Characterization of Fonio millet by NIR Hyperspectral Imaging

J. A. Fernández Pierna[1], Ph. Vermeulen[1], D. Stimant[2], B. Dupuis[2], P. Dardenne[1] and V. Baeten[1]

Walloon Agricultural Research Centre (CRA-W)
[1]Quality of Agricultural Products Department, Chaussée de Namur n°24, 5030 Gembloux, Belgium
[2]Section Systèmes Agricoles, Rue de Serpont, 100, 6800 Libramont, Belgium
fernandez@cra.wallonie.be

Introduction

Fonio millet (*Digitaria exilis* Stapf) plays an important role in food supply in West Africa. This crop is a small-grained, C4-metabolism cereal with a short life cycle. It is one of the first cereal crops to have been domesticated by West African farmers. It plays a central role in the emergence of traditional agriculture in the West African savannah, where it is now a staple food and an important part of the diet for several million people. This is true especially for the short-cycle varieties harvested at the end of the rainy season when granaries are empty. The cultivation of fonio millet in Conakry Guinea, Mali and Burkina Faso is indicative of its wide ecological adaptability.

The objective of this work is to define a methodology to investigate whether the NIR hyperspectral imaging method can be used for the characterization of the origin at the single kernel level, as well as regression of some typical parameters for fonio.

Methodology

Three experimental sites at Bareng and Kankan in Conakry Guinea, and Cinzana in Mali have been studied, which are representative of the distribution of fonio millet in West Africa. All the samples have been measured using a near infrared (NIR) hyperspectral camera. The instrument used is a MatrixNIR® Chemical Imaging System (Malvern instruments Ltd). It records sequential images with an InGaAs array detector (240x320 pixels) active in the 900-1700 nm range. A total of 76800 spectra are acquired by sampling area measured. In total, 174 samples have been measured in duplicate, driven to a total of 348 images. Each image contains between 150 and 200 seeds. In total 35833 spectra have been measured. From these samples, 89 are not peeled seeds and 85 are peeled seeds. The samples are grouped in classes according to their origin as indicated in Table 1.

Table 1 – Analysed samples according to the origin

Class name	Peeled seeds	Not peeled seeds
Bareng	35	42
Cinzana	10	12
Kankan	40	35

Selection of the about 100 more representatives seeds and extraction of their average spectrum

The mean spectrum of each grain is calculated by the application of a morphological mask obtained through a process of erosion in order to determine the contour of each grain. This mask is created by using binary images generated with the Isys software.

In order to reduce the number of kernels to about 100 in an image, a histogram (with the particle size distribution is created using a binary image. About 100 kernels containing a size distribution close to the mean are kept, as indicated in the Figures.

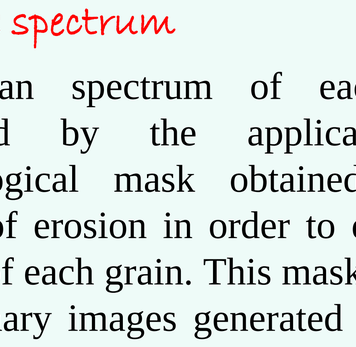

Figure 1 – Mask obtained for the image at 1300 nm

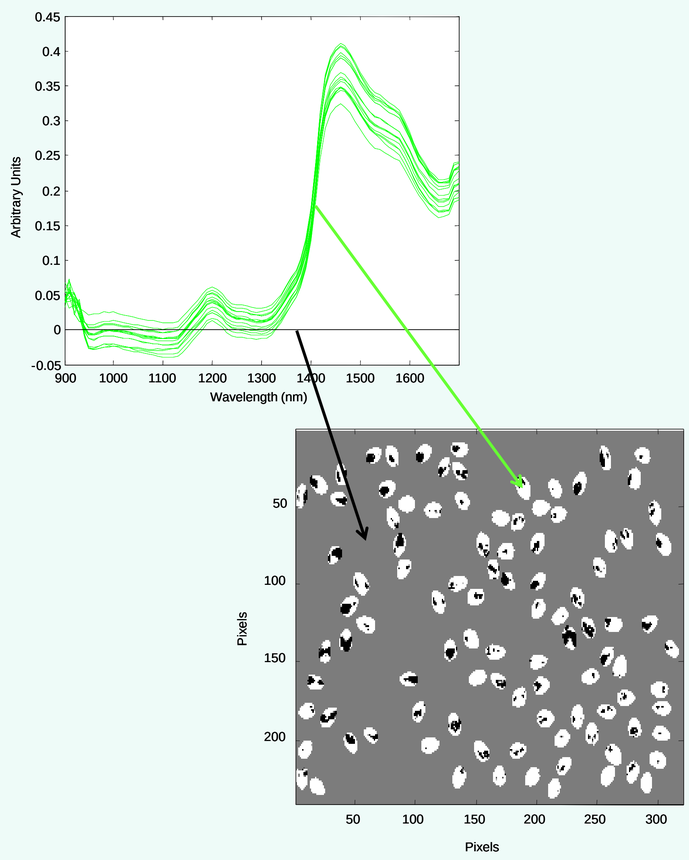

Figure 2 – Spectra from a single kernel of fonio after normalization and a representative spectrum from the background

Unsupervised study: PCA

Clear differences between peeled and unpeeled seeds can be found when performing a Principal Component Analysis (PCA) as indicated in Figure 3. When dealing only with unpeeled seeds, some differences appear concerning the origin of the samples (Figure 4).

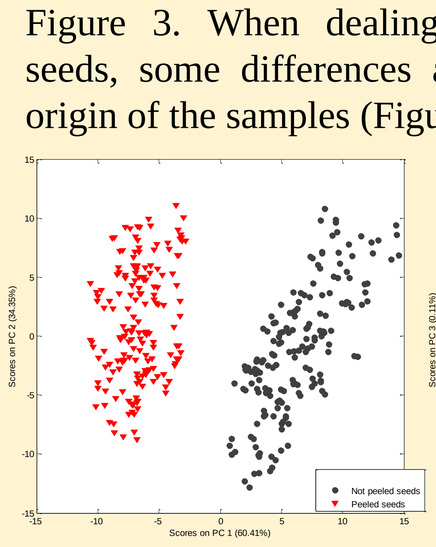

Figure 3 – PC1 vs. PC2 plot for all the spectra (unpeeled and peeled seeds)

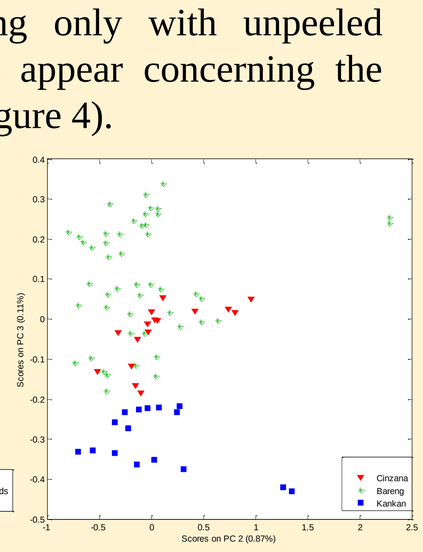

Figure 4 – PC2 vs. PC3 plot using the mean spectrum for each seed (unpeeled seeds)

Supervised study: discrimination models

This study includes the construction of individual models (PLS-DA) for each class in order to discriminate according to the origin of the samples. For these models, the validation procedure used is the leave-one-out cross-validation (LOOCV). Their performance is expressed in terms of the sensitivity and the specificity and the results are shown in Table 2.

Table 2 – Sensitivity and specificity for the three classes using PLS-DA after LOOCV

	Cinzana	Bareng	Kankan
Sensitivity	93.8%	91.7%	100%
Specificity	92.2%	100%	95.3%

Supervised study: regression models

This study has been performed on unpeeled samples, which corresponds to 8046 spectra (about 100 spectra by grain). The models have been constructed using PLS. and the data have been pre-processed by SNV (Standard Normal Variate) that corrects spectra for spectral noise and background effects. the validation procedure used is the leave-one-out cross-validation (LOOCV). The results for different properties (starch, dry matter, NDF, ADF, ADL ash and fat) are shown in Table 3 in terms of RMSECV and R^2.

Table 3 – Results for the different properties studied for the unpeeled samples

	R2	Latent variables	RMSEC	RMSECV
Starch	0.886	9	0.563	0.751
NDF	0.844	9	0.483	0.647
Dry matter	0.875	11	0.207	0.282
ADF	0.812	12	0.261	0.375
ADL	0.401	9	0.287	0.384
Ash	0.535	8	0.133	0.165
Fat	0.537	11	0.744	0.947

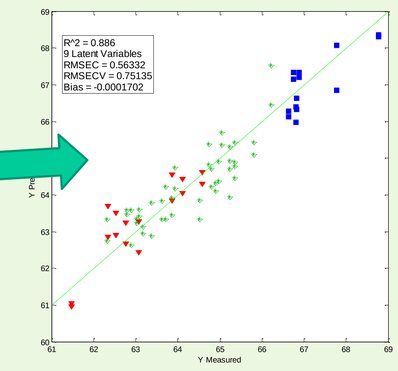

Figure 6 – Measured vales vs. Predicted values for the starch equation (unpeeled seed samples)

References

- ‘Genetic diversity and population diVerentiation of traditional fonio millet (Digitaria spp.) landraces from diVerent agro-ecological zones of West Africa’. H. Adoukonou-Sagbadja, · C. Wagner ,· A. Dansi , J. Ahlemeyer, · O. Daïnou, · K. Akpagana, · F. Ordon, ·W. Friedt. Theor Appl Genet (2007) 115:917–931
-'Hyperspectral imaging techniques : an attractive solution for the analysis of biological and agricultural materials'. V. Baeten, J.A. Fernández Pierna & P. Dardenne. In Techniques and Applications of Hyperspectral Image Analysis, pp. 289-311. (Hans F. Grahn & Paul Geladi Editors, John Wiley & Sons, Ltd.) 2007.
-‘NIR Imaging – Theory and applications’. J. A. Fernández Pierna, V. Baeten, J. Dubois, J. Burger, E.N. Lewis & P. Dardenne. In Comprehensive Chemometrics, volume 4, pp. 173-196 Oxford, Elsevier (Steve Brown, Romà Tauler and Beata Walczak eds.) 2009.

Walloon Agricultural Research Centre

RÉGION WALLONNE