Classification of grass samples containing or free of endophytes using near infrared and extended canonical variates analysis (ECVA)

R. Gislum,[a,*] P.H. Shetty[b] and J.R. Jørgensen[a]

[a]*Aarhus University, Faculty of Agricultural Sciences, Department of Genetics and Biotechnology, DK-4200 Slagelse, Denmark. E-mail: rene.gislum@agrsci.dk*

[b]*Department of Food Science and Technology, Pondicherry University, RV Nagar, Kalapet, Pondicherry 605014 India*

Introduction

Grass is used worldwide for forage either for grazing or as silage. The result of the increasing demand for meat and milk has put an immense pressure on the existing grass producing areas. Natural grass populations are often infected with toxic endophytes such as *Neotyphodium spp* resulting in considerable economic losses due to ill health and reduced livestock production. To avoid these drawbacks, the grasses are analyzed in the laboratory, but a faster screening method is warranted. Research in the past decade has shown that the endophytes are necessary for the persistence and maintenance production under biotic and abiotic stress, such as insect, herbivory and drought stress. Scientists have identified strains of endophytes which are non-toxic to livestock and still retain the grass host as persistent and resistant to biotic and abiotic stress. Some of these strains have also been commercialized successfully. With the increasing interest in the endophytes, the need for screening the grasses using non-invasive and rapid methods is also increasing.

Near infrared reflectance (NIR) spectroscopy has become a widely used method in the analysis of a range of agricultural products, as extensively reviewed by Williams and Norris, 1987.[1] and Osborne *et al*., 1993.[2] An example of how NIR can be used to quantify nitrogen concentration in grass samples and a description of how NIR can be further implemented for nitrogen application in agriculture can be found in Gislum *et al*.[3] Therefore, NIR is an obvious candidate in this study as it is already well known within agriculture. The use of the many variables from the NIR measurements to classify samples with or without endophytes requires special algorithms that are designed to handle the many multi-collinear data. The use of different classification methods which have been developed to handle multi-collinear data has been thoroughly described by Næs and Indahl.[4] Lately, a new algorithm named Extended Canonical Variates Analysis (ECVA) has been suggested and used in studies.[5,6]

An aim of the present experiment is to test the possibility to use near infrared reflectance spectroscopy (NIRS) as a screening method to determine if a grass sample contains endophytes.

Materials and methods

Reflectance spectra of 55 dried and ground grass samples were obtained using a QFA-Flex 400 FT-NIR instrument (Q-Interline, Roskilde, Denmark). The samples were packed in glass vials with a height of 6 cm and a diameter of 2.6 cm and measured using a rotating sample device. The sample was rotated at a speed of three rotations per minute, with a measuring sample window at the rotating sample device, which had a diameter of 6 mm, and the analysis surface was ~ 510 mm.[7] Spectra were collected at every 2 nm in the NIR region from 1100 to 2498 nm. One spectrum was obtained for each sample as an average of 64 sub-scans. The spectra were reported as log(1/*R*). Each sample was analysed for endophyte content and 28 samples contained endophytes. Principal component analysis (PCA) and extended canonical variates analysis (ECVA)[5] were performed using Matlab version 2008b. ECVA is an extension of the well known Canonical Variates Analysis method.[2] Models were made on raw spectra. The number of PLS components was seven.

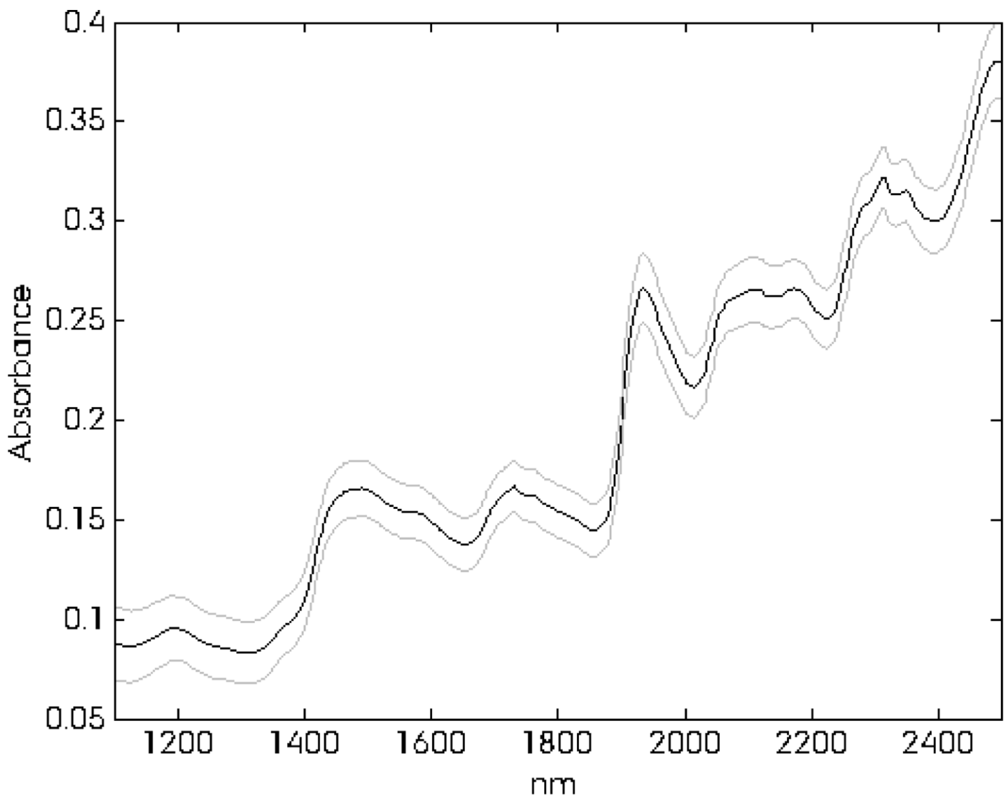

Figure 1. Average and +/– standard deviation NIR spectra of 55 grass samples.

Results and discussion

The average +/– standard deviation spectra of the 55 grass samples showed characteristics peaks for nitrogen, water and carbohydrates (Figure 1).

A PCA score plot showed no clear grouping of samples containing endophytes, or endophyte-free (data not shown). It was not surprising that the extended canonical variates were different between samples (Figure 2).

When Fisher's linear discriminant analysis was used for classification there were eight misclassifications, which equals 15% of the total number of samples, when seven partial least squares regression (PLSR) components were used (Figure 3).

The reason to choose seven PLSR components and not two to five PLSR components, which all had 13 misclassifications, can be discussed and more analysis should be made to define which information is achieved by including three extra PLSR components without lowering the number of misclassifications.

The use of NIR spectroscopy to classify grasses containing or free of endophytes is based on a co-factor as there is no theoretical proof that NIR should be able to detect the endophyte itself. The co-factor could be differences in water, nitrogen and/or carbohydrates in the grasses. Based on the extended canonical weight vector (data not shown) the area from 2240 nm to 2320 nm had large values and this range contains information from carbohydrates. The fact that the classification is based on a co-factor should not be a drawback in the future work, but it is of course important to describe the co-factor. The use of ECVA is thoroughly described by Nørgaard *et al.*[5] and further information can be found at www.models.kvl.dk.

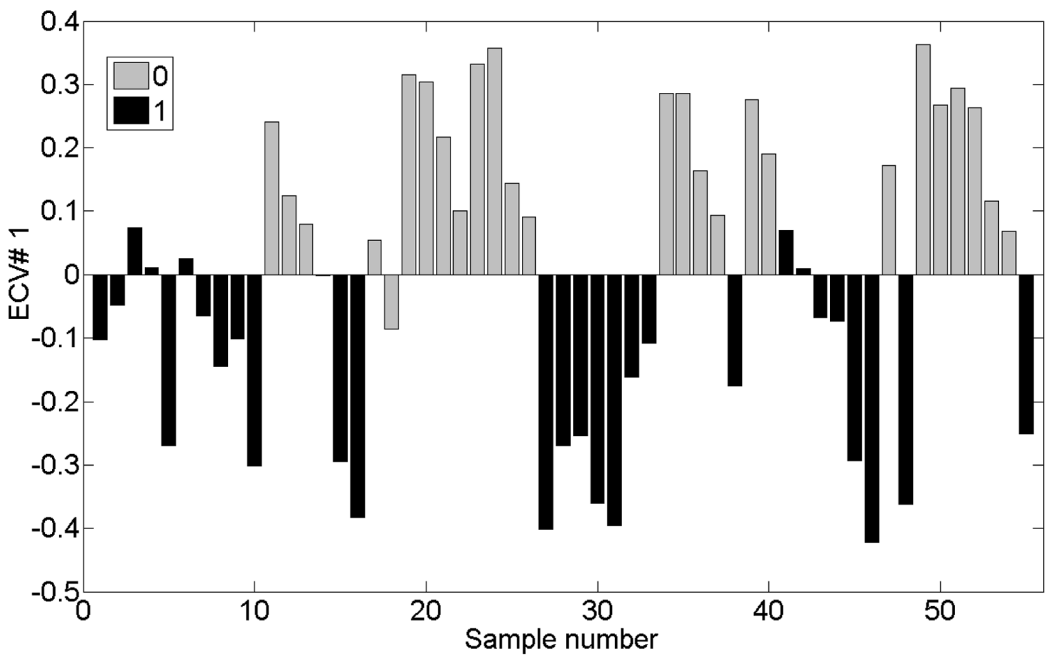

Figure 2. The extended canonical variates for group 0 (gray) and 1 (black) where group 0 has no endophytes and group 1 has endophytes.

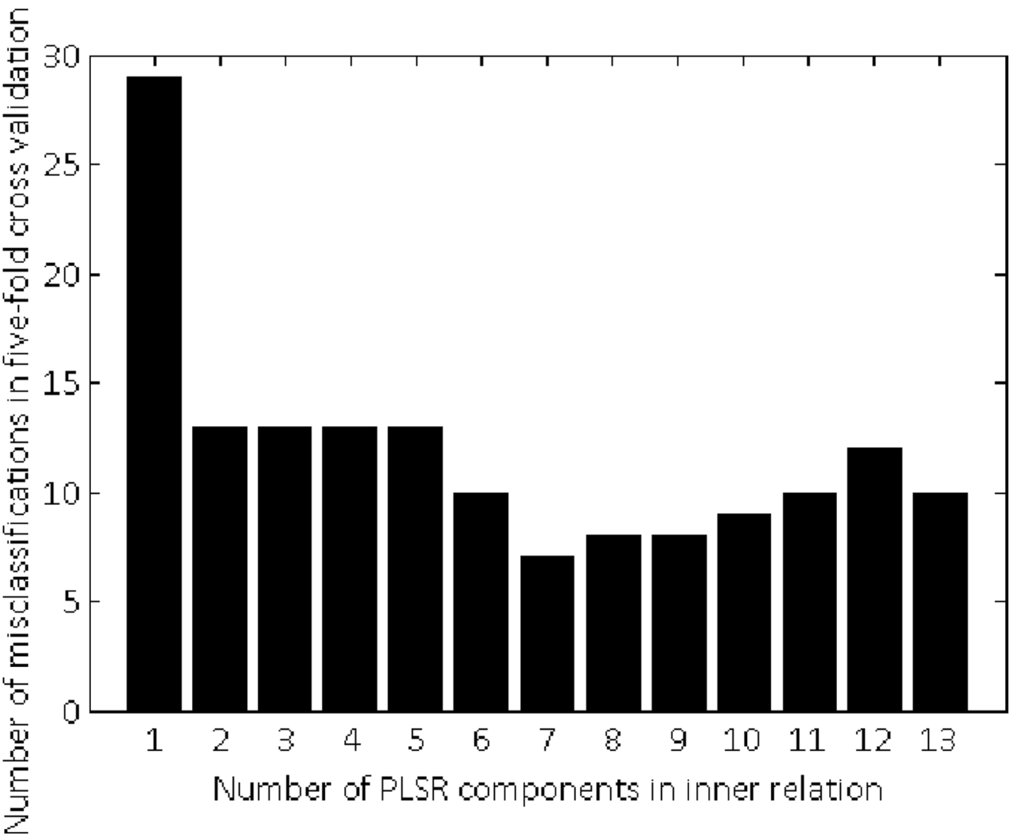

Figure 3. Number of misclassifications as a function of the number of PLSR components in the inner relation of ECVA on the NIR spectra.

References

1. P. Williams and K. Norris (Eds), *Near Infrared Technology in the Agricultural and Food Industries,* 2nd Edn. American Association of Cereal Chemists, Minnesota, USA (1987).
2. B. Osborne, T. Fearn and P. Hindle. *Practical NIR Spectroscopy with Applications in Food and Beverage Analysis,* 2nd Edn. Longman Scientific and Technical, Harlow, UK (1993).
3. R. Gislum, E. Micklander and J.P. Nielsen, *Field Crops Res.* **88,** 269 (2004).
4. T. Næs and U. Indahl, *J. Chemometr.* **12,** 205 (1998).
5. L. Nørgaard, R.Bro, F. Westad, S.B. Engelsen, *J. Chemometr.* **20,** 425 (2006).
6. L. Nørgaard, G. Sölétormos, N. Harrit, M. Albrechtsen, O. Olsen, D. Nielsen, K. Kampmann, R. Bro, *J. Chemometr.* **21,** 451 (2007).
7. A. Höskuldsson, Chemometr. Intell. Lab. Syst. **23,** 1 (1994).