Food Chemistry 240 (2018) 32–42

Contents lists available at ScienceDirect

Food Chemistry

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

Protein content prediction in single wheat kernels using hyperspectral imaging

Nicola Caporaso [a,b,*], Martin B. Whitworth [a], Ian D. Fisk [b,*]

[a] Campden BRI, Chipping Campden, Gloucestershire GL55 6LD, UK
[b] Division of Food Sciences, University of Nottingham, Sutton Bonington Campus, LE12 5RD, UK

ARTICLE INFO

Article history:
Received 16 March 2017
Received in revised form 29 June 2017
Accepted 10 July 2017
Available online 12 July 2017

Keywords:
Near-infrared spectroscopy
Wheat protein
Cereals
Rapid measurement
Chemical imaging
Single kernel assessment
Hyperspectral imaging

ABSTRACT

Hyperspectral imaging (HSI) combines Near-infrared (NIR) spectroscopy and digital imaging to give information about the chemical properties of objects and their spatial distribution. Protein content is one of the most important quality factors in wheat. It is known to vary widely depending on the cultivar, agronomic and climatic conditions. However, little information is known about single kernel protein variation within batches. The aim of the present work was to measure the distribution of protein content in whole wheat kernels on a single kernel basis, and to apply HSI to predict this distribution.

Wheat samples from 2013 and 2014 harvests were sourced from UK millers and wheat breeders, and individual kernels were analysed by HSI and by the Dumas combustion method for total protein content. HSI was applied in the spectral region 980–2500 nm in reflectance mode using the push-broom approach. Single kernel spectra were used to develop partial least squares (PLS) regression models for protein prediction of intact single grains.

The protein content ranged from 6.2 to 19.8% ("as-is" basis), with significantly higher values for hard wheats. The performance of the calibration model was evaluated using the coefficient of determination (R^2) and the root mean square error (RMSE) from 3250 samples used for calibration and 868 used for external validation. The calibration performance for single kernel protein content was R^2 of 0.82 and 0.79, and RMSE of 0.86 and 0.94% for the calibration and validation dataset, enabling quantification of the protein distribution between kernels and even visualisation within the same kernel. The performance of the single kernel measurement was poorer than that typically obtained for bulk samples, but is acceptable for some specific applications. The use of separate calibrations built by separating hard and soft wheat, or on kernels placed on similar orientation did not greatly improve the prediction ability. We simulated the use of the lower cost InGaAs detector (1000–1700 nm), and reported that the use of proposed HgCdTe detectors over a restricted spectral range gave a lower prediction error (RMSEC = 0.86% vs 1.06%, for HgCdTe and InGaAs, respectively), and increased R^2 value (R^2_c = 0.82 vs 0.73).

1. Introduction

Wheat is a staple commodity worldwide, used both for human consumption and for feed. Among wheat quality parameters, physical condition, moisture content, kernel hardness, Hagberg Falling Number (an indirect measurement of the effect of α-amylase activity in flour and ground wheat), and protein content are the most important. Protein content is important because it influences the technological performance in baked products, especially gluten formation for bread production, in combination with protein quality which is determined by varietal choice. Protein content has a significant impact on the final price, and many countries adopt it as a critical criterion to define wheat price.

Near-infrared (NIR) spectroscopy is a non-destructive and rapid method that can be used to investigate the chemical properties of complex food matrices and intact seeds or grains (Fox & Manley, 2014). The approach is based on the interaction of light radiation with the sample, in particular on molecular overtone and combination vibrations. NIR spectroscopy strongly relies on chemometrics for prediction of properties or classification of samples based on multivariate regression models, typically combined with spectral pre-treatment techniques. Common spectral pre-treatments aim to remove some interference due to the physical properties of the analyte, for example the particle size. These methods include

* Corresponding authors at: Division of Food Sciences, Sutton Bonington Campus, University of Nottingham, Sutton Bonington, Leicestershire, LE12 5RD, UK.
E-mail addresses: stxnc7@nottingham.ac.uk (N. Caporaso), ian.fisk@nottingham.ac.uk (I.D. Fisk).

http://dx.doi.org/10.1016/j.foodchem.2017.07.048
0308-8146/

multiplicative scatter correction (MSC), standard normal variate (SNV) and de-trending. A second category of spectral pre-treatments is based on the application of derivatives, including a smoothing step often using the Savitzky-Golay convolution method (Geladi, 2003; Rinnan, van den Berg, & Engelsen, 2009).

NIR spectroscopy measurement of protein content in wheat was initially based on batches of ground wheat or flours (Osborne, 1984). The feasibility of NIR prediction in batches of whole kernels without the need of grinding has been also demonstrated (Williams & Norris, 1987). Whole-grain applications became attractive to breeding programmes and industry due to the non-destructive scanning of samples and the speed of analysis, with the possibility of on-line or in-line data acquisition (Osborne, 1987; Williams & Sobering, 1993). Total protein prediction models based on NIR and visible spectroscopy of batches of whole wheat have high correlation coefficients (Cozzolino, Delucchi, Kholi, & Vázquez, 2006). However, the traditional NIR spectroscopy approach is based on bulk grains, and thus no indication of the protein variability among kernels is given (Bramble, Dowell, & Herrman, 2006). Some authors have tested single kernel NIR analysis, and applied it successfully to measure wheat protein content in transmittance mode (Delwiche, 1995). More recently, reflectance NIR spectroscopy was also applied, due to its better applicability at industrial level, and also tested for single kernels (Bramble et al., 2006).

The potential of single kernel NIR analysis strongly depends on its application and the quality parameter studied. For example it has limited potential to identify wheat varieties in breeding programmes, while more successful applications have been reported for physical grain quality determination, moisture, protein and kernel hardness measurement and loss of viability (Fox & Manley, 2014).

Protein variability among wheat kernels was reported for some USA wheat classes using NIR reflectance spectroscopy on a single kernel basis (Delwiche, 1998), and it was demonstrated that protein prediction of the batch can be improved by averaging a few hundred kernels from single kernel measurement (Delwiche & Hruschka, 2000). A single kernel characterisation of some European wheats was also performed by NIR transmittance spectroscopy (Nielsen, Pedersen, & Munck, 2003).

The possibility of high speed classification of single kernels for quality attributes is relevant in cereal breeding programmes, to improve the product quality. Whilst single kernel protein calibrations have been demonstrated, presentation of kernels individually results in practical difficulties for rapid analysis of bulk samples. Hyperspectral imaging (HSI) provides a potential approach to enable single kernel data to be acquired for larger numbers of kernels.

HSI combines NIR spectroscopy and digital imaging to give information about the spatial distribution of compounds. HSI creates three-dimensional "hypercube" datasets composed of two spatial dimensions and a single spectral dimension representing NIR spectra for each pixel of the image. As for bulk NIR spectroscopy, HSI heavily relies on chemometrics to extract useful chemical information from the hypercube (Gowen, O'Donnell, Cullen, Downey, & Frias, 2007). It has been applied successfully to measure the distribution of chemical composition in a wide range of food, including meat, fish, fruits, vegetables, and several applications to cereals (Gowen, O'Donnell, Cullen, Downey, & Frias, 2007; Wu & Sun, 2013). These include exploratory tests of HSI to measure or predict the milling quality of soft wheat (Delwiche, Souza, & Kim, 2013).

Although NIR calibrations show good performance for measurement of protein content in bulk wheat samples and are commonly applied at industrial level for laboratory and online measurement, limited work has been done on the application of HSI for wheat protein analysis. HSI offers potential advantages for assessment of uniformity in wheat and other granular food materials, united with the advantage of NIR spectrometry being contactless and rapid.

Therefore, the aim of our study was to develop an HSI calibration for total protein content in whole wheat kernels on a single grain basis and to assess the typical uniformity present in commercial wheat samples, and thus to apply the calibration to visualise the protein distribution within single kernels.

2. Material and methods

2.1. Wheat samples

Samples were obtained from a wide range of suppliers, mainly millers and breeders from the UK. Examples of Canadian, French, Italian, German and Eastern European wheat samples were also included. Samples came from the 2013 and 2014 harvests. They were selected to obtain a wide variation in terms of cultivars, environment and agronomic conditions, also including some genotypes from breeding trials and not yet registered or under registration. A total of approximately 190 wheat samples were used for the present experiment. From each sample, 10–12 kernels were randomly selected to be used for the analyses. Each kernel was presented for HSI measurement in both crease-up and crease-down orientations, resulting in a total of ~4200 kernel spectra. Each kernel was then analysed by the Dumas method to determine its protein content. The spectra and reference protein values were used for development and validation of the calibration.

2.2. Hyperspectral imaging

Data was acquired using a laboratory-scale hyperspectral imaging system described by Millar, Whitworth, Chau, and Gilchrist (2008) and Caporaso, Whitworth, and Fisk (2016). The instrument was supplied by Gilden Photonics Ltd. (Glasgow, U.K.) and includes a SWIR spectral camera (Specim Ltd., Oulu, Finland) containing a cooled 14 bit 320 × 256 pixel HgCdTe detector and N25E spectrograph providing 256 spectral bands over a wavelength range of ~980–2500 nm with a spectral resolution of about 6 nm. Samples were presented on a moveable sample stage and imaged using a push-broom approach. The camera was mounted above the stage at a distance of 220 mm and a 31 mm focal length lens was used, resulting in a swathe of 35 mm and a pixel size of 0.109 mm for 320 spatial pixels. Images were acquired at a rate of 100 frames s^{-1}, using a stage translation speed of 10.9 mm s^{-1}, providing the same spatial resolution parallel and perpendicular to the scan direction. A single 500 W incandescent illumination source was used for the first ~1000 kernels, and 2 lamps were used for the remaining samples. SpectralCube 3.0041 software (Specim) was used to control the camera and translation stage. The camera shutter was automatically closed for 1 s at the end of each scan and ~100 frames were recorded to establish the baseline signal of the detector (black reference). Separate scans of approximately 100 frames were also recorded for a white PTFE reference material with approximately 100% reflectance across the entire measured spectral range (white reference).

Wheat samples were presented with grains arranged in two rows on a black, NIR-absorbent plastic tray, and the hypercubes were obtained in diffuse reflectance mode. Images were first acquired for the dorsal side of the kernels, and the kernels were then manually rotated and a second image acquired for the ventral side. To minimise heating of the sample, the lamps were only turned on for the duration of the scan.