 2.5% for 1000-kernel weight, 2.5% for percentage large kernels, 2.0% for small kernels, 1.0% for wheat protein, 15.0% for semolina extraction, 10.0% for semolina bran specks, 30.0% for total color (7.5% for semolina dust color and 22.5% for spaghetti color), 30.0% for cooked spaghetti firmness, and 5.0% for cooked spaghetti residue. The other 11 weighting equations are different 5.0% for cooked spaghetti residue. The other 11 weighting equations are different from this one because the removal of just one variable from an equation proportionately redistributes the weighting value of each of the remaining variables. The weighting value assigned each variable depended not only on the number of variables in the equation, but also on their significance. For example, if no cooking tests were included in the equation, the weight previously allotted to the cooking variable was distributed proportionately among the remaining factors according to their respective weights in the all-inclusive equation. When both semolina dust color and spaghetti color were included, their weight values were combined in the ratio of 1:3, respectively. This weight represented the total color emphasis for that particular equation. Furthermore, when only one of the two color factors was included in the equation, only that color was emphasized, so that the relative weight of total color would remain constant for all 12 weighting equations.

DISCUSSION

The final program was established to evaluate durum quality objectively. The merits of the assigned coefficients and faulting limits were determined from the computer program evaluation of 923 samples. The correlation coefficient between the calculated and subjective final evaluations was significant ($r = 0.82$). It was also significant ($r = 0.78$) when the standard varieties were excluded. These standard varieties were automatically rated 4 by the subjective method, but could have been rated lower by the computer technique. Therefore, the lower correlation was expected when the standards were omitted; most of the samples were compared against only one standard variety and chances of the calculated rating being lower than 4 were reduced.

Rating differences between the calculated and the subjective final evaluations were obtained for each sample. The data are shown in Fig. 1 as percentages of the total number of samples for the given differences. The slight skewness toward the left exhibited by the histogram shows that the subjective evaluation was less severe. The distribution of the differences between the calculated and the subjective evaluations was so nearly symmetrical, however, that when differences did occur, chances of their being positive or negative were almost even. In addition, nearly 96% of the differences were within one rating unit.

A scatter diagram was constructed for the calculated *vs.* the subjective final ratings (Fig. 2). The size of the dot is proportional to the total number of samples respective to each rating. The data again show the slightly greater severity of the computer method over the subjective one. The greatest discrepancy between the calculated ratings and the average of the subjective ratings was at the lower end of the scale, where a calculated rating of 1 compared with an average subjective rating of 1.6. The calculated ratings of 2, 3, and 4 corresponded very well with the subjective rating averages of 2.1, 3.0, and 3.9, respectively.

To further test the program, an additional 132 samples (not including the standard varieties) from the crops of 1967–71 were randomly selected for evaluation. Data for each of the 12 weighting equations were available for some of these samples. The samples that had not been evaluated previously were subjectively evaluated prior to testing by the computer method. The coefficient of correlation between the subjective and objective final evaluations was +0.92, as compared to +0.78 for the original samples less the standards. The higher correlation for the 132 samples may have been due, in part, to more consistent subjective evaluation; when some of the samples were evaluated just prior to computer evaluation, the cut-off points of each variable had been established. Ninety-nine per cent of the evaluations for the two methods were within one rating unit, as compared to 96% for the original samples.

Figure 3 is a typical computer printout of data. Eleven separate deficiency columns were added to focus attention on problem areas of a sample. Deficiency was an independent variable and rated as follows: 3 = PB (probable deficiency), 2 = MN (minor deficiency), and 1 = MJ (major deficiency). A major deficiency for any one factor would automatically lower a final evaluation rating by one unit, all other factors being equal. This computer program, established within the confines of a comprehensive SAS computer package, does allow the performance of many supplementary statistical operations with the insertion of a few additional procedure statements, but limits the flexibility of the printout format. The inflexibility of the printout format in the computer package prohibits the inclusion of footnotes; thus, they must be added later if desired. Also, the units of measure of the variables must be added. Essentially, statistical operations can be increased at the expense of a completed printout. Although the SAS package does not permit it, footnotes and units of measure can be included in the printout automatically by establishing a new printout format separate

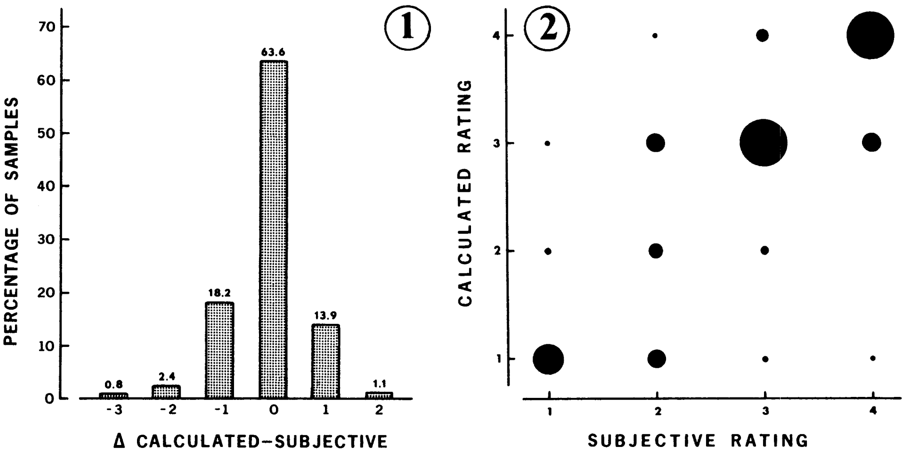

Fig. 1. Percentage distribution of differences between calculated and subjective final evaluations. Fig. 2. Scatter diagram of calculated *vs.* subjective final ratings (dot size proportional to number of samples).

DURUM QUALITY EVALUATION 1974 CROP

STATE=NORTH_DAKOTA STATION=LANGDON NURSERY=FIELD-PLOT

VARIETY	TW	KW	LG	MD	SM	PR	SEEX	SP	DU	VI	FR	RE 2/	VAL	DEFICIENCIES 5/ TW	KW	LG	SM	PR	MG	SP	DU	VI	FR	RE	SS
	#/Bu.	g	%	%	%	% 1/	%					%	3/						4/						6/
SERIES - 1																									
LEEDS	61.5	31.5	11	85	4	17.3	50.7	7	94	9.5	6.17	5.1	3					MN					MN		YS
WARD	60.5	33.6	15	79	6	16.2	53.7	13	91	9.5	5.29	6.1	4												YS
WELLS	59.5	25.9	6	77	17	16.3	51.2	10	90	9.5	5.31	7.8	3		MN	PB	MJ							MN	YS
ROLETTE	62.0	33.4	11	84	5	16.0	53.8	11	93	9.5	5.51	5.2	4										PB		
WASCANA	58.5	28.5	11	77	12	16.8	50.8	10	98	10.0	6.31	5.3	3	PB	PB		PB						MN		
SERIES - 2																									
LEEDS	62.0	40.5	49	49	2	15.0	46.8		94	10.0			4												YS
WARD	61.0	45.5	51	46	3	14.2	48.1		91	9.5			4												YS
WELLS	62.0	38.2	41	55	4	14.1	48.2		91	9.5			4		MN	PB									YS
HERCULES	61.5	48.8	64	33	3	14.1	50.3		91	9.0			4									PB			
ROLETTE	61.5	45.5	59	39	2	14.3	47.9		89	9.0			3								PB	PB			
SERIES - 3																									
LEEDS	62.5	37.5	23	75	2	14.7	48.4		96				4												YS
WARD	61.5	41.5	27	71	2	15.0	47.6		96				4												YS
WELLS	61.0	26.2	3	77	20	15.4	43.7		95				1		MJ	MJ	MJ		MN						YS
CROSBY	61.0	41.3	23	75	2	13.7	46.1		96				4												
MINDUM	62.5	37.6	21	76	3	15.0	49.2		86				2								MJ				

1/ 14% Moisture basis.
2/ TW = Test weight; KW = 1000-Kernel weight; LG = Large kernels; MD = Medium kernels; SM = Small kernels; PR = Wheat protein; SEEX = Semolina extraction; SP = Number of specks per 64.5 sq. cm; DU = Semolina dust color; VI = Spaghetti color; FR = Cooked spaghetti firmness in g cm; RE = Cooked spaghetti residue.
3/ VAL = Final evaluation; 1 = No promise; 2 = Little promise; 3 = Some promise; 4 = Good promise.
4/ MG = Milling deficiency based on percent semolina extraction.
5/ PB = Probable; MN = Minor; MJ = Major.
6/ SS = Station standard; YS indicates standard.

Fig. 3. Typical computer printout of data with footnotes added.

from the SAS package and including the new format in this computer program. This procedure may be desirable if the printout is to be used as a data table for a final report or publication. Labeling each sample for the appropriate weighting equation is an inconvenience. However, the method allows the manipulation of separate equations in one computer program and allows the use of a vast number of statistical operations, if desired.

There are many advantages of the described computer method over the subjective one. The computer method is:

1) timesaving—data are calculated and tabulated in one function;
2) consistent—samples are rated on the same basis from day to day without bias;
3) flexible—different combinations of variables can be evaluated within the same program, weighting values and variable limits can be changed as desired, and supplementary statistical functions can be performed;
4) explicit—specific deficiencies and their magnitude are pointed out;
5) concise—raw and refined data are included in one table.

There are five primary differences between the program mentioned by Gilles *et al.* (1) and ours. The former establishes the standard by manually calculating the average of the check varieties for each series of data, whereas ours is calculated automatically by the computer. The former is a separate program in itself, whereas ours is a program within a statistical package system enabling one to perform statistical functions quite easily within the same program. The former includes the use of a faulting card, whereas ours makes use of the percentage deviation from the standard for faulting; the former categorizes the faults as major or minor, whereas ours also includes a probable category; and the former has a separate printout sheet for the faults, but ours prints the data and faults on one sheet.

Acknowledgment

The authors thank J. E. Nelson for his helpful suggestions on the use of the Statistical Analysis System.

Literature Cited

1. GILLES, K. A., SIBBITT, L. D., HOLOIEN, M. O., and PETERSON, D. E. Computer evaluation and machine listing of wheat quality data. Cereal Chem. 42: 247 (1965).
2. NIERNBERGER, F. F., and JOHNSON, J. A. A computer approach to wheat quality analysis. Cereal Sci. Today 15: 117 (1970).
3. SHUEY, W. C., SPRICK, K. J., and SCOUTEN, C. W. A computerized method for wheat quality evaluation. Cereal Foods World 20: 46 (1975).
4. SERVICE, J. A user's guide to the statistical analysis system. Student Supply Stores, N.C. State Univ.: Raleigh (1972).

[Received August 19, 1975. Accepted March 1, 1976]