Talanta 121 (2014) 288–299

Contents lists available at ScienceDirect

Talanta

journal homepage: www.elsevier.com/locate/talanta

Review

Limitations and current applications of Near Infrared Spectroscopy for single seed analysis

Lidia Esteve Agelet*, Charles R. Hurburgh Jr.

Department of Agriculture and Biosystems Engineering, Iowa State University, USA

ARTICLE INFO

Article history:
Received 7 October 2013
Received in revised form
15 December 2013
Accepted 16 December 2013
Available online 6 January 2014

Keywords:
Near Infrared Spectroscopy
Single seed
Single kernel
Technologies
Quantification
Discrimination

ABSTRACT

Near Infrared Spectroscopy (NIRS) analysis at the single seed level is a useful tool for breeders, farmers, feeding facilities, and food companies according to current researches. As a non-destructive technique, NIRS allows for the selection and classification of seeds according to specific traits and attributes without alteration of their properties. Critical aspects in using NIRS for single seed analysis such as reference method, sample morphology, and spectrometer suitability are discussed in this review. A summary of current applications of NIRS technologies at single seed level is also presented.

Contents

1. Introduction

Plant breeding facilities are constantly looking to improve current varieties and to obtain new seeds with special traits. This is achieved with a careful selection of the best individual traits. Use of bulk samples in the selection process results in a larger seed production with only a fraction with the desired trait, since the heritability of a desired characteristic may be low. Analyzing individual seeds allows researchers to understand the future plant characteristics and the characteristics of its next generation [1], while retaining competitive agronomic performance, or obtaining high yields or resistance characteristics. Common Near Infrared Spectroscopy (NIRS) bulk sample analyzers provide measurements of samples of about 250 g of kernels on average. Single seed differences cannot be identified and no discrimination is possible.

Besides seed producers, farmers, feed processors, animal producers, food companies, and other seed-related industries can benefit from on-site single seed screening as well. The so-called 'dilution effect' of current analytical technologies allows low fractions of unwanted seeds to be mixed with the majority without the chance of identifying the impurity fraction, decreasing the overall batch value. For this reason seed inspection is very relevant for pricing commercial grains, as the undesired fraction of seeds is visually determined at single seed level. For processing and quality improvement purposes, NIRS analysis at the single seed level

* Correspondence to: Food Science Building, Iowa State University, Ames, IA 50014, USA. Tel.: +1 515 294 8319.
E-mail address: esteve.lidia@gmail.com (L. Esteve Agelet).

0039-9140/$ - see front matter © 2014 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.talanta.2013.12.038

followed by a sorting mechanism could help in increasing sample uniformity and purity. Whether the target is to segregate sound seeds from defective/damaged, to keep seeds with a specific concentration of a compound of interest, or to discriminate mixtures of varieties, the quality and economic value of a seed batch could increase considerably when the unwanted fraction of seeds is removed and the batch is uniform in its attributes [2].

This review gathers relevant aspects of seed and grain analysis by NIRS as a tool to quantify, segregate, and discriminate seeds on a fast and non-destructive manner. Current and promising applications are exposed together with limitations and challenges faced by the technology. Although NIRS is already well-known for successfully analyzing bulk grain and bean samples, Single Kernel NIRS (SKNIRS) still has to reach its full potential for industrial applications. SKNIRS needs to have recognized protocols and methods as bulk NIR analyses currently have.

2. Overview on near infrared spectroscopy

It has been over 60 years since the first practical application of NIRS as an analytical method. Karl Norris, pioneer of NIRS, developed the first applications of NIRS on grains and seeds in the 1960s [3,4]. Since then, instrumentation, statistical methods, and software have been improving and the number of applications have exponentially grown. NIRS is now a mature analytical method for grains and seeds, recognized by the American Association of Cereal Chemists (AACC 39-00) and the American Oil Chemist Association (AOCS am 1–9).

Near Infrared spectroscopy (NIRS) technologies have a performance comparable to other wet chemistry analytical methods, but with some important advantages such as short analysis time, small sample preparation, and non-destructiveness. The radiation from the near infrared (NIR) electromagnetic region (700–2500 nm) is absorbed by water and organic compounds such as carbohydrates, protein, oil or alcohols. The apparent absorbed energy by a sample, calculated from either transmitted or diffusively reflected radiation, can be related to the content of the compound. Shorter wavelengths, close to the visible region, are weakly absorbed compared to the longer wavelengths closer to the infrared region. For this reason, shorter wavelengths can penetrate deeper through samples that are not excessively thick and opaque. Fraser et al. [5] showed that, for apples, wavelengths up to 900 nm could penetrate up to 25 mm, while from 1400 to 1600 nm the penetration decreased to 1 mm.

Bulk sample NIR instruments working by transmittance mode mainly work on the region from 700 to 1100 nm. Those instruments measure the transmitted radiation through a fixed pathlength of a bulk sample of grains or beans, assuming that the decrease of the initial radiation in traveling through the sample is due to absorption. The pathlength is optimized according to the commodity being measured and the instrument setup, being common a pathlength around 15 mm for corn and soybeans. Instruments based on reflectance mode, on the other hand, measure the diffusely reflected radiation from the sample. The diffuse reflected signal is a fraction of the initial radiation source which after penetrating the sample few mm, has been interacting with the sample molecules, scattered in several directions, and traveled back to the surface. Only the diffuse fraction of the reflected radiation has interacted with the compound of interest. Other reflected fractions (such as specular) may only have interacted with the sample surface and thus does not contain chemical information related to the sample composition.

In order to correlate the sample absorbance to the concentration of a specific compound, the accurate amount of the compound under analysis must be known. For this reason NIR technologies are initially dependent on other chemical methods (also known as reference methods) to develop a calibration model and validate it properly. After some time, the calibration may need future updates because new sources of variability are most likely to appear when dealing with grains and seeds (variability due to fields, varieties, environmental changes etc.). More samples will have to be added in the model and more reference analyses will be needed. Therefore the selection of an appropriate primary chemical reference method and laboratory a crucial step when developing any NIRS application. Precision and accuracy of NIRS calibrations will be determined by the quality of the reference laboratory data. Combining reference data from different laboratories, even if the method is the same, is highly discouraged because errors from different laboratories differ.

There are many calibration algorithms, but most share the same principles as those that are widely known to perform well on quantitative analysis: multiple linear regression (MLR), principal component regression (PCR), and partial least squares (PLS) [6]. PLS and PCR often lead to very similar results, and MLR performs better when working with a short range of uncorrelated wavelengths or data points. PLS and PCR can be easily adapted for discrimination (i.e. PLS-DA) and are in fact derived from principal component analysis, a popular algorithm in pattern recognition and discrimination.

Once a calibration model is developed, it must be properly validated. True validation is done predicting independent samples, not related with samples included in the calibration set. Bagging and cross-validation are other validation alternatives when sample availability is a limitation. However, validation statistics from cross-validation may be overoptimistic, especially for models developed with few samples or not including all possible sources of variability. The use of suitable validation statistics is extremely important in order to report the calibration performance and determine its future use (screening, quality control etc.). The most widely utilized statistics for quantitative models are well summarized by William [7] and Fearn [8], and include among others those to quantify expected random errors (i.e. standard error of prediction (SEP) or cross validation (SECV)) and systematic error (bias). Other statistics such as the coefficient of correlation (R) – or its squared, the determination coefficient (R^2) and the ratio of the standard deviation of references over the SEP (RPD) give an idea of the overall calibration performance.

3. Impact of seed size and morphology

Single seeds are variable in shape and size. The variability of seed thickness translates in variability of the distance from the sample to the collecting sensor (focal length) and hence the sample distance (pathlength) that radiation travels in transmittance mode. Eq. (1) shows the relationship of the radiation reaching the sensor (F) with the focal length (f) and the irradiated sample diameter (D) [9]. According to that equation, large kernels reflect more radiation than small kernels as the focal length is larger. Once the reflectance is transformed to apparent absorbance, larger kernels will have lower optical density or absorbance offset compared to small kernels.

$$F = f/D \tag{1}$$

Because individual seeds can be small (i.e. 3 mm length for wheat kernels), the irradiated diameter (D) of commercial instruments working in reflectance mode is often larger than the seed diameter. In that case, the collected radiation includes scattering from kernel edges in detriment of relevant biochemical signal. For transmittance measurements, the irradiated diameter should not exceed the seed diameter because any radiation leakage through