FOSS

A White Paper from FOSS

Dynamic Sub-sampling™

Limiting the number of sub-scans for homogeneous samples can improve speed of near infrared analysis without compromising accuracy.

By: Robin Malm and Barbara Tobijaszewska

P/N 1026748, Issue 1, September 2014

Dedicated Analytical Solutions

Introduction

The use of Near Infra-red Transmittance analysis techniques as used by the FOSS Infratec™ 1241 has become a standard for whole grain analysis at the point of receival throughout the grain supply chain. One key to the success of this technique has been the speed of the analysis allowing a sample to be analysed in around one minute. A second success factor has been the ability to analyse a relatively large sample compared to alternative techniques. This is achieved through analysis of multiple subsamples combined with mathematical averaging of results to provide a robust prediction of the results.

Increasing the number of subsamples has been shown to improve the repeatability of a measurement. Conversely, reducing the number can have the opposite effect. In other words, the effect of sub-sampling with homogeneous samples is negligible whereas with inhomogeneous samples, the sub-sampling is critical.

What then if the grain analyser could be programmed to limit the number of sub-scans when handling homogeneous samples, thus improving speed of analysis without compromising accuracy? This goal formed the background for the development of Dynamic Sub-sampling.

Repeatability versus number of sub-samples

Sub-sampling and scanning are one of the major factors in the overall time of analysis of an instrument such as the Infratec. Any reduction in the number of sub-samples taken can reduce overall analysis time. However, any such improvement in speed must not compromise overall accuracy or more specifically a reduction in repeatability.

Instruments such as the Infratec™ 1241 have historically been used with a fixed number of sub-sample measurements, typically 10, of each sample. Figures 1 and 2, show the repeatability versus the number of subsamples. The repeatability was determined as the standard deviation between ten repeated measurements of the sample. It can be seen that sample 1, a highly heterogeneous sample, has a poor repeatability i.e. high standard deviation when using only five sub-samples. With this sample, increasing the number of sub-samples will improve the repeatability for both protein and moisture quite substantially. With sample 2, a homogeneous wheat sample, a good repeatability can be seen after five sub samples. In this case, increasing the number of subsamples will not give any significant improvement in repeatability for neither protein nor moisture.

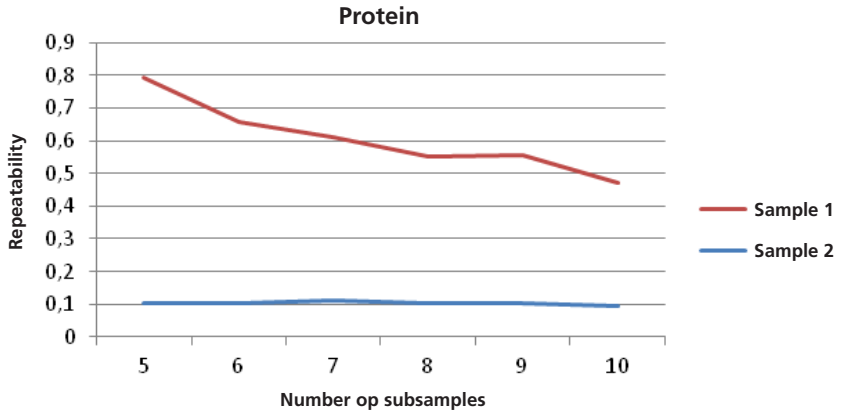

Figure 1. Repeatability versus number of sub samples for protein

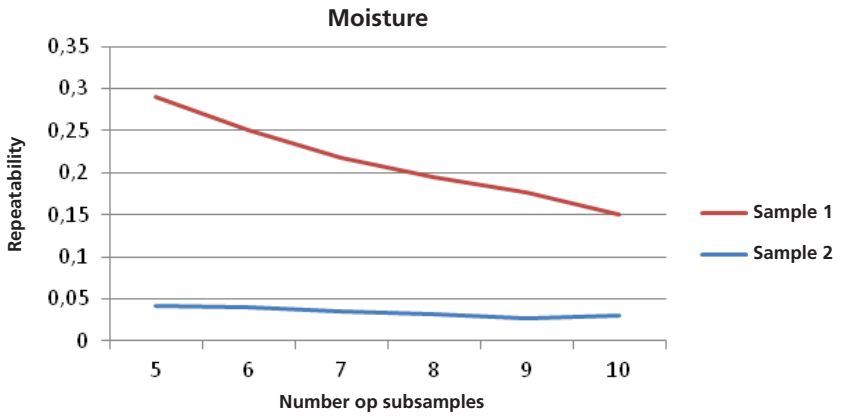

Figure 2. Repeatability versus number of subsamples for moisture

This discussion of sub-sampling versus repeatability shows the basis for reducing sub-sampling for homogeneous samples without impacting accuracy.

Dynamic Sub-sampling

Based on the repeatability versus the number of sub samples relation, the Dynamic Sub-sampling algorithm has been introduced for the Infratec™ NOVA. It works as follows:

- A preset minimum number of sub-samples is measured
- The standard deviation of the sub-sample predictions on selected parameters is calculated and compared to set limits. Then;
  - if the standard deviation is below the set limit and no outlier warnings present, the measurement is interrupted and the result is presented.
  - if any standard deviation is above the set limits or any outlier warnings present, the measurement is continued until the maximum number of subsamples is measured.

Setting the limit

The standard deviation between subsamples determines whether Dynamic Sub-sampling should be applied is. Standard deviation between subsamples is also used as an outlier measure, i.e. the C-outlier on Infratec 1241. However, for Dynamic Sub-sampling, the limit needs to be set lower than for the C-outlier measure in order to maintain good repeatability. As a rule of thumb, the limit for Dynamic Sub-sampling is set to about one third of the limit for the standard deviation used as an outlier measure.

Examples of applying Dynamic Sub-sampling

In the following, two examples of the effect on repeatability and accuracy by applying Dynamic Sub-sampling are shown. In these examples, the standard deviation between protein predictions was used to determine whether Dynamic Sub-sampling was applied or not. However, both protein and moisture were evaluated. The maximum number of subsamples was set to ten and minimum number of subsamples to five. The standard deviation limit was set to 0.3. This limit resulted in Dynamic Sub-sampling applied on about 70% of the measurements.

Repeatability

One wheat sample was measured ten times on each of the ten instruments. The repeatability was determined on each instrument as the standard deviation between the ten measurements. In Figure 3, the mean repeatability applying Dynamic Sub-sampling (With DS) and not applying Dynamic Sub-sampling (No DS) are shown. It can be seen that the repeatability is only slightly affected for protein when applying Dynamic Sub-sampling.

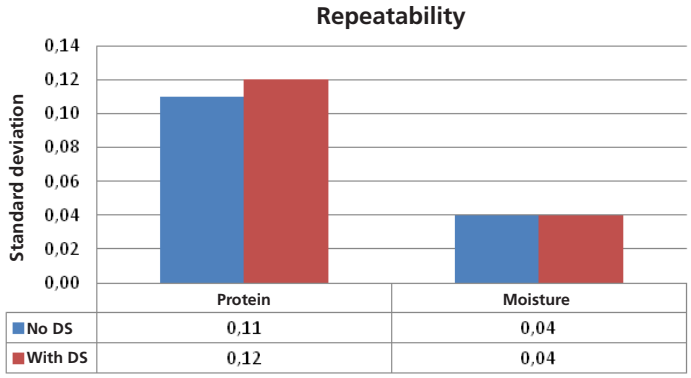

Figure 3. Impact of Dynamic Sub-sampling on repeatability.

Accuracy

To test the effect of Dynamic Sub-sampling on accuracy, data from 80 wheat samples measured on ten Infratec™ NOVA was used. As a reference, the predictions on ten Infratec 1241 instruments not applying Dynamic Sub-sampling was used.

In Figure 4, the mean Infratec 1241 predictions versus the Infratec NOVA predictions for all instruments are shown for protein and moisture not applying Dynamic Sub-sampling (No DS) and applying Dynamic Sub-sampling (With DS). It can be seen that the accuracy is only slightly affected for both protein and moisture resulting in a small increase in SEP values.

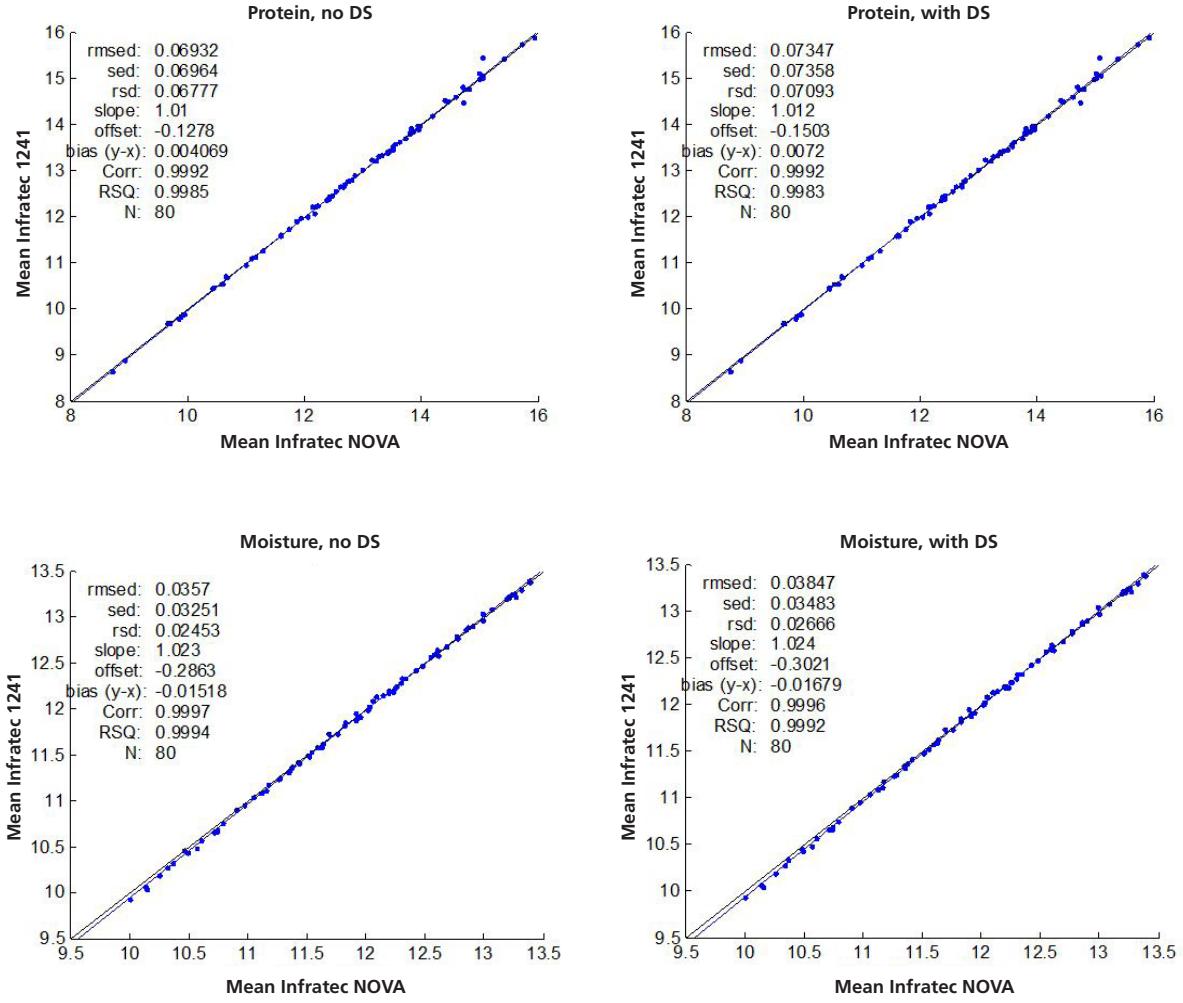

Figure 4. Accuracy for wheat not applying and applying Dynamic Sub-sampling.

Conclusions

While the speed of analysis offered by modern grain analysers such as the Infratec™ 1241 has been a significant improvement over previous techniques, speed remains a critical element of efficient grain receival and any gains that can be made can improve overall efficiency of the operation. With sub-sampling being one of the contributors to overall analysis time, the opportunity to reduce these where possible without impacting accuracy is a useful improvement that can add value to a grain company.

Dynamic Sub-sampling can be applied on Infratec NOVA in order to decrease measuring time on homogeneous samples without decreasing performance of prediction results.

FOSS

FOSS
Foss Allé 1
DK-3400 Hilleroed
Denmark

Tel.: +45 7010 3370
Fax: +45 7010 3371

info@foss.dk
www.foss.dk