International Diffuse Reflectance Conference 2016

Chemometrics ShootOut Rules
Traditional Format

This year's shootout is quite different from previous competitions. For the first time, a common set of wheat data from three different NIR instrument manufacturers will be analyzed. The objective is to evaluate the data with the goal of preprocessing the datasets to match the spectra from all three instrument manufacturers prior to developing a regression model for protein that results in equivalent results among the models as measured by the reproducibility.

We would like to thank the U.S. Department of Agriculture's Grain Inspection, Packers and Stockyards Administration and the instrument manufacturers that participated in the study for providing the data and Dr. Charles Hurburgh, Iowa State University, for facilitating the process.

The samples correspond to wheat grown throughout the United States with spectra collected on five instruments per NIR spectrometer manufacturer. The instrument models and instrument serial numbers have been coded. The spectra are in the range and spacing that their respective instrument manufacturers support. The reference protein results are on a 12% moisture basis.

An initial study undertaken by Iowa State University on behalf of the U.S. Department of Agriculture's Grain Inspection, Packers and Stockyards Administration yielded a reproducibility (standard deviation across NIR models) of 0.14% protein compared to an average standard deviation of 0.07 % protein across instrument copies of a given manufacturer. Reproducibility, as defined by the following equation

$$Reproducibility = \sqrt{\frac{\sum_{i=1}^{n}\left(\hat{y}_i - \bar{\hat{y}}\right)^2}{n-1}}$$

where n is the number of samples, will be used to judge the effectiveness of the calibration transfer method(s). Participants will be judged on the within and cross instrument manufacturer reproducibility for the various sets of data provided.

There are 1488 spectra in the calibration data set for 248 samples analyzed on three instruments for manufacturer A and three instruments for manufacturer B.

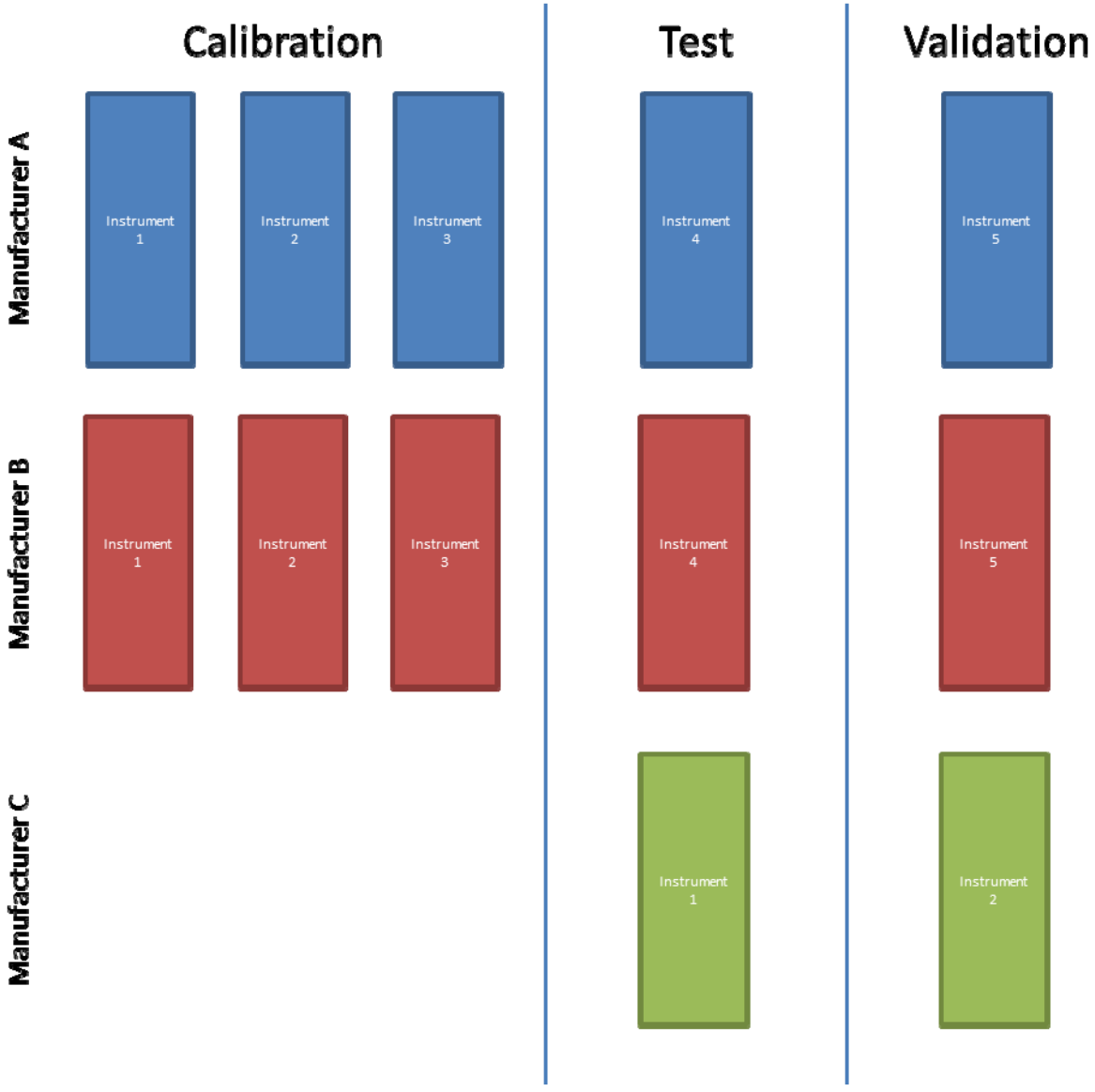
There are 744 spectra in the test data set for the same 248 samples as in calibration analyzed on a fourth instrument for manufacturer A and B in addition to one instrument for manufacturer C. The test set may be used to determine the preprocessing method but not be added to the calibration set for model development.

There are 450 spectra in the validation data set for an independent set of 150 samples analyzed once each on a new instrument for each manufacturer A, B, and C.

Reference protein values are provided for the calibration and test sets only.

The order of the Calibration and Test samples for all sets and instrument manufacturers is identical. However, the order of the Validation samples was randomized for each instrument manufacturer.

The figure below presents the available datasets.



Challenge description

This year's challenge will consist in developing the best model for the parameter provided using the calibration data. Because of the limited amount of information available, success in the shootout will rely on the participants' ability to build a model by relying only on their chemometrics skills, and not their knowledge of the data. However, the most important task will be to determine a method of pre-processing the data to minimize the spectral differences among instrument models so that a single calibration will yield both excellent accuracy and excellent reproducibility among "unknown" instruments of each manufacturer without secondary standardization. In addition, the quality of the presentation of the results and the reasoning behind the approach taken will be used to determine the winner. Participants are to:

1) Develop a preprocessing method to match the spectra from all three NIR instrument manufacturers to be used in a common calibration set.
2) Develop the best possible model for protein on the calibration set.
3) Test their model on a test set (we provide reference values).
4) Predict a validation set (we do NOT provide reference values).
5) Detail the reasoning when selecting pre-treatment methods, regression method, and number of latent variables.

It is explicitly prohibited to directly include samples from the test set in calibration in order to predict the validation set. However, information from the test set can be used to "tune" the calibration model through the use of standardization files, to derive shapes to perform orthogonalization, etc.

Participants who wish to compete for prizes **must submit** their predictions of the calibration, test and validation sets **by July 30, 2016** in an EXCEL file (or equivalent spreadsheet file) to: idrc.shootout@cnirs.org

Criteria for deciding winners include: (1) Prediction and reproducibility statistics of the test and validation sets, (2) novelty, uniqueness, and clarity of the presentation, (3) timing (staying within presentation time assigned), and (4) quality of answers to questions from the audience. An audience vote will be taken and the results of this vote will be considered by the judges for determining the winners.

Winners will be announced during the banquet on Thursday night. Prizes this year will be as follows: 1st Prize: $200, 2nd Prize: $100, 3rd Prize: $50. Decisions of the judges are final.

To ensure consistency among participants, judges will calculate the following statistics from the calibration, test, and validation sets:

1. Coefficient of determination
2. Root mean square error of calibration/cross-validation/prediction
3. Standard error of calibration/ cross-validation/ prediction
4. Bias of calibration/ cross-validation/ prediction
5. Reproducibility across all three NIR manufacturers.

To determine test set statistics, judges will use the following definitions of the above terms:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y} - y)^2}{n}} \qquad SE = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y} - y)^2 - \frac{(\sum_{i=1}^{n}\hat{y} - y)^2}{n}}{n-1}} \qquad Bias = \frac{\sum_{i=1}^{n}(\hat{y} - y)}{n}$$