FOSS

A WHITE PAPER FROM FOSS:

Best Practice in Sample Preparation for Calibration Development with EyeFoss™

ANALYTICS BEYOND MEASURE

June 2019. GB
fossanalytics.com

Contents

Introduction

EyeFoss™, developed by FOSS, is a dedicated image analysis instrument for grain defects and contaminants. Image analysis for grain, especially applied for payment purposes at grain receival to remove the subjective nature of visual testing, has been something of a "golden dream" in the industry. In the past 20 years, there have been many attempts to build such an instrument with varying levels of success. EyeFoss is the first such instrument to achieve success, being the first actively used for grower payment in a grain receival setting.

However, calibration development remains a key requirement for a product such as EyeFoss, as it is very difficult to create global calibrations. After all, while most people in the grain industry agree that fusarium-affected wheat or barley is undesirable; there is less consensus about what actually constitutes a fusarium-affected kernel.

Creating the right samples for calibration development and validation is the first step of a calibration development project. Derived from the experience gained by FOSS and our partner customers in pilot projects conducted to date, this whitepaper describes the best practice in this area.

EyeFoss Calibration Development Process Summary

1. The customer will collect and measure the **calibration development material** with EyeFoss for the products and parameters (i.e. foreign objects, damage) defined by the project. The customer is required to assess the calibration material using existing reference methods.

2. The customer will transfer the **image data and reference data** to the FOSS Innovation Centre (FIC) in Denmark. FOSS calibration experts at FIC utilise this data to developed **calibration models**.

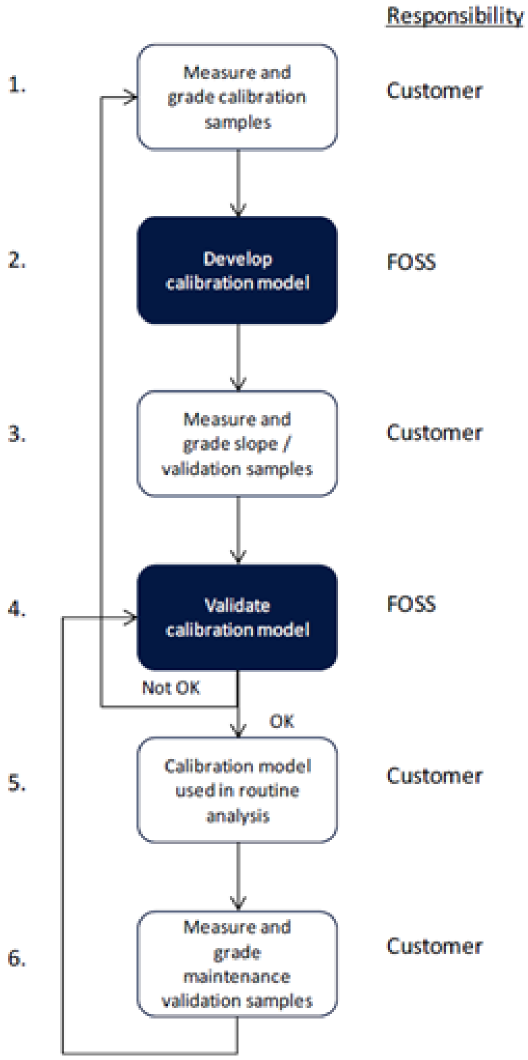
3. The customer prepares and conducts full reference analysis on **slope and validation samples**. These samples may contain multiple parameters. The slope samples define the slope correction needed between the EyeFoss reading and the reference method. The validation samples establish the performance of the EyeFoss predictions versus the reference method.

4. FOSS together with the customer validates the **calibration model** and jointly decides if the validation criteria (acceptance criteria) are met. One or more additional calibration loops (consisting of steps 1 to 4) may be required to achieve the desired performance.

5. The calibration model is installed with the slope defined in step 4 and is used in routine analysis.

6. The customer will be required to collect, measure and conduct reference analysis on maintenance validation samples on an ongoing basis, especially for the new season's grain. The validation may result in updates due to the appearance of new biological material.

For some of the common products and parameters, calibration data may already exist. This may be a useful starting point for calibration development if available and appropriate. In this case, the development process starts with collecting slope and validation samples in step 3 to assess the suitability of pre-existing data for purpose. It is unlikely that this data would be sufficient to create a suitable calibration model without at least some calibration development loops.



Sample Types:

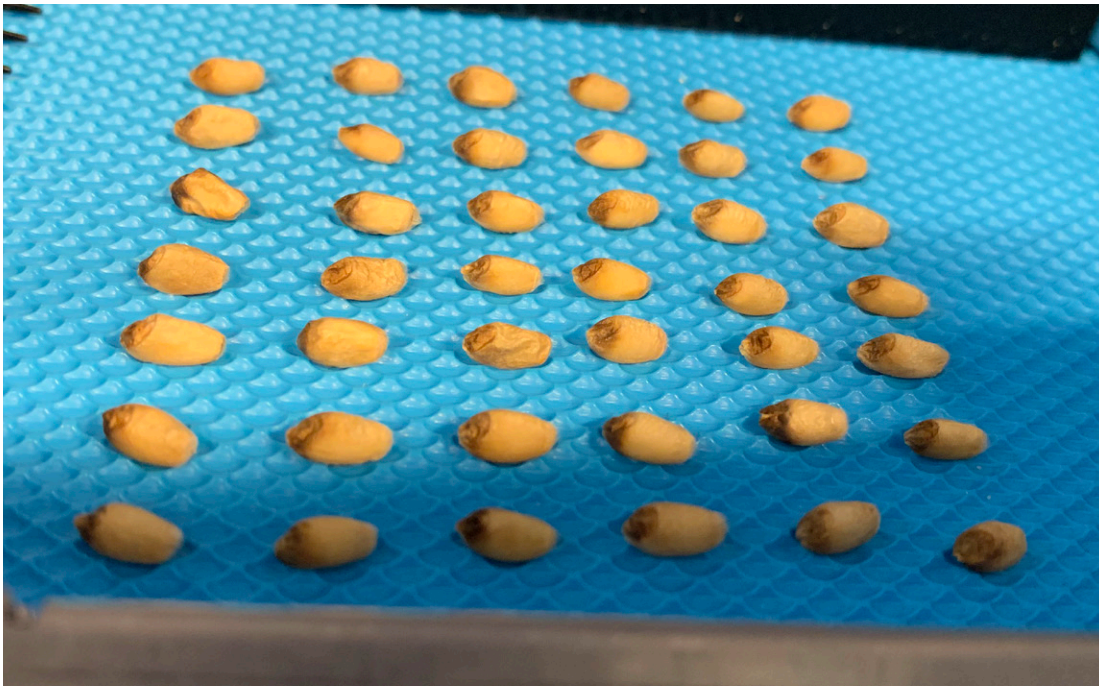
Calibration Samples:

There are two types of calibration samples; parameter samples, i.e. those affected by the parameter being targeted and non-parameter samples. In other words, the non-parameter calibration samples consist of sound and commercial quality material of the target grain type, for example, Wheat, Barley or Durum. Calibration material consists of single-kernel samples representing each of these two types of calibration samples. Typically, you should collect 5,000 to 10,000 sound kernels (non-parameter samples) and 3,000 of each of the defective kernels or foreign objects (parameter samples) for calibration development purposes. It may be a requirement to collect further samples for additional calibration development loops especially if problems occur in the development process. A smaller number of samples that cover the full range of the visual appearance of the parameter will be more useful than a large number of kernels all ap¬pearing the same.

A key to the creation of robust calibrations is to ensure that ranges of types of each defect or foreign object are included in the material. For example, taking the example of moldy wheat, calibration material should not only consist of highly moldy wheat, but should also include wheat kernels that, while considered moldy against analysis standards, have only limited discoloration.

Likewise, in the sound wheat sample it is important to include kernels that cover the range of visual representations of this product. For example, wheat, when subject to weather conditions that might lead to sprouting may become somewhat discolored and while still considered as sound wheat. As a result, the sound kernels in a sample containing sprouted grains may appear darker in color than wheat that has not been subject to any excessive moisture prior to harvest. To avoid falsely identifying defective kernels in a sample (false positives), it is critical to include these types of samples in calibration materials.

For calibration development, it is a requirement that it is possible to establish the identity link between the image and the corresponding parameter on single-object basis. Kernels may be scanned in bulk for some of the parameters chosen for the test (e.g. oats) as the same reference tag can be assigned to each of the objects individually when viewing the object in the calibration development tool. For other parameters (e.g. skinned barley or sprouted wheat), the kernels must be scanned individually with the affected part of each kernel facing the digital camera system.



As a minimum, a calibration created from material from the particular parameter (e.g. skinned) and the corresponding non-parameter or sound product (e.g. not-skinned). Non-parameter material can normally be scanned in bulk and can be used as non-material for several parameters provided it does not display the traits from any of the parameters themselves.

The magnitude of material required for a particular calibration is difficult to predict but will depend on how easy it is to discern the trait of the parameter from the corresponding non-parameter. For the more difficult parameters the general rule is that the more material the better the chances of a well performing calibration. Additional samples may be required if additional development rounds are required.

Slope Samples:

Slope samples are used for applying a slope and intercept correction to a calibration model. Such a correction is always necessary if the required output unit is different from the basic EyeFoss unit, (e.g. the raw EyeFoss output would be as a percent by count, while some companies require a conversion to percent by weight). In addition, a slope and intercept correction may help correcting for would-be imperfections in the calibration model.

Slope samples are normally full-sized samples of 500ml (required) which are visually graded and scanned in bulk. The range of the parameter in question shall at least encompass the range of commercial interest. The biological variance in the slope samples must not exceed that of the parameter- and non-parameter calibration material.

Twenty (20) 500ml slope samples are required per parameter. In some cases, it may be possible to combine several parameter phenomena in the same slope set. It is essential to conduct visual inspection of the full-sized slope samples (500ml).

Validation Samples:

Validation samples allow an independent validation of a slope-corrected calibration model for assessment against the acceptance criteria. While an independent set, they are of the same nature as the slope set.

Typically, twenty 500ml validation samples are required per parameter. As with the slope sets, in some cases, it may be possible to combine several parameter phenomena in the same slope set and it is essential to conduct visual inspection of the full-sized slope samples (500ml).

How to prepare samples:

Calibration Samples:

Parameter samples are prepared simply by picking through samples collected over time to remove the required number of kernels representing each parameter. These should be stored separately and not mixed with samples of other parameters.

Non-parameter samples can be prepared in the opposite manner to parameter samples that is by removing any parameter material. High quality non-parameter material is as important as parameter material in the development of EyeFoss calibrations and as such, great care is required to ensure there is no material remaining that might be mistaken for parameter material.

So far as is possible, both **parameter and non-parameter** material should cover the following variations:

- A wide range of visual appearances of the parameter
- As many seasonal variations as can be achieved
- As many regional variations as can be achieved.

Typically, in the first year of a calibration development process, material is only available from one season and from a limited area. As such, it will likely be necessary to build additional seasonal and regional variation into the calibration in subsequent seasons.

Slope Samples and Validation Samples:

While slope and validation samples MUST be separate sample sets with each having a separate purpose, the method of preparation is the same.

Sample Size

It is very common practice around the world to conduct standard visual analysis on a rather small sub-sample. This is primarily a factor of time available for analysis, as large samples of grain can take a significant amount of time to analyse. Typical sub-samples analysed around the world might be 500 to 1000 kernels or 50g to 100g of sample depending on the region and local rules / customs.

One of the key benefits of EyeFoss is its ability to analyse larger samples within a short period. For example, while a well-trained human analyst may be able to process a 100g sample for defects and foreign objects in 3 to 5 minutes, EyeFoss can analyse 4 times that in a similar amount of time. The statistics of sampling and the ability to get a more accurate analysis with a larger sample in a similar amount of time is one key to success with EyeFoss. However, in order to create high quality calibrations, it is important to use similar sized samples for slope and validation. This includes not only the analysis scan conducted on EyeFoss but also the reference work done by the human. While preparation of these samples will take considerable time, we have found that stronger calibrations quickly reward this effort.

A standard EyeFoss sample is 500ml in size. However, if such a volumetric measure is not common in the customer's local grain area, it is just as satisfactory to use a weighed sample. We recommend 400g for wheat and durum and 350g for barley, as these would roughly equate to a 500ml sample for standard test-weight product.

Obtaining Sufficient Material

A key challenge in preparing slope and validation samples is obtaining sufficient material. For a slope and validation sample set to be effective, it is critical that there are samples that represent quantities of the target parameter that replicate those expected in routine samples. Clearly, in this case, the majority of samples will fall around the critical level (sometimes called the grading limit). For example, if a national grain quality standard allows 1% sprouted kernels in wheat, a validation set of 20 samples would have a number of those samples around that limit. To make that set effective though it would also be important to have samples at the upper and lower levels typically seen. It is, however, not necessary to go to the extreme. There is little benefit from including samples above 20% sprouted for example.

The most effective way (in terms of impact on the calibration development) to collect sufficient material is to select samples that naturally have the amount of the target parameter needed. This is however not always possible, especially at the upper levels. Care however is required in "spiking" samples, although this does depend on the nature of the parameter.

For **contaminants** or **foreign material**, for example barley kernels in a wheat sample, spiking samples to obtain high-level samples is perfectly acceptable. The process is a simple one of adding barley kernels to a wheat sample to reach the desired level.

For **defect** parameters such as mouldy, sprouted etc., spiking is not an acceptable method. As noted previously in this whitepaper, many of the key defects we target occur when the grain experiences some level of elevated moisture either in the field or in storage. In these circumstances, even the wheat or barley kernels, which are considered sound in the sample, may be of a darker or different colour to kernels of the same variety that have not been subjected to these conditions. It would be very unusual for example to find 1% sprouted grains in a sample of pure, bright yellow wheat for example. Rather, the base colour of such a sample is likely to be a little "washed out" and darker perhaps with some levels of staining or darker brown discolouration, which in itself may be considered a defect in some markets.

The most effective means of preparing samples containing high levels of defect parameters then is a process of "reverse spiking" or concentration. In this process, select a larger sample containing a high level. The analyst would then remove sound kernels from the sample to obtain the desired concentration.

Mixed Samples

It may be desirable and in many cases, it is perfectly acceptable to have kernels representing multiple parameters in one slope or validation sample. To put it another way, it is not a requirement of EyeFoss calibration development that each sample contains only sound kernels and the kernels of the target parameter. For example, it would be perfectly acceptable for a wheat validation sample to contain sprouted kernels, barley kernels, mouldy kernals and fusarium. However, as spiking is not appropriate for defect parameters, it may be necessary to have multiple slope and/or validation sets to cover all the target parameters.

Reference Results

The final step in the preparation of slope and validation samples is to conduct reference analysis and to record the result. In the case of slope samples, compare these results against the EyeFoss predicted result to calculate any slope or intercept adjustments required. In the case of validation samples, these results are used to assess the performance of the calibration against the acceptance criteria.

Conduct reference analysis using the methods routinely employed in the region or area where the calibration is developed. This is important to ensure that EyeFoss calibrations best reflect the local customs and procedures for analysing grain. However, one key deviation from normal procedures is that it is essential for the performance of the instrument that this analysis is conducted on the full 500ml sample even in the event that normal procedures would require only a 50g or 100g sample (for example) to be assessed.

Conclusion

The sample sets described in this document form the basis of calibration development projects for EyeFoss. Care and attention paid to the preparation of these sample sets forms the basis for a successful project and the ultimate success of EyeFoss in a company's business.

Quick Guide to Sample Preparation:

Calibration Samples:

Parameter samples:

- 3.000* kernels (images) per parameter.
- Contains examples of the full range of appearances of the target parameter
- Prepared by individually selecting kernels from a range of samples.

Non-parameter samples:

- 5,000 to 10,000* kernels required.
- Prepared by removing any material of the target parameter.
- Care must be taken to ensure no kernels remain that may appear like any target parameter material.
- Kernels should be taken from multiple base samples.

**Additional material may be required in subsequent loops*

Slope Samples and Validation Samples:

- Two separate sets consisting of 20 samples each
- Sample Size: 500ml
- Each sample may contain more than one target parameter.
- Samples should cover the range of concentrations that are of commercial interest
- Contaminant samples may be spiked
- Defect samples should not be spiked but can be prepared by removing sound kernels to achieve higher concentrations.
- Reference analysis must be conducted on full 500ml sample.

Glossary:

Acceptance Criteria: The desired performance for a calibration, typically set at the beginning of a calibration development process. Acceptance criteria may be set for accuracy and/or repeatability and transferability.

Accuracy: A comparison between a result generated from a master EyeFoss and the reference result for that sample. Typically, accuracy is determined by assessing a group of validation samples.

Calibration Samples: Samples of the target parameter and samples of the non-parameter (i.e. the base grain) used to create calibrations.

Contaminant: Any kernel or object in the sample that is not of the base grain type. Also sometimes referred to as foreign material.

Crop Type:

Defect: Any kernel of the base grain type that has something wrong it with.

Grade: Grade is a term used to refer to the quality level of the product. For example, in a national or company quality standard, wheat or barley might be classified as "Grade 1" being the highest grade of that product attracting the highest price.

Grading: Grading is a term most commonly used in this context to refer to the act of sampling and analysing a sample of grain for the purpose of determining its grade.

Grading Limit: A grading limit is a level or amount of a parameter allowed in a sample before that sample would no longer fit into a particular grade. In relation to EyeFoss, grading limits are normally expressed as a percent by either count or weight and are usually a maximum amount. For example, a sample may be allowed to contain 1% sprouted grain to be acceptable in "Grade 1" and any level above that would see that sample be allocated to "Grade 2". A national or company quality standard may set multiple grading limits for a parameter. In the example given a standard may allow up to 1% sprouted in "Grade 1", up to 2% in "Grade 2" and so on. The first grading limit, where a sample would fall from the highest possible grade to the next is referred to as the **Primary Grading Limit.** This is the grading limit which is most commonly used to determine the acceptance criteria.

Parameter: A parameter is a type of defect or contaminant being targeted. For example, if the sample is wheat, it may be barley in wheat. Likewise, it may be sprouted or stained wheat. Parameters may be either a defect or a contaminant. A **non-parameter** sample is a sample or kernel of the base grain that does not exhibit characteristics of the parameter being targeted.

Slope samples: are used for applying a slope- and intercept correction to a calibration model.

Sound grain: Grain kernels of the main sample type which is considered of good quality and not effected by any defect to the extent that those kernels would be considered defective.

Validation Samples: are used as an independent validation set of a slope-corrected calibration model.

FOSS

FOSS
Nils Foss Allé 1
DK-3400 Hilleroed
Denmark

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

info@foss.dk
www.fossanalytics.com