PRECISION OF INTER-LABORATORY MEASUREMENTS

Charles L. Schmalz, General Chemist
Jerry W. Steinmeyer, Quality Control Supervisor

THE AMALGAMATED SUGAR COMPANY
OGDEN, UTAH

presented at the
TWENTY-SIXTH GENERAL MEETING
AMERICAN SOCIETY OF SUGAR BEET TECHNOLOGISTS

FEBRUARY 24-27, 1991

Section D

PRECISION OF INTER-LABORATORY MEASUREMENTS

C.L. Schmalz and J.W. Steinmeyer

INTRODUCTION

A fundamental principle of analytical chemistry is the need for equipment calibration and reagent standardization. This is drilled into chemists throughout their training. Student experiments always include steps to check apparatus and solutions.

Unfortunately the lesson does not seem to stick. Upon leaving school most chemists promptly ignore equipment calibration. They carefully standardize reagents but only rarely run laboratory cross checks. They know their results are correct, after all their instruments are as good as they were when new (ten years ago?). Many analysts go through their careers blithely assuming the values they report are the same (more or less) as those another laboratory would get.

A word about terminology -- When analysts speak of precision, they are referring to how well two or more results agree. Accuracy, on the other hand, refers to how closely particular results match the actual value being measured. In another context, precision relates to how tightly a gun groups and accuracy to how close it comes to the bulls-eye.

The subject of analytical precision came up at the 1989 USNC meeting in New Orleans. One of the authors was in attendance and innocently stated the industry needed a precision study to confirm suspected poor inter-laboratory agreement. All present agreed and promptly designated you know who to conduct the study.

OBJECTIVE

Experienced sugar chemists have long suspected the precision of routine results. There has been a feeling (reenforced by comments from the plant operators) that they were poorer than generally acknowledged. The goal of this project was to demonstrate the level of precision in sugar analysis across the industry. It was intended as a preliminary survey to set the stage for possible future studies. Hopefully the results obtained would also increase awareness within industry of the problem. The work was not intended to be definitive nor was it a formal collaborative study as defined by the AOAC.

In order for the results to be meaningful a few conditions had to be met. The test substances should be stable and uniform to minimize sampling errors. Analyses evaluated should be those performed every day in any sugar laboratory. People and equipment should be those used to obtain results routinely reported to customers and factory personnel. A final objective was to minimize interference with normal laboratory operations, to make the burden of testing as light as possible.

EXPERIMENTAL

The test material chosen was white sugar, something all of us should be able to evaluate with reasonable certainty. Solution color and conductivity ash methods were selected for evaluation. Both are fairly simple, objective procedures, some of the most straightforward tests in use. They are also both instrumental determinations. As an afterthought, gravimetric moisture was added, even though it is anything but a reliable procedure.

Hundred pound bags of white sugar were obtained from the production stream at two Amalgamated Sugar Company factories. These were as uniform and homogeneous as any material available. Sample sets were prepared containing approximately one pound of each in labeled plastic bags. These were distributed, along with written instructions and a report form, to 17 cane and beet sugar companies in the U.S. and Canada. In addition a set was sent to SPRI in New Orleans. Three beet sugar companies were sent multiple samples for distribution to their factory laboratories. A total of 34 sets were sent out.

The analysts were asked to use standard USNC procedures, copies of which were included with the samples. In keeping with the objective of the study, measurements were to be made by technicians or chemists routinely testing sugar and using normal testing equipment.

In order to insure against calculation errors, the report form asked for the raw instrument readings as well as details of the tests. The make and model of the instruments used was also requested.

Test sugars were not selected for any particular values. Nether were any attempts made to measure the absolute values being evaluated. Accuracy of results is a subject beyond the scope of this study. As it happened one sugar had a moderately high color, the other medium. Ash levels were on the low side and similar in both samples. Who knows what the moisture values were.

RESULTS

Response to the project was quite good. Of the 34 samples sent out, 28 reports were returned by 15 of the 18 organizations' contacted. Essentially all results were usable. Individual values are appended. Reporting laboratories are coded to preserve anonymity. The results are summarized here by determination.

1. Solution Color

Following modern practice, the USNC method calls for the use of adsorbance measurements. Percent transmittance is not dead, though. Eight respondents reported transmittance. Three sample cell path lengths were reported by respondents as follows:

Labs Using	Path (mm)
1	23
17	50
9	100

The question of optimum sample cell length for color measurement has long been debated among sugar analysts. The results below confirm that there is little or no difference between the two most popular path lengths.

Average Color
(ICUMSA units)

	Sample 1	Sample 2
50 mm Path	35.0	27.6
100 mm Path	35.3	27.4
Both Lengths	35.1	27.5

Range of Color Values and
Standard Deviation
(all path lengths)

	Sample 1	Sample 2
High Color	41	31
Low Color	31	24
Standard Deviation	2.7	2.2

Reporting of turbidity was somewhat confused. For this reason the values were not included here. As best we can determine, path length does appear to affect turbidity measurements. Typical turbidity values reported were 1 to 2 ICUMSA units for either sample when 100mm cells were used. Laboratories using 50mm cells typically reported zero turbidity. Bear in mind that all measurements were made on relatively simple instruments (B&L Spectronic or similar) with mediocre collimation of the light beam.

A number of different colorimeters were used. One basic instrument type was in common use though. The B&L Spectronic 70 was used by 8 laboratories and 10 others used similar models (mostly variations on the Spec 70's simple optical bed) by B&L or

its successor, Milton Roy. The popularity of this type of instrument is due to its ability to handle long path cells and relatively low cost (about one third of the next, more sophisticated instrument type available).

2. Ash

On casual inspection, precision of the ash determination looks fairly good. This is somewhat misleading because of the small values involved. The precision is actually about half of that for color.

Conductivity Ash
(percent)

	Sample 1	Sample 2
Average	.0087	.0079
High value	.0114	.0100
Low value	.0043	.0056
Standard Deviation	.0017	.0014

Of note is the close agreement between the two samples when run in a given laboratory. No laboratory reported differences greater than .0025 and many reported the same value for both samples. On the other hand, there was very poor agreement between laboratories.

Conductivity instrumentation used was of two general types. The most popular being the classic manually balanced bridge, Leeds and Northrup or Beckman (Industrial Instruments) models were the older bridges and Yellow Springs Instruments being the popular newer models. A number of direct reading digital conductivity meters were also used, nearly all being Cole Parmer 1482 or the essentially identical Amber Scientific 1052. No details of instrument calibration were recorded.

Quite a bit of deviation from the USNC method's solution concentration was observed. The method calls for a concentration of 25g/100ml. About half of the respondents used that value. Other concentrations used ranged from 10 to 31 grams per 100ml. This may or may not have a significant effect on results.

3. Moisture

Moisture determinations, as expected, had the widest range of values and poorest precision. The CV for moisture was nearly 50 percent compared with 8 percent for color and 19 percent for ash. Some of the variation could be due to sample handling and storage conditions. Most of it is, more likely, due to shortcomings inherent in the test.

Moisture Values
(percent)

	Sample 1	Sample 2
Average	.015	.015
High value	.036	.027
Low value	.004	.002
Standard deviation	.006	.007

The very small quantity of water in white sugar presents a great challenge to the analyst. Ten grams of white sugar typically contains less than three milligrams of moisture, a weight difficult to measure precisely, given typical analytical balances and the vibration found in many sugar laboratories. The problem is compounded by dry sugar's high affinity for moisture and its ability rapidly adsorb water from the environment.

Most laboratories used the standard 10 gram sample size. A few are moving to larger samples, in the 20 to 50 gram range. The heavier sample reduces weighing errors but can cause problems cooling the sample after removal from the oven.

Drying equipment was evenly divided among vacuum ovens, convection ovens, and forced draft models. Vacuum ovens were operated at 75-85°C, the others at 100 to 105°C. For the most part, drying times were in the 3 to 4 hour range called for in the procedure. There was considerable variation in cooling times, ranging from 15 to 60 minutes. Experience has shown cooling time to be a critical variable in sugar moisture work.

Calcium sulfate (Drierite) was the most popular desiccant, silica gel was used by a few analysts. One laboratory used phosphorus pentoxide. No correlation could be seen between desiccant type and moisture results.

Of particular interest was one laboratory, which apparently ran the sugar moisture on an infrared moisture balance. This is a technique normally used for high moisture samples like pressed and dried beet pulp and everyone knows won't work for sugar. Surprisingly they got fairly good results, .019%, in only three minutes. More work in this area is obviously warranted.

DISCUSSION

One valid question is just how much analytical precision is needed or required. We might be justified in assuming the precision observed here is acceptable by the industry. It is the level at which we're currently operating and most involved seem satisfied (or maybe just blissfully ignorant). While the spreads noted are wider than we would like to see, results are apparently acceptable.

The economic impact or cost benefits ratio of analytical results may be an overlooked item. When values are well within specifications the cost of errors is not too great. However, when results approach specification limits large costs can be involved.

Let's examine the color of sample #1 for example. Many sugar users specify a 35 RBU color limit. Most of us have a stated maximum color level above which we divert production to remelt. For the moment let's assume this is also 35 RBU. Sample #1 averaged 35.1 RBU color. Most laboratories would ship or accept 35.1 sugar. One laboratory in our study measured the color of this sample at 41, two others got 38. Sugar at 38 or 41 RBU would probably be diverted to a less color sensitive customer, it might even be remelted. Now, significant (and in this case needless) costs or delays enter the picture.

Another point raised by this study deals with meeting customer specifications. It's obvious that a producer wishing to avoid problems must err considerably on the safe side and run to very tight quality standards at added cost. Also what's your reaction when a customer wants to reject an out of specification shipment? Do you wonder if maybe it really is out and was missed by poor precision in your laboratory? Maybe you mentally question the customer's analysis, knowing how far off your's can be.

If we can improve precision, it possible we can relax production specifications. We will most certainly increase confidence in results.

What kind of precision is practically obtainable, given industry facilities and personnel. Like many sugar companies Amalgamated uses an in-house check sample program to evaluate their laboratory results. It is a typical round-robin test. A common sample is distributed to each laboratory. Five individuals at each location run the sample, the four technicians routinely doing the test and the supervising chemist. Comparison of results within a given location give an indication of technician skill and comparisons between laboratories serve as a calibration check.

In the case of color our check tests have standard deviations typically less than 1.0 RBU within a single laboratory and 1.5 to 2.0 RBU across the company. From this we might conclude the industry wide precision could be improved 25 to 50 percent in the case of color measurements. Our experience with ash measurements indicates results could be much better than those observed in this study. Our in-house results have standard deviations of about .0003% within a single laboratory. Unfortunately our standard deviations between laboratories are about the same magnitude as this study, .0015%, leading us to question the procedures used to standardize conductivity equipment. As for moisture, Amalgamated does not even try to run check samples for sugar moisture.

The authors feel technician skill level may have been a factor affecting this study's precision. Quite a bit of the analytical

work was done by bench technicians (as was intended) and confusion arising from following unfamiliar procedures appears to have caused problems in a few cases. Practice testing with samples of published values and pretest training might have improved precision somewhat.

RECOMMENDATIONS

In its role of a preliminary survey this project illustrates a lack of precision across the industry. It points out the need for a continuing exchange of samples. Additional work is needed, particularly to develop an industry wide program of laboratory evaluation or a supply of standard reference materials.

We also need further work on specific methods. The most obvious need is for a better means to monitor moisture in sugar. The ten to fifteen fold range of values reported demonstrates the present method's shortcomings. A viable alternative might be the Karl Fischer method, particularly with some of the modern titrimeters now available. At the very least, gravimetric method test conditions need to be rigorously defined and then followed in practice. Particular emphasis should be given to the steps between removal from the oven and final weighing.

Conductivity ash measurement, potentially a very reliable and reproducible determination, was somewhat disappointing. As noted, equipment standardization procedures might be the place to start improvements. Another point to examine would be conductivity cell care and maintenance practices.

Color proved to be the most precise measurement studied. It is our feeling results should have been better though. One problem in need of study is a comparison between the photometric and scattering errors of 50mm vs. 100mm path cells. The simple colorimeters in common use do not have very good light beam collimation. It may be that what we gain with the longer path length we loose in scattering error.

The study also demonstrated the need for more precise and detailed instructions in future studies of this type. Despite instructions to follow USNC procedures and fill in the values requested on the report forms, there was quite a bit of deviation. It appears that many laboratories followed their own standard procedures and recorded the data in their own format. These deviations usually had little impact on final values reported as most laboratories use appropriately scaled calculation multipliers. They did cause some confusion and made checking results difficult, however.

This study could not really address the question of instrument calibration. A project to do so would be worthwhile. Our in-house tests have shown that the variation between locations is frequently over twice that observed within a single laboratory.

CONCLUSIONS

This project has demonstrated that, at present, there is considerable variation between laboratories running replicates of the same sugar sample. Some work within individual laboratories (instrument maintenance, personnel training, etc.) might result in improved precision. At the corporate level, an ongoing program of inter-laboratory comparison should be part of every company's quality assurance program. It appears that instrument calibration is a major factor in the variation observed and checks between laboratories is the simplest way to detect such problems. Finally, the industry might benefit from a program to establish a source of standard reference samples of known values.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the contribution of the numerous sugar technicians whose analytical work provided the data reported in this project. The support of the participating sugar companies is also greatly appreciated. We would also like to thank the USNC for encouragement to conduct the study and permission to present it here.

USNC 1989 SUGAR TESTING ROUND ROBIN
SUMMARY OF COLOR, ASH, AND MOISTURE ANALYSES

Location	Color (RBU)		Ash (%)		Moisture (%)	
	Sx #1	Sx #2	Sx #1	Sx #2	Sx #1	Sx #2
A	36.7	29.9	0.0100	0.0090	0.012	0.011
B	36.2	29.4	0.0100	0.0100	0.013	0.012
C	37.0	29.8	0.0100	0.0100	0.024	0.027
D	29.0	24.0	0.0087	0.0068	0.015	0.013
E	34.0	26.0	0.0114	0.0087	0.022	0.024
F	36.0	29.0	0.0066	0.0067	0.018	0.021
G	38.0	30.0	0.0088	0.0078	0.012	0.011
H	33.0	26.0	0.0070	0.0070	0.019	0.022
I	34.3	26.3	0.0087	0.0077	0.007	0.007
J	36.0	28.0	0.0076	0.0062	0.020	0.023
K	36.7	29.3	0.0071	0.0068	0.027	0.022
L	31.0	21.5	0.0081	0.0073	0.036	0.027
M	41.0	30.0	0.0088	0.0076	0.015	0.010
N	35.0	27.0	0.0080	0.0070	0.016	0.014
O	31.0	24.0	0.0086	0.0087	0.004	0.006
P	35.0	26.9	0.0080	0.0100	0.009	0.002
Q	32.0	28.0	0.0060	0.0060	0.011	0.013
R	31.0	25.0	0.0070	0.0060	0.012	0.012
S	38.0	29.1	0.0100	0.0100	0.012	0.008
T	34.1	27.5	0.0090	0.0070	0.017	0.020
U	36.0	27.0	0.0110	0.0091	0.006	0.006
V	34.5	26.6	0.0100	0.0090	0.011	0.008
W	36.4	28.8	0.0096	0.0085	0.027	0.020
X	36.8	29.5	0.0100	0.0090	0.010	0.011
Y	32.8	26.9	0.0080	0.0070	0.016	0.016
Z	38.9	24.9	0.0111	0.0092	0.013	0.009
AA	38.0	31.0	--	--	0.019	0.019
BB	34.0	28.0	0.0043	0.0056	0.022	0.027
Mean	35.1	27.5	0.0087	0.0079	0.015	0.015
Std. Dev.	2.74	2.24	0.0017	0.0014	0.006	0.007