Universitat de Lleida

Document downloaded from:

<http://hdl.handle.net/10459.1/71219>

The final publication is available at:

<https://doi.org/10.1016/j.jfca.2020.103763>

Copyright

cc-by-nc-nd, (c) Elsevier, 2021

Està subjecte a una llicència de [Reconeixement-NoComercial-SenseObraDerivada 4.0 de Creative Commons](https://creativecommons.org/licenses/by-nc-nd/4.0/)

Journal Pre-proof

PREDICTION OF BIOACTIVE COMPOUNDS IN BARLEY BY NEAR-INFRARED REFLECTANCE SPECTROSCOPY (NIRS)

Elena Albanell (Conceptualization) (Methodology) (Writing - original draft) (Writing - review and editing) (Funding acquisition), Mariona Martínez (Methodology) (Formal analysis) (Writing - review and editing), Massimo De Marchi (Supervision) (Funding acquisition) (Writing - review and editing), Carmen L. Manuelian (Conceptualization) (Data curation) (Formal analysis) (Writing - review and editing)

PII: S0889-1575(20)31468-X

DOI: <https://doi.org/10.1016/j.jfca.2020.103763>

Reference: YJFCA 103763

To appear in: *Journal of Food Composition and Analysis*

Received Date: 21 July 2020

Revised Date: 15 November 2020

Accepted Date: 12 December 2020

Please cite this article as: Albanell E, Martínez M, De Marchi M, Manuelian CL, PREDICTION OF BIOACTIVE COMPOUNDS IN BARLEY BY NEAR-INFRARED REFLECTANCE SPECTROSCOPY (NIRS), *Journal of Food Composition and Analysis* (2020), doi: <https://doi.org/10.1016/j.jfca.2020.103763>

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2020 Published by Elsevier.

PREDICTION OF BIOACTIVE COMPOUNDS IN BARLEY BY NEAR-INFRARED REFLECTANCE

SPECTROSCOPY (NIRS)

Elena ALBANELL^{a,*}, Mariona MARTÍNEZ^b, Massimo DE MARCHI^c, Carmen L. MANUELIAN^{c,*}

^a Group of Ruminant Research (G2R), Department of Animal and Food Sciences, Universitat Autònoma de Barcelona (UAB), 08193, Bellaterra, Spain

^b Food Technology Department, Universitat de Lleida-Agrotecnio Center, Lleida, Alcalde Rovira Roure 191, 25198 Lleida, Spain

^c Department of Agronomy, Food, Natural resources, Animals and Environment (DAFNAE), University of Padova, Viale dell'Università 16, 35020 Legnaro (PD), Italy

e-mail address: elena.albanell@uab.cat ; marionamar@pvcf.udl.cat ; massimo.demarchi@unipd.it ; carmenloreto.manuelianfuste@unipd.it

Highlights

- Development NIRS prediction models by two instruments and different spectral ranges
- NIRS was applied to predict Bound Phenols and Anthocyanin in barley flours
- The anthocyanin prediction model did not improve if the visible range was included

ABSTRACT

Barley grains contain a variable amount of biologically active compounds such as non-starch polysaccharides and phenol compounds. These compounds are important in nutrition due to their significant health benefits and technological role in food. We developed predictive models for β -glucans (BG), arabinoxylans (AX), bound phenols (BP), free phenols (FP), and anthocyanins (AN) based on near-infrared spectroscopy (NIRS) using two different NIRS instruments with different

spectral range and spectral steps. Regressions of modified partial least squares (MPLS) and several combinations of scattering correction and derivative treatments were tested. The optimal calibration models generated high coefficients of determination for BG and BP, but not for AN content. The instrument with the highest resolution only gave better results for BG prediction models, and the addition of the visible range did not prove to be ostensibly advantageous to the determination of any of the active compounds of study, not even in the case of AN analysis.

Keywords: Anthocyanin; Arabinoxylan; Barley; β -glucan; Near-Infrared; Phenolic compounds

1. Introduction

Barley (*Hordeum vulgare* L.) is one of the most ancient crops and, thanks to domestication, today is one of the major crops in the world in terms of acreage and production, which significantly contributes to the world's food supply (Gupta et al., 2010). It is grown for many uses such as food source for human beings and animals as well as industrial food processing (e.g., malting, brewing, and baking) (Gordon et al., 2019).

Nowadays, there is an increasing trend to provide consumers safe foods based on cereal grains and/or components with high nutritive value to promote their health. Cereal grains contain a variable amount of biologically active compounds such as dietary fiber (e.g., arabinoxylans, β -glucans, cellulose, and lignin), sterols, tocopherols, phenolic compounds, vitamins, and microelements (Bartłomiej et al., 2012; Blakeney & Flinn, 2005). Compared to other grains, barley has a relatively high concentration of (1 $\rightarrow$ 3),(1 $\rightarrow$ 4)- β -glucans (β -glucan) (BG) and arabinoxylans (AX) which depend on genetic and environmental factors (Gupta et al., 2010). Both BG and AX are associated with cell walls of cereals. Type and composition of these non-starch polysaccharides and phenolic compounds is of increasing importance in the nutrition of both humans and animals due to their significant health benefits and technological role in food. Consumption of cereals containing BG and/or phenolic compounds contributes to a decreased risk of chronic health problems, such as those associated with cardiovascular diseases and diabetes, by reducing blood cholesterol and glucose levels, and it can also contribute in preventing some types of cancer (Gupta et al., 2010; Han et al., 2017; Li et al., 2003).

Conventional analyses to determine bioactive compounds of barley are based on chemical and enzymatic analyses which provide the basis to calculate the accuracy and precision of new methods/analytical tools. However, these conventional methods require highly trained technicians/workforce, are time-consuming and labor-intensive, use several chemical reagents, and destroy the sample during the analysis. As a consequence, the development of a simple, rapid,

and cost-effective method for measuring bioactive compounds in barley would be of great benefit for the food industry.

Near-infrared spectroscopy (NIRS) is a technique that collects the reflected light of a sampled material in the near-infrared region of the electromagnetic spectrum and can provide information on the chemical composition of agricultural products. As well as being a rapid analytical method and a relatively cost-effective alternative to conventional analyses, it does not destroy the sample, thus allowing further analysis. The use of NIRS for quality control of cereals is well established in the literature (Osborne, 2000). Some authors have shown the possibility to predict BG content in intact barley seeds using different near-infrared instruments, resolution, and wavelength range: De Sá & Palmer (2006) used the wavelength range between 900-1700 nm; Sohn et al. (2008) used three NIRS instruments with different resolutions, one Fourier transform near-infrared (FT-NIR) and two dispersive instruments (8 nm and 10 nm bandpass); Ringsted et al. (2017) used long wavelength NIRS (2260-2380 nm). In naked barley, Schmidt et al. (2009) quantified BG content and compared different types of near-infrared instruments: FT-NIR, NIT (near-infrared transmission), and Vis-NIR (visible and near-infrared; 400-2498 nm, scanning at 2-nm intervals). Moreover, a previous study suggested the feasibility of NIRS to predict AX in a set of grains and pulses (Blakeney & Flinn, 2005), but, to the best of our knowledge, there are no studies that measured AX in barley flour using NIRS. Although the NIRS technique has been successfully applied for the determination of total phenols in sorghum (Alfieri et al., 2019; Dykes et al., 2014), rice (Zhang et al., 2008), and barley grains (Han et al., 2017), studies describing the use of NIRS technique to measure anthocyanins in barley are lacking.

This study aimed to develop NIRS prediction models for rapid measurement of non-starch polysaccharides (BG and AX), phenols compounds (free and bound phenolic acids – FP and BP), and anthocyanins (AN) of barley flour as well as to compare the accuracy of two different NIRS instruments.

2. Material and methods

2.1. Barley samples

A total of 186 barley samples were selected from a previous genetic study to obtain genotypes with high BG and amylose contents and enhanced antioxidant activity. All barley varieties used in the present study were grown in Spain (Semillas Batlle, SA, Bell-lloc d'Urgell, Spain) during 2016, 2017, and 2018 and included hulled and hull-less genotypes. Barley grain samples were stored at room temperature in a desiccator and in the dark until use. Prior to the analysis, barley grain samples were ground in a Foss Cyclotec 1093™ (FOSS, Hillerød, Denmark) mill equipped with a 0.5 mm screen to produce whole grain flour. A sub-sample of each sample flour was used for wet chemical analysis and the remnant flour was used for NIRS analysis. The ground samples were stored at 4°C until analysis.

2.2. Chemical analysis

Content of BG was determined with the β -glucan (mixed-linkage) from Megazyme (Bray, Ireland). Briefly, 0.2 mL of aqueous ethanol (50% v/v) and 4 mL of sodium phosphate buffer (20 mmol/L, pH 6.5) were added to 80 mg of whole grain flour. After that, the samples were vortexed, incubated, and tempered at 50 °C. Then, 0.2 mL lichenase enzyme (10U) was added to the samples tube and incubated. After incubation, sodium acetate buffer was added and centrifuged at 1000 $\times$ g. For each sample, 0.1 mL of the supernatant was transferred to other tubes, and 0.1 mL of β -glucosidase (0.2U) was added. The reaction blanks were also prepared with 0.1 mL of each sample and 0.1 mL of sodium acetate buffer was added. The mixtures were incubated a 50 °C for 50 min and then, glucose oxidase/peroxidase reagent was added and the mixture was incubated at 50 °C for 20 min. After that, the absorbance was read at 510 nm using a spectrophotometer Jenway 6300 (Essex, England). The results were expressed as g/100 g of whole grain flour.

The arabinoxylans content was analyzed with the D-Xylose (Xylan & Arabinoxilan) kit from Megazyme (Bray, Ireland). Fifty mg of each barley flour sample was placed in 50 mL centrifuge tubes and 2.5 mL of HCl 1.3 M were added. Then, the samples were incubated, vortexed, cooled to room temperature, and 2.5 mL of NaOH 1.3 M was added to each. For each sample, 10 μ L of the supernatant was transferred to a well on the microplate. Absorbance readings at 340 nm were taken using a Multiscan GO spectrophotometer (Thermo Scientific, Vantaa, Finland). The results were expressed as g/100 g of whole grain flour.

Free phenolic compounds were extracted using 1 mL of the extraction solution (79.5% methanol, 19.5% distilled water, and 1% formic acid) in 50 mg of the sample. Samples were mixed for 10 min and centrifuged at $9000 \times g$ for 10 min. The extraction process was repeated two more times and the supernatants from the three extractions were pooled before the spectrophotometric analysis. The BP were extracted from the residue obtained after the extraction of the FP by alkaline hydrolysis with NaOH 2M. The samples were kept for 12 h at room temperature to complete the hydrolysis. Then, the samples were mixed for 10 min and centrifuged at $9000 \times g$ for 10 min; the supernatant was transferred to clean tubes and acidified with HCl 37% (w/w) to reach pH 2. After centrifuging, 1 mL of supernatant was mixed with a solution of 99.9% methanol and 0.1% formic acid, and centrifuged (10 min at $9000 \times g$). The extraction process was repeated one more time. The supernatants from the two extractions were pooled and directly analyzed by spectrophotometric method. Both FP and BP determinations were carried out according to the Folin-Ciocalteu method adapted to a microplate format (Bobo-García et al., 2015). The absorbance was read at 760 nm using a Multiscan GO spectrophotometer (Thermo Scientific, Vantaa, Finland). Results were expressed as mg of gallic acid equivalents/g of whole grain flour.

The Abdel-Aal & Hucl (1999) method was used for the analysis of total AN compounds. Briefly, a 1.6 mL of solution of 85% ethanol and 15% hydrochloric acid 1M was added to 100 mg of flour

sample. Then, samples were mixed for 15 min and centrifuged at $27200 \times g$ for 30 min.

Absorbance readings at 535 nm were taken using a Multiscan GO spectrophotometer (Thermo Scientific, Vantaa, Finland). The AN were quantified with a standard curve obtained for the cyanidin-3 glucoside. The results were expressed as μg of cyanidin-3 glucoside equivalents/g of whole grain flour.

2.3. NIRS analysis and calibration procedure

Two NIR spectrometers were used in this study. A NIRSystems 5000 scanning monochromator (FOSS, Hillerød, Denmark) with spectral range 1100-2500 nm using a PbS detector, located at the Universitat Autònoma de Barcelona (Bellaterra, Spain); and a Vis-NIR spectrometer DS2500 (FOSS, Hillerød, Denmark) equipped with a monochromator with spectral range 400-2500 nm and a dual Si and PbS detector, located at the food laboratory of the Department of Agronomy, Food, Natural resources, Animals and Environment of the University of Padova (Legnaro, Italy). In both spectrometers the spectral data were collected in a diffuse reflection mode. The first one used a closed ring cup cell (35 mm diameter) with quartz glass windows containing 2-3 g of the sample and absorbance was recorded as $\log(1/\text{reflectance})$ ($\log(1/R)$) every 2 nm from 1108 to 2492 nm which gave 692 data points for each sample. Each sample was scanned twice, by manually rotating the sample cup approximately 180° relative to the previous scan. In order to obtain better accuracy in calibration, we used the mean of the two scans of each sample (average spectral data). The DS2500 used an open ring cup cell (40 mm diameter intern; FOSS cup) containing 3-4 g of the sample and absorbance was recorded as $\log(1/R)$ every 0.5 nm from 408 to 2492 nm in which 4138 data points were given for each sample. Each spectrum was the average of 32 sub-spectra collected during the automatic rotation of the FOSS cup.

Chemometric analysis was performed using WinISI 4.10 software (Infrasoft International, Port Matilda, PA, USA). Prior to calibration, $\log(1/R)$ spectra were corrected for the effects of scatter

using the standard normal variate (SNV) and detrend (D) algorithms and by multiplicative scatter correction (MSC) to reduce the effects of the particle size. The prediction models were performed by the modified partial least squares regression (MPLS) and several combinations of scattering correction (NONE, D, SNV, SNV + D, MSC) and derivative (0,0,1,1; 1,4,4,1; 1,5,5,1; 1,8,8,1; 1,10,10,1; 2, 4,4,1; 2,5,5,1; 2,8,8,1; and 2,10,10,1; where the first digit is the number of the derivative, the second is the gap over which the derivative is calculated, the third is the number of data points in the first smoothing, and the fourth is the number of data points in the second smoothing) were tested.

One hundred and fifty samples were used for calibration (approximately 80% of the total samples) and 36 random samples (around 20% of the total samples) were separated previously and used as an external source of samples for validation. Different MPLS regressions were developed using different spectral range and spectral step: 1) NIRSystems 5000, range: 1108-2492 nm, 2 nm steps, 692 data points; 2) DS2500, range: 408-2492 nm, 0.5 nm steps, 4138 data points; 3) DS2500, range: 408-1092 nm, 0.5 nm steps, 1369 data points; and 4) DS2500, range: 1108-2492 nm, 0.5 nm steps, 2769 data points. In all scenarios, the same samples were used as a validation set. In the calibration set, chemical (t) or spectral (H) outliers (about 8% of the samples) were removed. The t outliers are samples that have a relationship between their reference values and spectra that is different from the relationship of the other samples in the set and with large residuals (t values > 2.2). The optimum model was selected by the following statistics: minimum standard error of calibration (SEC), minimum standard error of prediction (SEP), greatest coefficient of determination for calibration (R^2_{CAL}), greatest coefficient of determination for validation (R^2_{VAL}), the ratio of performance to deviation near 3.0 (RPD, defined as the ratio of standard deviation for the validation samples to the value of SEP), and the range error ratio near 10 (RER, defined as the ratio of the range in the reference data from the validation set to the SEP) (Williams, 2014; Williams & Sobering, 1996).

Moreover, in the validation set, reference values of each model were linearly regressed on the respective predicted value to obtain the linear regression coefficient (slope), to test if the slope differed from one, and to obtain the residuals using PROC REG of SAS ver. 9.4 (SAS Institute Inc., Cary, NC, USA). Then, the normality of the residuals was assessed and the average of the residuals (bias) was tested using PROC UNIVARIATE of SAS ver. 9.4 (SAS Institute Inc., Cary, NC, USA) to see if differed from zero.

3. Results & discussion

Table 1 shows the compositional data of the ground barley samples employed for the development of the calibration and validation equations in this study. Barley samples were selected to include the greatest possible genetic variability and were obtained in three different years to provide a broad range of bioactive components, which are considered basic principles to obtain good calibration models (De Marchi et al., 2018; Næs et al., 2002). Therefore, the distribution of samples obtained across the entire range (Table 1) is a good starting point for a robust calibration. All parameters were well represented in both calibration and validation matrices covering similar ranges and a relatively broad range as is recommended (Næs et al., 2002).

Average absorbance spectra for ground barley samples, captured by the two instruments (NIRSystems 5000 and DS2500), with the most relevant absorption bands indicated are shown in Figure 1. In the visible region of the spectrum (400-700 nm), absorption was probably affected by the pericarp color of the barley samples (Dykes et al., 2014) as it suggests the observed peak at 667 nm (red light). In the NIRS region of the spectra, a first and defined peak was observed at 1200 nm and a wide peak at 1460 nm. Further peaks were found close to 1720 and 1770 nm. Moreover, there was a narrow peak at 1930 nm, a wide peak at 2100 and three more peaks at 2280, 2320 and 2490 nm. The average NIRS raw spectrum obtained followed a similar shape to those

previously reported in whole kernel barley (Sohn et al., 2008) and ground barley samples (Dykes et al., 2014; Shi et al., 2019). According to previous reports, the characteristic bands of BG constituents can be detected in the regions from 1194 nm to 1290 nm and from 2260 nm to 2380 nm (Seefeldt et al., 2009), and phenolics and flavonoids should be detected in the regions 1415-1512 nm, 1650-1750 nm, and 1955-2035 nm (Han et al., 2017; Verardo et al., 2015; Zhang et al., 2008).

In the present study, calibrations by MPLS regression were performed for barley flour using the average spectral per sample and two NIR spectrometers with different spectral range and wavelength accuracy (Table 2). Different pre-treatments of spectral data were tested for their ability to remove or reduce disturbing effects not related to the chemical absorption of light. The model for each component shown in Table 2 is the best result from 45 models developed resulting from the evaluation of different pre-treatments (five scatter correction techniques and nine math treatments). The optimal spectral pre-treatments were the second derivative (2,4,4,1 and 2,5,5,1) of log (1/R) –except for BG content which was the first derivative (1,4,4,1)– combined with SNV-D or MSC. These mathematical treatments are commonly used in food and cereal analysis by NIRS (Stubbs et al., 2010; Woodcock et al., 2008). The linear regression of measured versus predicted values for the best calibration equation for each bioactive compound (BG, AX, BP, FP, and AN) are represented in Figure 2. After applying those calibration models to the validation dataset, slope of measure vs predict values did not differ from the unity, except the one selected for AN using DS2500, range 408-2492 nm at 0.5 intervals; residuals were normally distributed and bias did not differ from zero. The best predictive models built among all the components analyzed were obtained for BG and BP content, showing an R^2_{CAL} of 0.97 and 0.90, respectively. Predictions are considered excellent when coefficient of determination is greater than 0.91, good when ranges from 0.82 to 0.90, approximate when is between 0.66 and 0.81, and poor when is less than 0.66 (Karoui et al., 2006). The worse predictive models were observed for AX and FP with $R^2_{\text{CAL}} \leq 0.70$.

As reported by Manley (2014), these apparently poor results could be related to the accuracy of the reference method and/or the narrow range of values encountered (Table 1 and Table 2). The error of laboratory increases when more steps are involved and lower is the concentration of the trait to be predicted, which directly influences the accuracy of the prediction models (De Marchi et al., 2018).

We obtained excellent calibration models for BG content with both devices (instruments) and wavelength range (R^2_{CAL} , 0.95 to 0.97; SEC, 0.33 to 0.40; SD/SECV, 3.71 to 4.36; Table 2). Only a slightly lower accuracy of the models for BG content was observed with the wavelength range from 408 to 1092 nm. That specific range is the one that considered the shorter wavelength region of the instrument Vis-NIR. The calibration model was still considered good (R^2_{CAL} , 0.83), but it presented greater SEC (0.72) and lower SD/SECV (2.13). These results were in agreement with those reported by Schmidt et al. (2009) using a FT-NIR (R^2_{CAL} = 0.96-0.98) and by Seefeldt et al. (2009) in barley flour samples with a dispersive NIRS in the region 1194-1240 nm (R^2_{CAL} = 0.94). Ringsted et al. (2017) showed the effect of BG range on the model performance to predict BG content in barley. Models with narrower BG range showed lower R^2_{CAL} values than models with wider BG range –like in our case–, which achieved greater R^2_{CAL} values (Ringsted et al., 2017). Some studies demonstrated the potential to discriminate barley flour or barley slice samples using long-wavelength (from 2200 to 2500 nm) based on their BG content (Ringsted et al., 2017; Seefeldt et al., 2009) which supports the slightly better results we obtained using the NIR region (1108–2492 nm) than the vis–NIR region (408–1092 or 408-2498 nm).

Validation statistics (Table 3) for the best-fitting equations previously selected for BG (Table 2) showed a great accuracy of the models using both devices and different wavelength range (RPD, 3.2 to 3.9; RER, 11.8 to 14.5; SEP, 0.394 to 0.486). However, validation statistics considering only from 408 to 1092 nm were poorer (RPD, 2.1; RER 7.6; SEP, 0.754). Based on Williams (2014), for complex feed matrix, a prediction model with an $RPD \leq 1.9$ is considered to be not suitable; values

between 2.0 and 2.4 are considered poor and only adequate for rough screening purposes, values between 2.5 and 2.9 could be applied for screening, and values ≥ 3.0 (or $RER > 10$) indicate good prediction and can be used for quantitative analysis. Those results confirm the high precision of the BG prediction model developed using both devices and a wavelength above 1092 nm since they provided higher values than the minimum recommended for prediction uses. In contrast, Sohn et al. (2007) reported a much lower R^2_{CAL} (0.67) and RPD (1.6) for BG using a dispersive NIR spectroscopy (NIR region, 1100-2498 nm) than we did, but the study was focused on barley for bioethanol production, thus, barley samples rich in BG content were not included. A better prediction model adequate for quality control (RPD = 3.5) was reported by Blakeney & Flinn (2005) using a sample set of grains and pulses, while a RPD of 4 to 5.6 was reported in naked barley by Schmidt et al. (2009) using FT-NIR. Although, a lower RPD (2.1-2.3) was achieved using dispersive NIR (1100-2500 nm) (Schmidt et al., 2009).

Calibrations developed for AX content were unsatisfactory ($R^2_{CAL} \leq 0.67$; Table 2) considering the studied instruments and wavelength ranges. In the same way, validation statistics revealed that prediction models were not suitable for prediction purposes ($RPD \leq 1.5$; $RER \leq 6.0$; $SEP \leq 0.9$; Table 3). A better accuracy for AX content ($R^2_{CV} = 0.84$; RPD = 2.51) was reported previously on a set of samples of cereals grains and pulses including a wider diversity (i.e., with diverse genetic and environmental backgrounds) (Blakeney & Flinn, 2005). Although NIR spectroscopy was capable of detecting AX content in plant materials (Blakeney & Flinn, 2005; Kačuráková et al., 1994), to the best of our knowledge, no study has obtained satisfactory prediction models for AX in barley flour.

In the current study, the best calibration equation for phenolic acids was obtained for BP using the 1108-2492 nm spectral range and 2 nm steps ($R^2_{CAL} = 0.90$; SD/SECV = 2.5; Table 2). The use of the instrument with more resolution (0.5 nm steps) and that included the visible region (408-2492 nm spectral range) slightly decreased the R^2_{CAL} and the SD/SECV (0.86 and 2.3, respectively; Table

2). According to previous reports, the characteristic bands of phenolics and flavonoids should be detected in wavelength regions from 1415 nm to 1512 nm, 1650 to 1750 nm, and from 1955 to 2035 nm (Cozzolino et al., 2004; Verardo et al., 2015; Zhang et al., 2008), thus, the inclusion of the visible wavelength seems to impair the prediction model. The R^2_{VAL} and the RPD index for BP content using the 1108-2492 nm spectral range and 2 nm were 0.9 and 3.2, respectively, which indicated the effectiveness of the method. However, the accuracy of the prediction models in external validation was lower for FP ($R^2_{\text{VAL}} = 0.6$; RPD = 1.7) and AN ($R^2_{\text{VAL}} = 0.7$; RPD = 2.1) content, which suggested a greater variation between the reference values and NIR predictions for FP and for AN than for BP content. Sometimes, low coefficient of determinations can be observed in samples with low or non-detectable concentrations of the component and/or narrow range, thus making more difficult to quantify and to predict them. Anthocyanins belong to the flavonoids group and provide the purple, blue, and red shades in many plants, including barley seeds. Despite the relationship between anthocyanins and color, the predictive models obtained did not improve ostensibly when the visible range was included.

Phenolic compounds in cereal grains are either in free or bound forms. Generally, free phenolic compounds are proanthocyanidins or flavonoids, whereas bound phenolic compounds are ester-linked to cell wall polymers and mainly consist of ferulic acid and its oxidatively coupled dimers (Bonoli et al., 2004). In ground rice samples, Zhang et al. (2008) determined total phenolic and flavonoid content achieving R^2_{VAL} values of 0.86 and 0.38 and RPD values of 2.60 and 1.27, respectively, using a dispersive NIRS. In whole sorghum grain, Dykes et al. (2014) obtained values of R^2_{VAL} of 0.86 and 0.68 for total phenols and for 3-deoxyanthocyanidins, respectively, using a vis/NIR spectrometer. In grain ground barley samples, Verardo et al. (2015) reported R^2_{VAL} values of 0.96 to determine proanthocyanidin using the spectral range (1100-2500 nm) and FT-NIR spectroscopy, and more recently, Han et al. (2017) obtained R^2_{VAL} and RPD values for total phenolics content of 0.95 and 3.2, respectively, and for free *p*-coumaric acid content of 0.96 and

3.6, respectively, using a dispersive NIR instrument; values slightly higher to the ones found in the present study. To the best of our knowledge, there are no references available on barley grain to compare the accuracy of our prediction equations for BP, FP, and AN.

4. Conclusion

The study investigated the feasibility of using two different NIRS instruments, resolution, and wavelength range, combined with chemometrics for the determination of non-starch polysaccharides, phenols compounds, and anthocyanins in a wide range of barley flour. Different spectral pretreatment techniques have been tried to construct acceptable models. It was concluded that the NIRS models developed were sufficiently accurate to predict BG and BP contents in barley grain, poor to predict AN, and not reliable for the prediction of FP and AX content. Our results highlighted that both spectral range and spectral steps influence the accuracy of the models, obtaining the best results retaining only the 1108 – 2492 nm region and using the 0.5 nm steps for BG and 2 nm steps for the rest of components. Despite the relationship between AN and color, the predictive models obtained apparently did not improve when the visible range was included. The AX prediction model was the component with worst results, probably due to the complexity of this fraction and/or the narrow range of our values.

Author statement

All co-authors have approved the manuscript and there is no financial interest to report. We confirm that this manuscript has not been published elsewhere and is not under consideration by another journal.

CRedit authorship contribution statement

Elena Albanell: Conceptualization, Methodology, Writing - original draft, Writing - review &

editing, Funding acquisition. **Mariona Martínez:** Methodology, Formal analysis, Writing - review &

editing. **Massimo De Marchi**: Supervision, Funding acquisition, review & editing. **Carmen L.**

Manuelian: Conceptualization, Data curation, Formal analysis, Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research was funded by MINECO/FEDER (AGL2015–69435-C3–3-R) project (Spain). Vis-NIR spectrometer DS2500 was funded by Progetto di Eccellenza "Centro per l'Agricoltura, la Sostenibilità e gli Alimenti" (CASA), CUP C26C18000190001, MIUR (Italy).

References

- Abdel-Aal, E. S. M., & Hucl, P. (1999). A rapid method for quantifying total anthocyanins in blue aleurone and purple pericarp wheats. *Cereal Chemistry*, 76(3), 350–354.
<https://doi.org/10.1094/CCHEM.1999.76.3.350>
- Alfieri, M., Cabassi, G., Habyarimana, E., Quaranta, F., Balconi, C., & Redaelli, R. (2019). Discrimination and prediction of polyphenolic compounds and total antioxidant capacity in sorghum grains. *Journal of Near Infrared Spectroscopy*, 27(1), 46–53. <https://doi.org/10.1177/0967033518825351>
- Bartłomiej, S., Justyna, R. K., & Ewa, N. (2012). Bioactive compounds in cereal grains – occurrence, structure, technological significance and nutritional benefits – a review. *Food Science and Technology International*, 18(6), 559–568. <https://doi.org/10.1177/1082013211433079>
- Blakeney, A. B., & Flinn, P. C. (2005). Determination of non-starch polysaccharides in cereal grains with near-infrared reflectance spectroscopy. *Molecular Nutrition and Food Research*, 49(6), 546–550.
<https://doi.org/10.1002/mnfr.200500038>
- Bobo-García, G., Davidov-Pardo, G., Arroqui, C., Vírveda, P., Marín-Arroyo, M. R., & Navarro, M. (2015). Intra-laboratory validation of microplate methods for total phenolic content and antioxidant activity on polyphenolic extracts, and comparison with conventional spectrophotometric methods. *Journal of the Science of Food and Agriculture*, 95(1), 204–209. <https://doi.org/10.1002/jsfa.6706>
- Bonoli, M., Verardo, V., Marconi, E., & Caboni, M. F. (2004). Antioxidant phenols in barley (*Hordeum vulgare* L.) flour: Comparative spectrophotometric study among extraction methods of free and bound phenolic compounds. *Journal of Agricultural and Food Chemistry*, 52(16), 5195–5200.
<https://doi.org/10.1021/jf040075c>
- Cozzolino, D., Kwiatkowski, M. J., Parker, M., Cynkar, W. U., Damberg, R. G., Gishen, M., & Herderich, M. J. (2004). Prediction of phenolic compounds in red wine fermentations by visible and near infrared spectroscopy. *Analytica Chimica Acta*, 513(1), 73–80. <https://doi.org/10.1016/j.aca.2003.08.066>
- De Marchi, M., Penasa, M., Zidi, A., & Manuelian, C. L. (2018). Invited review: Use of infrared technologies for the assessment of dairy products—Applications and perspectives. *Journal of Dairy Science*,

101(12), 10589–10604. <https://doi.org/10.3168/jds.2018-15202>

De Sá, R. M., & Palmer, G. H. (2006). Analysis of β -glucan in single grains of barley and malt using NIR-spectroscopy. *Journal of the Institute of Brewing*, 112(1), 9–16. <https://doi.org/10.1002/j.2050-0416.2006.tb00701.x>

Dykes, L., Hoffmann, L., Portillo-Rodriguez, O., Rooney, W. L., & Rooney, L. W. (2014). Prediction of total phenols, condensed tannins, and 3-deoxyanthocyanidins in sorghum grain using near-infrared (NIR) spectroscopy. *Journal of Cereal Science*, 60(1), 138–142. <https://doi.org/10.1016/j.jcs.2014.02.002>

Gordon, R., Chapman, J., Power, A., Chandra, S., Roberts, J., & Cozzolino, D. (2019). Mid-infrared spectroscopy coupled with chemometrics to identify spectral variability in Australian barley samples from different production regions. *Journal of Cereal Science*, 85(May 2018), 41–47. <https://doi.org/10.1016/j.jcs.2018.11.004>

Gupta, M., Abu-Ghannam, N., & Gallagher, E. (2010). Barley for brewing: Characteristic changes during malting, brewing and applications of its by-products. *Comprehensive Reviews in Food Science and Food Safety*, 9(3), 318–328. <https://doi.org/10.1111/j.1541-4337.2010.00112.x>

Han, Z., Cai, S., Zhang, X., Qian, Q., Huang, Y., Dai, F., & Zhang, G. (2017). Development of predictive models for total phenolics and free p-coumaric acid contents in barley grain by near-infrared spectroscopy. *Food Chemistry*, 227, 342–348. <https://doi.org/10.1016/j.foodchem.2017.01.063>

Kačuráková, M., Ebringerová, A., Hirsch, J., & Hromádková, Z. (1994). Infrared study of arabinoxylans. *Journal of the Science of Food and Agriculture*, 66(3), 423–427. <https://doi.org/10.1002/jsfa.2740660323>

Karoui, R., Mouazen, A. M., Dufour, É., Pillonel, L., Picque, D., Bosset, J.-O., & De Baerdemaeker, J. (2006). Mid-infrared spectrometry: A tool for the determination of chemical parameters in Emmental cheeses produced during winter. *Lait*, 86, 83–97. <https://doi.org/10.1051/lait:2005040>

Li, J., Kaneko, T., Qin, L. Q., Wang, J., & Wang, Y. (2003). Effects of barley intake on glucose tolerance, lipid metabolism, and bowel function in women. *Nutrition*, 19(11–12), 926–929. [https://doi.org/10.1016/S0899-9007\(03\)00182-5](https://doi.org/10.1016/S0899-9007(03)00182-5)

Næs, T., Isaksson, T., Fearn, T., & Davies, T. (2002). *A User-Friendly Guide to Multivariate Calibration and*

Classification (NIR Publications (ed.); Chichester).

Osborne, B. G. (2000). Near-Infrared Spectroscopy in Food Analysis. *Encyclopedia of Analytical Chemistry*, 1–14. <https://doi.org/10.1002/9780470027318.a1018>

Ringsted, T., Ramsay, J., Jespersen, B. M., Keiding, S. R., & Engelsen, S. B. (2017). Long wavelength near-infrared transmission spectroscopy of barley seeds using a supercontinuum laser: Prediction of mixed-linkage beta-glucan content. *Analytica Chimica Acta*, 986, 101–108. <https://doi.org/10.1016/j.aca.2017.07.008>

Schmidt, J., Gergely, S., Schönlechner, R., Grausgruber, H., Tömösközi, S., Salgó, A., & Berghofer, E. (2009). Comparison of different types of NIR instruments in ability to measure β -glucan content in naked barley. *Cereal Chemistry*, 86(4), 398–404. <https://doi.org/10.1094/CCHEM-86-4-0398>

Seefeldt, H. F., Blennow, A., Jespersen, B. M., Wollenweber, B., & Engelsen, S. B. (2009). Accumulation of mixed linkage (1 $\rightarrow$ 3) (1 $\rightarrow$ 4)- β -d-glucan during grain filling in barley: A vibrational spectroscopy study. *Journal of Cereal Science*, 49(1), 24–31. <https://doi.org/10.1016/j.jcs.2008.06.012>

Shi, H., Schwab, W., Liu, N., & Yu, P. (2019). Major ergot alkaloids in naturally contaminated cool-season barley grain grown under a cold climate condition in western Canada, explored with near-infrared (NIR) and fourier transform mid-infrared (ATR-FT/MIR) spectroscopy. *Food Control*, 102(January), 221–230. <https://doi.org/10.1016/j.foodcont.2019.03.025>

Sohn, M., Himmelsbach, D. S., Barton, F. E., Griffey, C. A., Brooks, W., & Hicks, K. B. (2007). Near-infrared analysis of ground barley for use as a feedstock for fuel ethanol production. *Applied Spectroscopy*, 61(11), 1178–1183. <https://doi.org/10.1366/000370207782597148>

Sohn, M., Himmelsbach, D. S., Barton, F. E., Griffey, C. A., Brooks, W., & Hicks, K. B. (2008). Near-infrared analysis of whole kernel barley: Comparison of three spectrometers. *Applied Spectroscopy*, 62(4), 427–432. <https://doi.org/10.1366/000370208784046768>

Stubbs, T. L., Kennedy, A. C., & Fortuna, A. M. (2010). Using NIRS to predict fiber and nutrient content of Dryland cereal cultivars. *Journal of Agricultural and Food Chemistry*, 58(1), 398–403. <https://doi.org/10.1021/jf9025844>

Verardo, V., Cevoli, C., Pasini, F., Gómez-Caravaca, A. M., Marconi, E., Fabbri, A., & Caboni, M. F. (2015).

Analysis of oligomer proanthocyanidins in different barley genotypes using high-performance liquid chromatography-fluorescence detection-mass spectrometry and near-infrared methodologies. *Journal of Agricultural and Food Chemistry*, 63(16), 4130–4137. <https://doi.org/10.1021/acs.jafc.5b01425>

Williams, P. (2014). Tutorial: The RPD statistic: a tutorial note. *NIR News*, 25(1), 22.

<https://doi.org/10.1255/nirn.1419>

Williams, P. C., & Sobering, D. C. (1996). How do we do it: a brief summary of the methods we use in developing near infrared calibrations. *Near Infrared Spectroscopy: The Future Waves*, 185–188.

Woodcock, T., Downey, G., & O'Donnell, C. P. (2008). Better quality food and beverages: The role of near infrared spectroscopy. *Journal of Near Infrared Spectroscopy*, 16(1), 1–29.

<https://doi.org/10.1255/jnirs.758>

Zhang, C., Shen, Y., Chen, J., Xiao, P., & Bao, J. (2008). Nondestructive prediction of total phenolics, flavonoid contents, and antioxidant capacity of rice grain using near-infrared spectroscopy. *Journal of Agricultural and Food Chemistry*, 56(18), 8268–8272. <https://doi.org/10.1021/jf801830z>

Figure 1

Spectra of ground barley. In black, average vis/NIR spectrum using the DS2500; in grey, average NIR spectrum using the NIRSystems 5000.

Figure 2

Linear regression plot of measured versus predicted values of the calibration dataset for: (A) β -glucan (BG), g/100 g; (B) Arabinoxylan (AX), g/100 g; (C) Bound phenol (BP), mg of gallic acid equivalents/g; (D) Free phenol (FP), mg of gallic acid equivalents/g; and (E) Anthocyanin (AN), μ g of cyanidin-3 glucoside equivalents/g. R^2_{CAL} : coefficient of determination in calibration.

Journal Pre-proof

Table 1

Flour barley population statistics of calibration and validation data set used.

Parameter	Calibration set				Validation set			
(Units)	N	Range	Mean	SD	n	Range	Mean	SD
β -glucan (g/100 g)	142	1.18-9.95	5.90	1.81	36	3.46-9.17	6.06	1.55
Arabinoxylan (g/100 g)	145	1.76-7.93	4.71	1.32	35	2.10-7.02	4.74	1.20
Bound Phenol (mg/g) ^a	150	1.70-6.76	3.65	1.05	36	1.91-6.37	3.76	1.18
Free Phenol (mg/g) ^a	150	1.38-4.15	2.41	0.58	36	1.38-3.51	2.35	0.52
Anthocyanin (μ g/g) ^b	139	2.76-66.47	19.45	14.56	32	3.43-54.03	17.68	14.07

N = number of samples for calibration; n = number of samples for external validation; Range = interval between the maximum and minimum value of data set; SD = standard deviation.

^aPhenolic content was expressed as mg of gallic acid equivalents/g.

^bAnthocyanins compounds were expressed as μ g of cyanidin-3 glucoside equivalents/g.

Table 2

Calibration and cross-validation fitting statistics for flour barley bioactive compounds prediction models using NIRSystems 5000 and DS2500 instruments.

Trait	^a Math treatment	^b Scatter correction	R ² _{CAL}	SEC	R ² _{CV}	SECV	SD/SECV
NIRSystems 5000 (1108-2492 nm at 2 nm intervals; 692 data points)							
β -glucan	1,4,4,1	SNV+D	0.966	0.346	0.941	0.435	4.16
Arabinoxylan	2,4,4,1	MSC	0.665	0.744	0.581	0.816	1.62
Bound Phenol	2,4,4,1	SNV+D	0.898	0.340	0.830	0.425	2.47
Free Phenol	2,4,4,1	SNV+D	0.672	0.319	0.515	0.377	1.54
Anthocyanin	2,5,5,1	MSC	0.839	5.807	0.614	8.578	1.70
DS2500 (408-2492 nm at 0.5 nm intervals; 4,138 data points)							
β -glucan	1,4,4,1	SNV+D	0.949	0.402	0.921	0.488	3.71
Arabinoxylan	2,4,4,1	SNV+D	0.674	0.721	0.590	0.794	1.66
Bound Phenol	2,4,4,1	SNV+D	0.855	0.386	0.785	0.458	2.29
Free Phenol	2,4,4,1	SNV+D	0.704	0.301	0.615	0.335	1.73
Anthocyanin	2,5,5,1	MSC	0.660	6.920	0.564	8.174	1.78
DS2500 (408-1092 nm at 0.5 nm intervals; 1,369 data points)							
β -glucan	1,4,4,1	SNV+D	0.825	0.724	0.751	0.850	2.13
Arabinoxylan	2,4,4,1	SNV+D	0.673	0.712	0.552	0.841	1.57
Bound Phenol	2,4,4,1	MSC	0.844	0.402	0.785	0.460	2.28
Free Phenol	2,4,4,1	MSC	0.633	0.300	0.401	0.375	1.55
Anthocyanin	2,5,5,1	MSC	0.856	5.309	0.616	8.292	1.76
DS2500 (1108-2492 nm at 0.5 nm intervals; 2,769 data points)							
β -glucan	1,4,4,1	SNV+D	0.967	0.332	0.944	0.415	4.36
Arabinoxylan	2,4,4,1	SNV+D	0.563	0.811	0.502	0.856	1.54
Bound Phenol	2,4,4,1	SNV+D	0.807	0.438	0.765	0.484	2.17
Free Phenol	2,4,4,1	SNV+D	0.682	0.304	0.589	0.347	1.67
Anthocyanin	2,5,5,1	MSC	0.689	7.231	0.491	9.002	1.62

^aMath treatment: derivative order, subtraction gap, first smoothing, second smoothing. ^bSNV standard normal variate, D detrend, MSC multiple scatter correction. R²_{CAL} coefficient of determination for calibration; SEC standard error of calibration; R²_{CV} coefficient of determination for cross validation; SECV standard error of cross validation.

Table 3

Validation fitting statistics bioactive for flour barley bioactive compounds prediction models using NIRSystems 5000 and DS2500 instruments.

Trait	R ² _{VAL}	SEP	Bias	Slope	RPD	RER
NIRSystems 5000 (1108-2492 nm) 2 nm						
β -glucan	0.912	0.486	-0.049	0.918	3.19	11.75
Arabinoxylan	0.568	0.834	-0.246	0.805	1.44	5.90
Bound Phenol	0.893	0.371	-0.053	0.983	3.18	12.02
Free Phenol	0.632	0.307	0.025	1.239	1.69	6.94
Anthocyanin	0.675	6.707	0.899	0.808	2.10	7.54
DS2500 (408-2492 nm) 0.5 nm						
β -glucan	0.920	0.435	-0.048	0.963	3.56	13.13
Arabinoxylan	0.578	0.868	-0.325	0.821	1.38	5.67
Bound Phenol	0.842	0.449	-0.046	0.988	2.63	9.93
Free Phenol	0.635	0.327	0.031	0.821	1.59	6.51
Anthocyanin	0.541	8.688	-0.220	0.682	1.62	5.82
DS2500 (408-1092 nm) 0.5 nm						
β -glucan	0.766	0.754	0.023	0.962	2.06	7.57
Arabinoxylan	0.545	0.833	-0.248	0.871	1.44	5.91
Bound Phenol	0.764	0.552	-0.017	0.922	2.14	8.08
Free Phenol	0.624	0.328	0.091	0.935	1.59	6.49
Anthocyanin	0.630	8.837	-0.361	0.887	1.59	5.73
DS2500 (1108-2492 nm) 0.5 nm						
β -glucan	0.937	0.394	0.013	1.014	3.93	14.49
Arabinoxylan	0.459	0.825	-0.112	0.739	1.45	5.96
Bound Phenol	0.830	0.471	0.004	1.092	2.51	9.47
Free Phenol	0.637	0.324	0.046	0.849	1.60	6.57
Anthocyanin	0.471	9.644	-2.210	0.834	1.46	5.25

R²_{VAL} coefficient of determination for external validation; SEP standard error of prediction; RPD ratio of performance to deviation (SD/SEP); RER range error ratio (Range/SEP).