Microchemical Journal 220 (2026) 116378

Contents lists available at ScienceDirect

Microchemical Journal

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

Smart screening of durum wheat: a comparative study of NIR portable sensors for protein and moisture content

M. Monti [a], S. Pellacani [b], L. Strani [b], A. D'Alessandro [c], M. Cocchi [b], B. Giussani [a,*], C. Durante [b]

[a] Science and High Technology Department, University of Insubria, Via Valleggio 9, 22100 Como, Italy
[b] Department of Chemical and Geological Sciences, University of Modena and Reggio Emilia, Via Campi 103, 41125 Modena, Italy
[c] Chemometrics Science Group Research, Development and Quality, BARILLA G. e R. Fratelli, Societa per Azioni, Via Mantova 166, 43122 Parma, Italy

ABSTRACT

Durum wheat (*Triticum durum* Desf.) is a strategic crop for the Italian agro-industrial sector, where rapid, non-destructive assessment of protein and moisture content is crucial for quality evaluation and processing performance. Portable Near-Infrared (NIR) spectrometers provide a promising approach for in-field and in-line analyses, overcoming limitations of conventional laboratory methods.

This study evaluates two portable devices, NeoSpectra Scanner and PoliSPEC-NIR, for predicting protein and moisture in intact durum wheat kernels. Over 100 samples from various Italian regions were analyzed using different acquisition modes, and Partial Least Squares (PLS) regression models were developed and externally validated to predict protein content and moisture in samples.

Both sensors demonstrated satisfactory predictive capabilities, with higher accuracy for protein and best results achieved by PoliSPEC-NIR in scanning mode (RMSEP 0.35 g/100 g for protein content prediction and 0.21 g/100 g for moisture prediction). Acquisition mode and surface coverage significantly influenced model robustness, highlighting the need for standardized measurement protocols. These findings support the operational feasibility of miniaturized NIR spectrometers for rapid, non-destructive quality monitoring along the durum wheat value chain.

1. Introduction

Durum wheat (*Triticum durum* Desf.) is a strategic crop for the Italian agro-industrial sector, as it represents the primary raw material for high-quality pasta production [1]. Determining essential compositional parameters, such as protein and moisture content, is critical for qualitative grain classification and evaluating its technological value, directly influencing processing performance and final product quality [2,3].

Although reference analytical methods (e.g., Kjeldahl method for protein determination and gravimetric or chemical methods, such as Karl Fischer titration or coulometry, for moisture analysis) ensure high precision, they are time-consuming, require chemical reagents and laboratory equipment, and are unsuitable for operational contexts that demand rapid and non-destructive analyses [4–6].

In recent years, Near-Infrared spectroscopy (NIR) has gained considerable attention in cereal science as a rapid, non-invasive, and multiparametric analytical technique [7–9]. However, the routine application of conventional benchtop NIR spectrometers remains limited due to their size, operational complexity, and high cost. The emergence and growing availability of miniaturized Near-Infrared (NIR) spectrometers represent a significant advancement in analytical technologies applied to food and feed sciences, enhancing the applicability of NIR spectroscopy, a powerful, non-destructive technique well-suited for the rapid characterization of complex matrices, including both unprocessed foodstuffs (e.g., grains, fruits, raw meat, milk) and highly processed products (e.g., flours and protein powders), as well as their individual constituents such as moisture, fat, protein, or sugar content [10–17]. A growing body of literature underscores the expanding application of portable NIR spectrometers for both in-field and in-line analyses, offering a practical response to several limitations inherent to conventional laboratory-based instrumentation [18–20]. Importantly, the use of miniaturized NIR devices is fully consistent with the principles of green analytical chemistry, enabling rapid, non-invasive, and sustainable measurements directly at the point of need, while eliminating reagent consumption and reducing environmental impact [21–23].

Currently available portable NIR spectrometers cover a wide spectrum of technical specifications and instrumental configurations, enabling their application across a range of analytical tasks and matrix types with diverse physicochemical characteristics, even in contexts where data variability and preprocessing strategies play a critical role [24,25]. Therefore, comparative evaluation of analytical performance

---

* Corresponding author.
E-mail address: barbara.giussani@uninsubria.it (B. Giussani).

https://doi.org/10.1016/j.microc.2025.116378
Received 22 October 2025; Received in revised form 22 November 2025; Accepted 26 November 2025

and selecting the most suitable instrumentation are especially critical to effectively integrate these technologies within the agri-food supply chain [26,27].

In this context, the present study compares the analytical performance of two portable NIR spectrometers, NeoSpectra Scanner and PoliSPEC-NIR, in the non-destructive analysis of intact durum wheat kernels for predicting protein and moisture content [28–31]. Over 100 samples collected from various Italian regions were analyzed using different acquisition modes (point-based, rotary, and scanning) to evaluate the predictive effectiveness of chemometric models developed via Partial Least Square (PLS) regression. The ultimate goal is to assess the feasibility of operationally deploying these devices for rapid screening applications aligned with the needs of precision agriculture and digitalization of the cereal value chain.

2. Materials and methods

2.1. Samples

The analyzed durum wheat kernel samples originated from various Italian regions, ensuring a broad representation of the geographical and agronomic variability of the territory. A total of 103 samples were collected, distributed as follows: Emilia-Romagna (30 samples), Marche (23), Tuscany (11), Puglia (10), Lombardy (8), Veneto (6), Umbria (6), Lazio (4), Campania (3), while Molise and Sicily contributed with 1 sample each. The regional distribution of the samples is illustrated in Fig. 1. The variety in origin contributed to the overall variability observed in the samples, allowing for an in-depth evaluation of the properties of interest. One of the target parameters considered was moisture content, expressed in g/100 g, which ranged between 8.83 g/100 g and 12.29 g/100 g, with an average of 10.35 g/100 g. Protein content, a crucial parameter for determining the technological quality of wheat, ranged between 11.12 and 17.28 g/100 g, with an average of 13.40 g/100 g (Table 1).

All samples were stored at room temperature in sealed containers to ensure stability and preserve their physicochemical characteristics until analysis. No preliminary treatments were performed on the samples before spectroscopic measurements.

2.2. NIR sensors

Since miniaturized spectrometers on the market exhibit different technological characteristics and acquisition configurations, the main features and analysis details of the two NIR devices used in this study are presented. The NeoSpectra Scanner and PoliSPEC-NIR vary in spectral range, resolution, optical geometry, and data acquisition methods. These differences also extend to hardware components, calibration procedures, and data processing approaches, necessitating specific

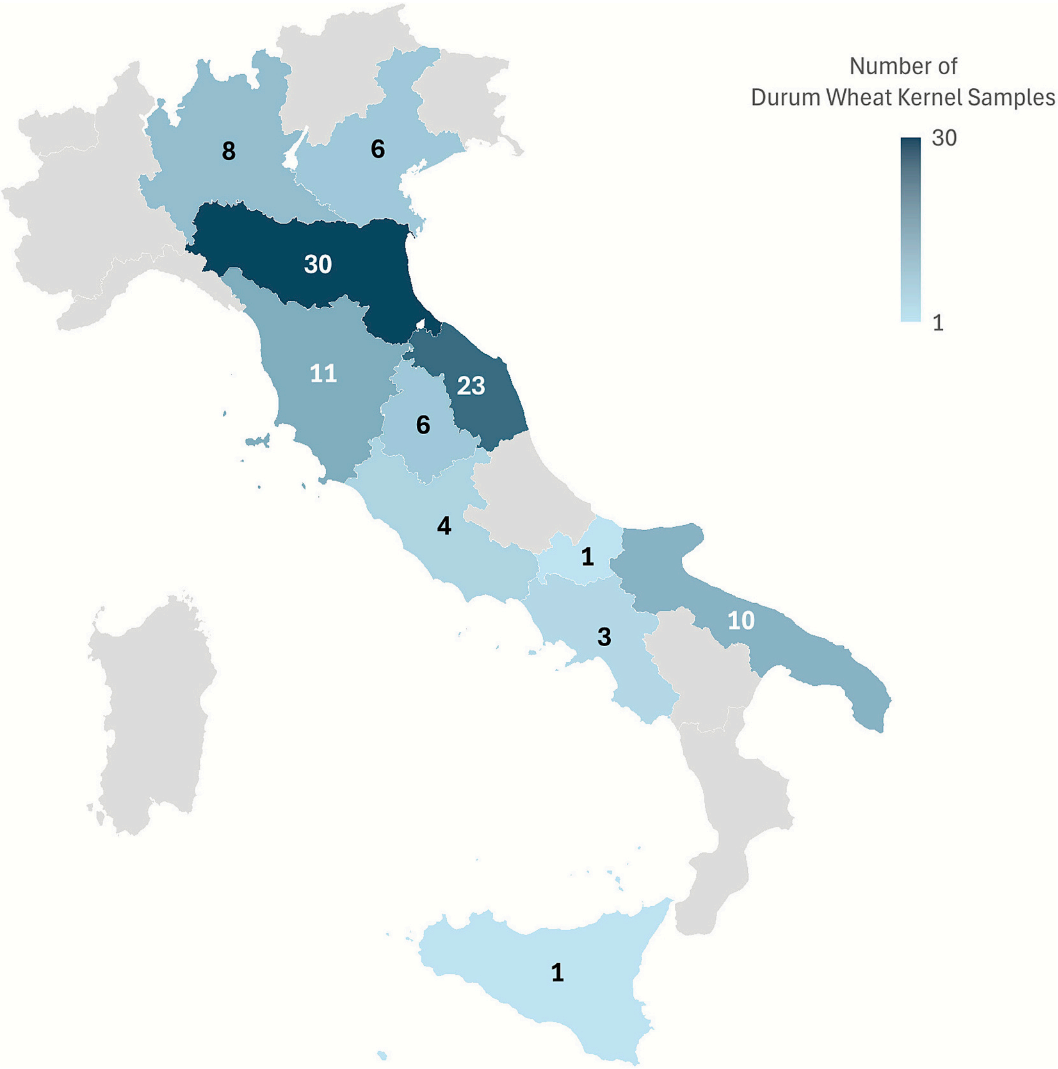

Fig. 1. Number of durum wheat kernel samples collected per Italian region. Darker colors indicate regions with a higher number of samples.

Table 1
Descriptive statistics of protein and moisture content in the calibration and validation sets used for model development. N°: number of samples; Min (g/100 g): minimum; Max (g/100 g): maximum; Avg (g/100 g): average; STD (g/100 g): standard deviation.

	Sample set	N°	Min (g/100 g)	Max (g/100 g)	Avg (g/100 g)	STD (g/100 g)
Protein	Calibration set	73	11.12	17.28	13.40	0.96
	Validation set	30	11.59	17.01	13.42	1.13
Moisture	Calibration set	73	8.83	11.36	10.28	0.53
	Validation set	30	9.30	12.29	10.54	0.70

optimization for each instrument. In general, the NeoSpectra Scanner, being a miniaturized device, is more compact and lightweight, making it suitable for rapid and mobile analyses, whereas the PoliSPEC-NIR, classified as a handheld spectrometer, offers greater spectral resolution and stability. A summary of their key technical specifications and acquisition settings is provided in Table 2.

2.3. NeoSpectra Scanner

The NIR spectra of the kernel samples were acquired using the NeoSpectra Scanner (Si-Ware, Cairo, Egypt), a miniaturized portable spectrophotometer with dimensions of 178 mm × 91 mm × 62 mm and a weight of approximately 730 g. The device operates within a wavelength range of 1350–2500 nm (7400–4000 cm^{-1}) with a spectral resolution of 16 nm and is equipped with a MEMS Michelson interferometer, a single InGaAs photodetector, and seven halogen lamps as the illumination source. The spectrometer was controlled via a proprietary mobile application installed on a Lenovo Smart Tab M8 Android tablet, connected via Bluetooth. Each spectrum was recorded with a scan time of 5 s, without data interpolation, and the instrument was used exclusively while charging. Calibration was performed before each measurement using a 100 % reflectance reference with a Spectralon® standard. The acquired spectral data were automatically saved in ".Spectrum" format, converted into an Excel file, and subsequently processed using MATLAB R2024b (MathWorks Inc., Natick, MA, USA) and PLS_Toolbox Version 9.5.0 (Eigenvector Inc., Manson, WA, USA). Before measurement, kernel samples were analyzed in their intact form without grinding or other physical treatments, in line with non-destructive testing practices suitable for industrial applications. They were simply removed from their storage packaging and subjected to gentle, non-invasive cleaning to eliminate big visible residues that could interfere with spectral accuracy. For the point-based acquisition mode, each sample was divided into an aliquot of approximately 40 g, which was transferred into a 100 mL becker to facilitate handling and avoid direct contact. The kernel samples were then manually positioned on the spectrometer sensor and distributed as evenly as possible. In this configuration, 15 spectral replicates were acquired for each sample, with the kernel samples being repositioned between each replicate to capture intra-sample variability.

The optical window of the NeoSpectra Scanner has a diameter of 10 mm, corresponding to an active measurement area of 79 mm^2 as declared by the producer, which limits the spatial coverage during each acquisition and emphasizes the importance of sample repositioning and the replicate measurements to account for heterogeneity.

For the rotary acquisition mode, the full amount of kernel per sample, approximately 200 g, was loaded into the rotor accessory, which gently rotated the sample above the sensor during spectral acquisition. This dynamic presentation aimed to enhance representativity by exposing different portions of the sample to the detector surface. In both modes, the measurement protocol was identical in terms of the number of replicates (15), scan time, and instrument settings. The only differences concerned the kernel quantity used and the employment of the rotor accessory in the second mode.

2.4. PoliSPEC-NIR

The kernel samples were also analyzed with the PoliSPEC-NIR (ITPhotonics, Italy), a handheld and robust spectrophotometer designed for both portable applications and in-line process monitoring. It operates within the near-infrared spectral range of 900–1700 nm. The instrument is housed in durable anodized aluminum and features an integrated front heatsink to efficiently dissipate heat from both the illumination system and the actively cooled InGaAs detector. Designed primarily for reflectance measurements, the PoliSPEC-NIR utilizes a diffuse 0° optical geometry and includes an automatic internal reference system, which can also be manually controlled via dedicated buttons in portable mode. Spectral data were acquired with a numerical resolution of 3.2 nm and an optical resolution (HWHM) of 3.25 nm. As with the NeoSpectra Scanner, measurements were performed under stable operating conditions. In this case, the instrument was powered via a 12 Vdc supply and operated using its internal rechargeable battery to ensure portability. Similarly, samples were analyzed in their whole kernel form, without grinding or structural alteration, to maintain their native characteristics. Two acquisition methods were applied. In the point-based acquisition mode, an aliquot of each sample was poured into a plastic sample holder, ensuring a uniform filling to maintain consistent scattering effects across measurements. Three spectra were then acquired by positioning the instrument at three different positions, repeating the procedure on three separate extractions from the same bag, resulting in a total of nine spectra per sample. Compared to the NeoSpectra, the PoliSPEC-NIR features a larger optical window with a diameter of 20 mm, corresponding to an active measurement area of approximately 314 mm^2. This surface is nearly four times larger than that of the NeoSpectra Scanner (79 mm^2), resulting in a broader sampling area. Such a wider detection surface improves the representativity

Table 2
Key technical specifications and characteristics of the selected miniaturized NIR spectrometers.

Spectrometer (Vendor)	Source	Wavelength Selector	Window size (Area)	Detector	Wavelength Region	Spectral Resolution (at Wavelength)	Weight (Dimension)	Connectivity	Approximate Order of Unit Cost ($)
NeoSpectra Scanner (Si-Ware Systems)	Tungsten halogen	MEMS Michelson interferometer	79 mm^2	InGaAs (single element)	1350–2500 nm	16 nm	730 g (17.8 × 9.1 × 6.2 cm)	Bluetooth	<9000
PoliSPEC-NIR (ITPhotonics)	Tungsten halogen	Fixed grating	314 mm^2	InGaAs (256 pixels, Peltier-cooled)	900–1700 nm	3.25 nm	3.2 kg (21.6 × 21.3 × 8.5 cm)	WiFi + RS422/RS485 (optional Ethernet)	<29,000

of each single measurement and potentially reduces the impact of sample inhomogeneity.

In the scanning acquisition mode, which has been done manually but could be automatized in routine in-situ analysis, the sample was poured in a rectangular container with a glass bottom, ensuring a uniform filling. The spectral acquisition was performed by moving the instrument continuously over the container surface along the longer dimension for ten seconds, recording approximately forty spectra per sample, with minimal variations depending on the scanning speed. To ensure data consistency and reduce variability, the operation was repeated twice consecutively for each sample.

3. Reference value determination for protein and moisture

3.1. Protein content determination

Protein content in grains was determined following the method UNI EN ISO 16634-2:2016. The determination of the total nitrogen content was based on combustion according to the Dumas' principles: the sample (about 80 mg precisely weight at 0.1 mg precision) was burnt in a suitable high temperature system in presence of oxygen and resulting nitrogen oxides were reduced by catalyst to molecular nitrogen. From the combustion gases were selectively removed by traps oxygen, water and carbon dioxide. The remaining nitrogen is then quantified by using gas chromatography with a thermal conductivity detector. For cereals the conversion factor from nitrogen to proteins to be used is 5.7. The protein content is typically expressed as (Nx5.7) g/100 g on dry matter. Analysis was done in triplicate for each sample. The estimated error for this reference technique is 0.4 g/100 g for protein.

3.2. Moisture determination

Moisture of grains was evaluated according to the UNI EN ISO 712-1:2024 method. In synthesis: wheat grains samples were coarsely grinded, avoiding overheating, then about 10 g of grounded products was exactly weighted with a precision of 1 mg in a porcelain plate and put in the oven at 130 °C until a constant weight was reached. The difference in weight reported to 100 g represents the moisture of the sample. Analysis was done in duplicate for each sample. Error of 0.3 g/100 g for moisture was estimated for this reference technique.

3.3. Statistical data analysis

The spectral data acquired from the NeoSpectra Scanner and the PoliSPEC-NIR were analyzed using Partial Least Squares Regression (PLS) to develop predictive models for two key quality parameters of durum wheat kernel samples: moisture and protein content. Prior to model calibration, various spectral preprocessing strategies were tested to improve data quality, reduce baseline shifts, correct for scattering effects, and enhance relevant chemical information. Different preprocessing combinations were evaluated, involving techniques such as smoothing, standard normal variate (SNV) [32], variable sorting for normalization (VSN) [33], multiplicative scatter correction (MSC) [34], mean centering, detrending, and first or second derivatives (e.g., Savitzky-Golay) [35]. It is important to note that the selection of the optimal pre-treatment is based on the analysis of experimental replicates. The best pre-treatment is the one that minimizes differences between spectroscopic replicates. This approach is grounded in the principle that an effective pre-treatment should remove all sources of variation between replicates—except for the underlying chemical information, which should remain consistent.

Practically, in order to select the optimal preprocessing, the overall acquired spectra, after each spectral preprocessing, were mean centered per sample and subjected to PCA retaining the number of PCs explaining the 95 % of variance. In this way, the origin in scores space corresponds to the average value of the sample replicates and the closer the replicates in the scores space are, the more effective is the preprocessing. Then, the interquartile range (IQR) relative to the reduced Hotellings-T^2 distance for each sample is calculated, the most performing preprocessing corresponds to lower median values and smaller box size.

External validation was used. To divide the dataset into calibration and validation sets, a sample selection procedure based on the duplex algorithm run on the preprocessed spectral data matrix (X) augmented with references values data (Y) was used [36]. This method ensures a representative distribution of both the independent variables (X: spectral data) and the dependent variables (Y: reference values), supporting robust model development and generalization.

The dataset was divided into 73 samples for model calibration and 30 samples for external validation. Once the sample sets were defined (using data collected with the PoliSPEC-NIR sensor), the same calibration and validation subsets were consistently used across all models independently of the instrument or acquisition method to allow a fair and direct comparison between the NeoSpectra Scanner and the PoliSPEC-NIR, as well as between point-based and scanning/rotary acquisition modes.

First model evaluation was performed through cross-validation, using a random subset approach with 10 splits and 20 iterations, which provides a balanced assessment of predictive performance and allow the choice of the correct number of latent variables to avoid overfitting. Key metrics used to assess the models included the coefficient of determination (R^2), the Root Mean Square Error of Calibration (RMSEC), the Root Mean Square Error of Cross-Validation (RMSECV), and the Root Mean Square Error of Prediction (RMSEP) for external validation. These indicators were used to compare the predictive accuracy and robustness of the models developed under different configurations, ensuring the reliability of the extracted chemical information from the NIR spectra. The choice and reporting of these statistical indicators were carried out in accordance with established recommendations for analytical calibration and data reporting [37].

4. Results and discussion

4.1. Spectra interpretation

Fig. 2 shows the raw NIR spectra acquired from the two spectrometers, NeoSpectra Scanner and PoliSPEC-NIR. The two instruments operate in different spectral ranges, with the NeoSpectra Scanner covering the 1350–2500 nm region and the PoliSPEC-NIR working within 900–1700 nm. The graphs reveal slight differences in signal intensity and spectral characteristics if comparing the overlapped region, which can be attributed to the distinct optical configurations of the instruments, including illumination geometry, spectral resolution and detector characteristics. Both spectra exhibit absorption bands associated with key wheat components, such as water and proteins. The analysis of the most relevant spectral regions revealed substantial differences between the two instruments, attributable to both their spectral coverage and optical resolution. In the near-infrared region, absorption features mainly arise from overtone and combination vibrations of O—H, fundamental modes. Higher-order overtones (second, third and above) are intrinsically weak and contribute only marginally to NIR spectra. [38,39] As a consequence, the most relevant contributions in the spectra originate from first overtones and combination bands of these groups. For the spectroscopic range covered by NeoSpectra Scanner, the most influential wavelengths were located in the ranges of 1450–1470 nm, 1900–1950 nm, 2050–2180 nm, and 2300–2350 nm. The regions around 1450–1470 nm and 1900–1950 nm correspond to the first overtone of O—H stretching and the O—H stretching + bending combination band of free and bound water, respectively. The interval 2050–2180 nm and the region between 2300 and 2350 nm are generally associated with combination bands involving N—H and C-H/C=O groups, which reflect contributions from proteins and other organic constituents. In the case of the range covered by PoliSPEC-NIR, the most

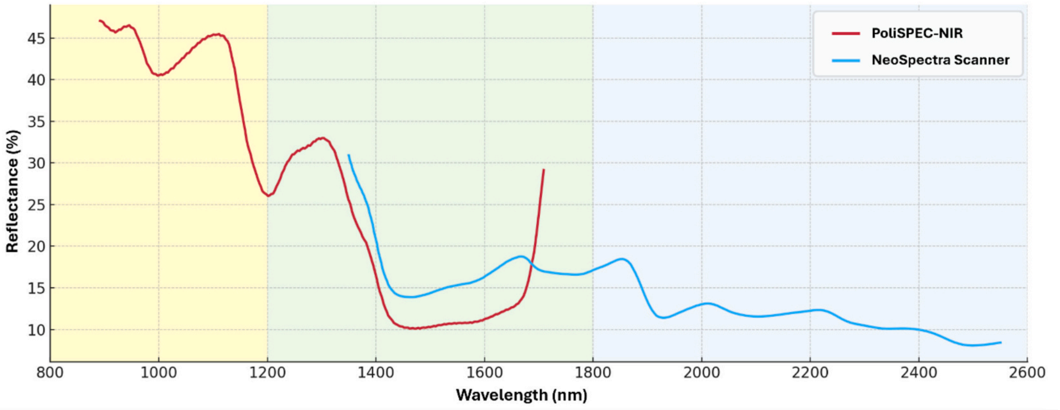

Fig. 2. Average spectra of durum wheat kernel acquired using different spectrometers and optical configurations. The background colors highlight the distinct NIR regions: Region 1 (800–1200 nm) in yellow, Region 2 (1200–1800 nm) in green, and Region 3 (1800–2600 nm) in light blue. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

predictive spectral regions were identified within the 1100–1200 nm and 1300–1450 nm ranges. Specifically, the region around 1420 nm is associated with the first overtone of O—H stretching, indicative of free moisture, while the band between 1370 and 1390 nm corresponds to C—H combination bands (CH_2, CH_3) typically associated with aliphatic structures, including those related to proteins [40–44]. Overall, the spectral differences observed between instruments reflect their wavelength coverage, optical resolution and sampling characteristics, and the identified regions are consistent with the vibrational behavior of water, proteins and structural organic components in durum wheat kernels.

4.2. *Prediction of protein content*

The prediction of protein content in wheat samples was performed using Partial Least Squares Regression (PLS) applied to spectral data acquired with the NeoSpectra Scanner and PoliSPEC-NIR. Given the differences in the two sensors, independent models were developed for each spectrometer and acquisition mode to assess their predictive performance. Fig. 3 resumes the results from the best models, while regression coefficients are reported in supplementary material.

In the point-based mode with the NeoSpectra Scanner, the calibration model was built by selecting 5 latent variables (LVs) according to the minimum RMSECV. The calibration phase yielded an R^2 value of 0.82, with RMSEC of 0.37 g/100 g and RMSECV of 0.49 g/100 g. Validation resulted in an RMSEP of 0.59 g/100 g, and a bias of −0.0044 g/100 g. In the rotary mode, the best calibration model was achieved using 4 latent variables, giving a calibration R^2 value of 0.79, RMSEC of 0.48 g/100 g, and RMSECV of 0.64 g/100 g. Validation yielded an RMSEP of 0.47 g/100 g, and a bias of −0.072 g/100 g.

The PoliSPEC-NIR was tested using two different acquisition methods: point-based and scanning. The point-based mode, similar to that used with the NeoSpectra Scanner, involved collecting spectra from a fixed position on the sample. In this mode, the calibration model was developed using 73 samples, selecting 6 latent variables (LVs) according to the minimum RMSECV. The calibration phase yielded an R^2 of 0.82, with RMSEC of 0.40 g/100 g and RMSECV of 0.51 g/100 g. Validation on 30 independent samples resulted in an RMSEP of 0.43 g/100 g and a bias of 0.002 g/100 g. In the scanning mode, the calibration model was built with 7 latent variables, achieving a calibration R^2 of 0.86, with RMSEC of 0.35 g/100 g and RMSECV of 0.48 g/100 g. Validation on 30 independent samples yielded an RMSEP of 0.35 g/100 g and a bias of 0.046 g/100 g.

The calibration and validation result for both spectrometers and acquisition modes are summarized in Table 3, providing an overview of their predictive performance in protein content estimation.

4.3. *Prediction of moisture content*

To estimate the moisture content in durum wheat kernel samples, Partial Least Squares Regression (PLS) was applied to the spectral data obtained from the NeoSpectra Scanner and PoliSPEC-NIR devices. Fig. 4 summarizes the results of the best-performing models, while the regression coefficients are provided in the supplementary material.

Concerning the models obtained with the NeoSpectra Scanner sensor, in the point-based mode, the optimum PLS calibration model was constructed using 3 latent variables (LVs). The calibration phase yielded an R^2 value of 0.67, an RMSEC of 0.33 g/100 g, and an RMSECV of 0.37 g/100 g. Validation resulted in an R^2 of 0.38, an RMSEP of 0.38 g/100 g, and a bias of −0.18 g/100 g. In the rotary mode, the best calibration model was obtained with 5 latent variables, yielded an R^2 value of 0.61, an RMSEC of 0.38 g/100 g, and an RMSECV of 0.43 g/100 g. Validation resulted in an R^2 of 0.38, an RMSEP of 0.38 g/100 g, and a bias of −0.12 g/100 g.

Concerning the PoliSPEC-NIR spectrometer, in the point-based configuration the calibration model was generated using 5 latent variables (LVs) according to the minimum RMSECV, yielded a calibration R^2 value of 0.80, with RMSEC of 0.23 g/100 g and RMSECV of 0.28 g/100 g. Validation resulted in an RMSEP of 0.23 g/100 g and a bias of 0.007 g/100 g. In the scanning mode, the calibration model was developed using 6 latent variables, achieving a calibration R^2 of 0.82, with RMSEC of 0.22 g /100 g and RMSECV of 0.29 g/100 g. Validation gave an RMSEP of 0.21 g/100 g and a bias of 0.013 g/100 g.

All calibration and validation results for the moisture models, across both spectrometers and acquisition modes, are compiled in Table 4.

4.4. *General discussion*

The results obtained from both sensors are encouraging and promising. Considering the importance of being able to perform rapid measurements outside the laboratory, a slightly higher error could be acceptable. Just imagine the possibility of taking sample measurements at a manufacturing site or in a storage center. In these situations, the speed of the analysis—without any sample preparation or preprocessing—could compensate for a minor loss of accuracy. Another potential application could be at the truck unloading point at the mill, where rapid pre-screening could help determine in which silos to store the incoming material. Moreover, both instruments showed prediction errors that were not significantly different from those of the official reference methods (reported in the Protein content determination and in Moisture determination paragraph), with errors of 0.3 g/100 g for moisture and 0.4 g/100 g for protein.

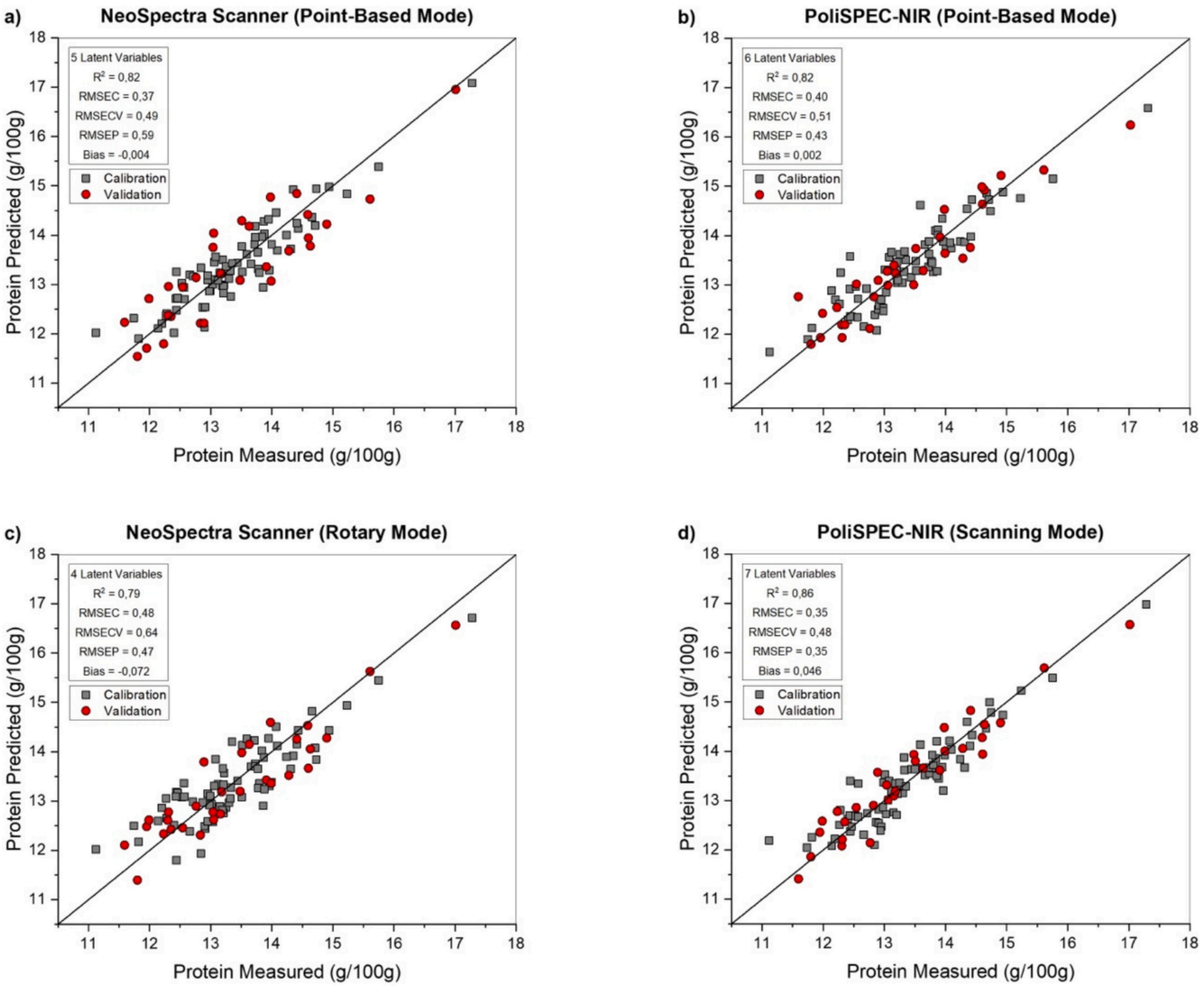

Fig. 3. **a)** Predicted vs. measured protein content for the NeoSpectra Scanner (point-based mode); **b)** Predicted vs. measured protein content for the PoliSPEC-NIR (point-based mode); **c)** Predicted vs. measured protein content for the NeoSpectra Scanner (rotary mode); **d)** Predicted vs. measured protein content for the PoliSPEC-NIR (scanning mode).
R^2: Coefficient of Determination; RMSEC: Root Mean Square Error of Calibration (g/100 g); RMSECV: Root Mean Square Error of Cross-Validation (g/100 g); RMSEP: Root Mean Square Error of Prediction (g/100 g).

Table 3
Calibration and validation statistics of partial least squares models for protein prediction using different preprocessing methods and acquisition modes for the two instruments.

Instrument	Mode	Y	Preprocessing	LV	R^2	RMSEC (g/100 g)	RMSECV (g/100 g)	RMSEP (g/100 g)	Bias (g/100 g)
NeoSpectra Scanner	Point-based mode	Protein	WLS + MSC + Mean Center	5	0.82	0.37	0.49	0.59	−0.004
NeoSpectra Scanner	Rotary mode	Protein	WLS + MSC + Mean Center	4	0.79	0.48	0.64	0.47	−0.072
PoliSPEC-NIR	Point-based mode	Protein	EMSC + Mean Center	6	0.82	0.40	0.51	0,43	0.002
PoliSPEC-NIR	Scanning mode	Protein	2nd derivative (SG order 2; Window 15) + Mean Center	7	0.86	0.35	0.48	0.35	0.046

LV: Latent Variables; R^2: Coefficient of Determination; RMSEC: Root Mean Square Error of Calibration; RMSECV: Root Mean Square Error of Cross-Validation; RMSEP: Root Mean Square Error of Prediction; WLS: Weighted Least Squares baseline correction; MSC: Multiplicative Scatter Correction; EMSC: Extended Multiplicative Scatter Correction; SG: Savitzky–Golay filter.

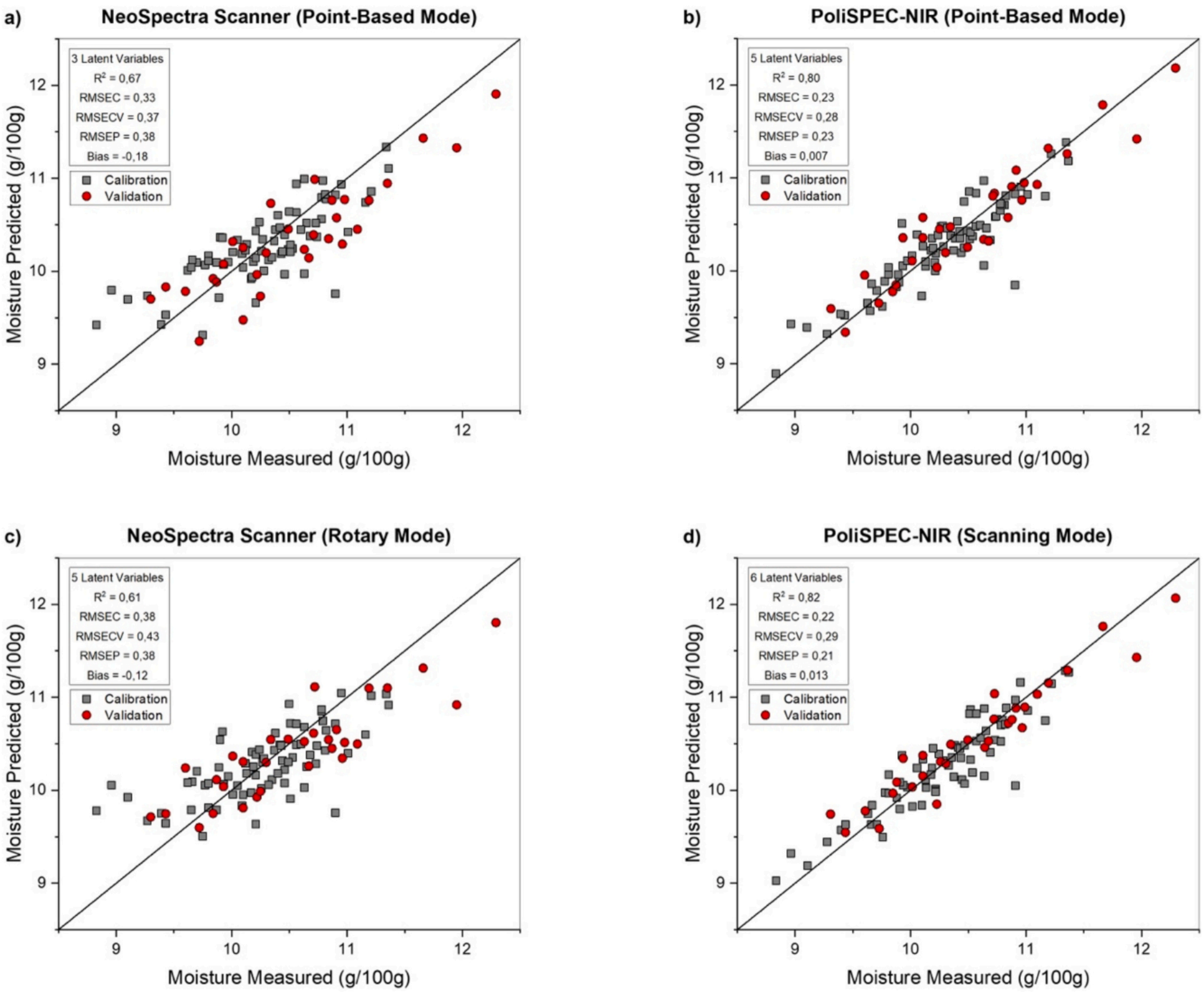

Fig. 4. **a)** Predicted vs. measured moisture content for the NeoSpectra Scanner (point-based mode); **b)** Predicted vs. measured moisture content for the PoliSPEC-NIR (point-based mode); **c)** Predicted vs. measured moisture content for the NeoSpectra Scanner (rotary mode); **d)** Predicted vs. measured moisture content for the PoliSPEC-NIR (scanning mode).

R^2: Coefficient of Determination; RMSEC: Root Mean Square Error of Calibration (g/100 g); RMSECV: Root Mean Square Error of Cross-Validation (g/100 g); RMSEP: Root Mean Square Error of Prediction (g/100 g).

Table 4

Calibration and validation statistics of partial least squares models for moisture prediction using different preprocessing methods and acquisition modes for the two instruments.

Instrument	Mode	Y	Preprocessing	LV	R^2	RMSEC (g/100 g)	RMSECV (g/100 g)	RMSEP (g/100 g)	Bias (g/100 g)
NeoSpectra Scanner	Point-based mode	Moisture	WLS + MSC + Mean Center	3	0,67	0,33	0,37	0,38	−0,18
NeoSpectra Scanner	Rotary mode	Moisture	Smoothing (Window 11) + Mean Center	5	0,61	0,38	0,43	0,38	−0,12
PoliSPEC-NIR	Point-based mode	Moisture	2nd derivative (SG order 2; Window 15) + Mean Center	5	0,80	0,23	0,28	0,23	0,007
PoliSPEC-NIR	Scanning mode	Moisture	2nd derivative (SG order 2; Window 15) + Mean Center	6	0,82	0,22	0,29	0,21	0,013

LV: Latent Variables; R^2: Coefficient of Determination; RMSEC: Root Mean Square Error of Calibration; RMSECV: Root Mean Square Error of Cross-Validation; RMSEP: Root Mean Square Error of Prediction; WLS: Weighted Least Squares baseline correction; MSC: Multiplicative Scatter Correction; EMSC: Extended Multiplicative Scatter Correction; SG: Savitzky–Golay filter.

It is worth to recall that the two devices differ significantly in their characteristics. The first, the NeoSpectra spectrometer, is a miniaturized device with an acquisition window of approximately 1 cm^2. This small window size limits its capacity to capture the inhomogeneities within a sample such as wheat grains, it is therefore essential for the NeoSpectra Scanner to analyze several aliquots of the sample in order to capture as much of the sample variability as possible. In contrast, the PoliSPEC-NIR sensor features a much larger acquisition window, enabling it to better capture and average the intrinsic variability of such heterogeneous samples. This difference likely explains why the models developed using the PoliSPEC-NIR sensor typically show lower prediction errors compared to those based on the NeoSpectra Scanner instrument. Beyond the technical specifications—namely, the NeoSpectra Scanner being a miniaturized scanner and the PoliSPEC-NIR a handheld device—there is also a notable difference in their price ranges.

Regarding the NeoSpectra Scanner sensor, there does not appear to be a clear trend associated with the use of the automated rotating system, at least for the determination of moisture and protein contents in this type of sample. Concerning the PoliSPEC-NIR sensor, the use of scanning mode appears to slightly enhance model performance compared to the point-based mode. For both moisture and protein content prediction, the results are improved, although this comes at the cost of an increased number of factors.

It is important to note that, in practical applications, the NeoSpectra Scanner sensor requires only a few grams of sample per analysis, whereas the PoliSPEC-NIR instrument demands a significantly larger sample volume, particularly in scanning mode, where the device is physically swiped across the sample.

5. Conclusions

The results of this study highlight the potential of portable NIR spectrometers for the non-destructive estimation of protein and moisture content in intact durum wheat kernels. Both instruments, NeoSpectra Scanner and PoliSPEC-NIR, demonstrated satisfactory predictive capabilities, with better performance observed for protein than for moisture, and generally higher accuracy achieved with the PoliSPEC-NIR. The different acquisition modes tested (point-based, rotary/scanning) affected model quality, suggesting that sample homogenization and greater surface coverage improve predictive robustness. Notably, the scanning mode yielded the best results for the PoliSPEC-NIR, with lower RMSEP values compared to other configurations - RMSEP 0.35 g/100 g for protein content prediction and 0.21 g/100 g for moisture prediction. Overall, the developed PLS models showed good accuracy and generalizability, supporting the feasibility of deploying miniaturized spectrometers in operational contexts for rapid screening applications. Nevertheless, predictive accuracy remains influenced by instrumental and procedural factors, highlighting the need for targeted calibration and standardized measurement protocols for field use. In conclusion, portable NIR devices represent a promising solution for quality monitoring of durum wheat along the value chain, supporting more efficient decision-making in agricultural and industrial settings, in line with the principles of precision agriculture and the digital transformation of the agri-food sector.

CRediT authorship contribution statement

M. Monti: Writing – original draft, Validation, Methodology, Investigation, Formal analysis. **S. Pellacani:** Investigation, Formal analysis. **L. Strani:** Validation, Investigation, Formal analysis. **A. D'Alessandro:** Resources, Methodology, Funding acquisition. **M. Cocchi:** Writing – review & editing, Supervision, Methodology, Funding acquisition, Data curation, Conceptualization. **B. Giussani:** Writing – review & editing, Supervision, Methodology, Conceptualization. **C. Durante:** Supervision, Methodology, Data curation.

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.

Acknowledgements

The authors would like to express their deepest gratitude to Quantum Design Italy (https://qdindustria.it/, accessed on 1 October 2025) that provided a PoliSPEC-NIR portable instrument to perform NIR measurements.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.microc.2025.116378.

Data availability

Data will be made available on request.

References

[1] M. Sissons, Role of Durum Wheat Composition on the Quality of Pasta and Bread, 2025.

[2] A. Khalid, A. Hameed, M.F. Tahir, Wheat quality: a review on chemical composition, nutritional attributes, grain anatomy, types, classification, and function of seed storage proteins in bread making quality, Front. Nutr. 10 (2023), https://doi.org/10.3389/fnut.2023.1053196.

[3] A. Tabatabaeefar, Moisture-dependent physical properties of wheat, Int. Agrophysics (2003) 207–211.

[4] P. Sáez-Plaza, T. Michałowski, M.J. Navas, A.G. Asuero, S. Wybraniec, An Overview of the Kjeldahl Method of Nitrogen Determination. Part I. Early History, Chemistry of the procedure, and titrimetric finish, Crit. Rev. Anal. Chem. 43 (2013) 178–223, https://doi.org/10.1080/10408347.2012.751786.

[5] A. Bennett, J.R. Hudson, Determination of moisture in cereals: review of methods in common use, J. Inst. Brew. 60 (1954) 29–34, https://doi.org/10.1002/j.2050-0416.1954.tb02744.x.

[6] Wernecke Roland, Wernecke Jan, Industrial Moisture and Humidity Measurement, Wiley (2013), https://doi.org/10.1002/9783527652419.

[7] K. Chadalavada, K. Anbazhagan, A. Ndour, S. Choudhary, W. Palmer, J.R. Flynn, S. Mallayee, S. Pothu, K.V.S.V. Prasad, P. Varijakshapanikar, C.S. Jones, J. Kholová, NIR instruments and prediction methods for rapid access to grain protein content in multiple cereals, Sensors 22 (2022), https://doi.org/10.3390/s22103710.

[8] N. Caporaso, M.B. Whitworth, I.D. Fisk, Near-infrared spectroscopy and hyperspectral imaging for non-destructive quality assessment of cereal grains, Appl. Spectrosc. Rev. 53 (2018) 667–687, https://doi.org/10.1080/05704928.2018.1425214.

[9] Z. Du, W. Tian, M. Tilley, D. Wang, G. Zhang, Y. Li, Quantitative assessment of wheat quality using near-infrared spectroscopy: a comprehensive review, Compr. Rev. Food Sci. Food Saf. 21 (2022) 2956–3009, https://doi.org/10.1111/1541-4337.12958.

[10] D. Cozzolino, Near infrared spectroscopy and food authenticity, in: Advances in Food Traceability Techniques and Technologies, Elsevier, 2016, pp. 119–136, https://doi.org/10.1016/B978-0-08-100310-7.00007-7.

[11] J. Riu, G. Gorla, D. Chakif, R. Boqué, B. Giussani, Rapid analysis of milk using low-cost pocket-size NIR spectrometers and multivariate analysis, Foods 9 (2020), https://doi.org/10.3390/foods9081090.

[12] S. Grassi, C. Alamprese, Advances in NIR spectroscopy applied to process analytical technology in food industries, Curr. Opin. Food Sci. 22 (2018) 17–21, https://doi.org/10.1016/j.cofs.2017.12.008.

[13] E. Arias, V. Sierra, N. Prado, P. González, G. Fiorentini, J. Díaz, M. Oliván, Development of a portable near-infrared spectroscopy tool for detecting freshness of commercial packaged pork, Foods 11 (2022) 3808, https://doi.org/10.3390/foods11233808.

[14] C.S.P. Santos, R. Cruz, D.B. Gonçalves, R. Queirós, M. Bloore, Z. Kovács, I. Hoffmann, S. Casal, Non-destructive Measurement of the internal quality of Citrus fruits using a portable NIR device, J. AOAC Int. 104 (2021) 61–67, https://doi.org/10.1093/jaoacint/qsaa115.

[15] K.H.S. Peiris, S.R. Bean, X. Wu, S.A. Sexton-Bowser, T. Tesso, Performance of a handheld MicroNIR instrument for determining protein levels in Sorghum grain samples, Foods 12 (2023) 3101, https://doi.org/10.3390/foods12163101.

[16] F. Tao, L. Liu, C. Kucha, M. Ngadi, Rapid and non-destructive detection of cassava flour adulterants in wheat flour using a handheld MicroNIR spectrometer, Biosyst. Eng. 203 (2021) 34–43, https://doi.org/10.1016/j.biosystemseng.2020.12.010.

[17] M. Lukacs, J.-L.Z. Zaukuu, G. Bazar, B. Pollner, M. Fodor, Z. Kovacs, Comparison of multiple NIR spectrometers for detecting low-concentration nitrogen-based

adulteration in protein powders, Molecules 29 (2024) 781, https://doi.org/10.3390/molecules29040781.

[18] C. Piccini, K. Metzger, G. Debaene, B. Stenberg, S. Götzinger, L. Borůvka, T. Sandén, L. Bragazza, F. Liebisch, In-field soil spectroscopy in Vis–NIR range for fast and reliable soil analysis: a review, Eur. J. Soil Sci. 75 (2024), https://doi.org/10.1111/ejss.13481.

[19] H. Huang, H. Yu, H. Xu, Y. Ying, Near infrared spectroscopy for on/in-line monitoring of quality in foods and beverages: a review, J. Food Eng. 87 (2008) 303–313, https://doi.org/10.1016/j.jfoodeng.2007.12.022.

[20] S. Sohaib Ali Shah, A. Zeb, W.S. Qureshi, M. Arslan, A. Ullah Malik, W. Alasmary, E. Alanazi, Towards fruit maturity estimation using NIR spectroscopy, Infrared Phys. Technol. 111 (2020) 103479, https://doi.org/10.1016/j.infrared.2020.103479.

[21] M. Sajid, J. Płotka-Wasylka, Green analytical chemistry metrics: a review, Talanta 238 (2022) 123046, https://doi.org/10.1016/j.talanta.2021.123046.

[22] Y. He, L. Tang, X. Wu, X. Hou, Y. Lee, Spectroscopy: the best way toward green analytical chemistry? Appl. Spectrosc. Rev. 42 (2007) 119–138, https://doi.org/10.1080/05704920601184259.

[23] P. Anastas, N. Eghbali, Green chemistry: principles and practice, Chem. Soc. Rev. 39 (2010) 301–312, https://doi.org/10.1039/B918763B.

[24] B. Giussani, M. Monti, J. Riu, From spectroscopic data variability to optimal preprocessing: leveraging multivariate error in almond powder adulteration of different grain size, Anal. Bioanal. Chem. 417 (2025) 1393–1405, https://doi.org/10.1007/s00216-024-05710-1.

[25] B. Giussani, S.O. El Ghouch, M. Monti, J. Riu, Understanding variability and calibration challenges in NIR miniaturized sensors: the impact of particle size and analytical session in almond powder analysis, Measurement 256 (2025) 118481, https://doi.org/10.1016/j.measurement.2025.118481.

[26] C.A. Teixeira Dos Santos, M. Lopo, R.N.M.J. Páscoa, J.A. Lopes, A review on the applications of portable near-infrared spectrometers in the agro-food industry, Appl. Spectrosc. 67 (2013) 1215–1233, https://doi.org/10.1366/13-07228.

[27] A.M. Mouazen, W. Saeys, J. Xing, J. De Baerdemaeker, H. Ramon, Near infrared spectroscopy for agricultural materials: an instrument comparison, J. Near Infrared Spectrosc. 13 (2005) 87–97, https://doi.org/10.1255/jnirs.461.

[28] C. Fan, Y. Liu, T. Cui, M. Qiao, Y. Yu, W. Xie, Y. Huang, Quantitative prediction of protein content in corn kernel based on near-infrared spectroscopy, Foods 13 (2024), https://doi.org/10.3390/foods13244173.

[29] K. Chadalavada, K. Anbazhagan, A. Ndour, S. Choudhary, W. Palmer, J.R. Flynn, S. Mallayee, S. Pothu, K.V.S.V. Prasad, P. Varijakshapanikar, C.S. Jones, J. Kholová, NIR instruments and prediction methods for rapid access to grain protein content in multiple cereals, Sensors 22 (2022), https://doi.org/10.3390/s22103710.

[30] T. Ma, G. Zhang, X. Wang, S. Yi, C. Wang, Non-destructive method of small sample sets for the maize moisture content measurement during filling based on NIRS, international journal of agricultural and biological, Engineering 17 (2024) 236–244, https://doi.org/10.25165/j.ijabe.20241704.8738.

[31] R. Zheng, Y. Jia, C. Ullagaddi, C. Allen, K. Rausch, V. Singh, J.C. Schnable, M. Kamruzzaman, Optimizing feature selection with gradient boosting machines in PLS regression for predicting moisture and protein in multi-country corn kernels via NIR spectroscopy, Food Chem. 456 (2024) 140062, https://doi.org/10.1016/j.foodchem.2024.140062.

[32] R.J. Barnes, M.S. Dhanoa, S.J. Lister, Standard Normal variate transformation and De-trending of near-infrared diffuse reflectance spectra, Appl. Spectrosc. 43 (1989) 772–777, https://doi.org/10.1366/0003702894202201.

[33] G. Rabatel, F. Marini, B. Walczak, J. Roger, VSN: variable sorting for normalization, J. Chemom. 34 (2020), https://doi.org/10.1002/cem.3164.

[34] T. Isaksson, T. Næs, The effect of multiplicative scatter correction (MSC) and linearity improvement in NIR spectroscopy, Appl. Spectrosc. 42 (1988) 1273–1284, https://doi.org/10.1366/0003702884429869.

[35] Abraham Savitzky, M.J.E. Golay, Smoothing and differentiation of data by simplified least squares procedures, Anal. Chem. 36 (1964) 1627–1639, https://doi.org/10.1021/ac60214a047.

[36] M. Daszykowski, B. Walczak, D.L. Massart, Representative Subset Selection, 2002.

[37] A.C. Olivieri, Practical guidelines for reporting results in single- and multi-component analytical calibration: a tutorial, Anal. Chim. Acta 868 (2015) 10–22, https://doi.org/10.1016/j.aca.2015.01.017.

[38] J. Grabska, K.B. Beć, Y. Ozaki, C.W. Huck, Anharmonic DFT study of near-infrared spectra of caffeine: vibrational analysis of the second overtones and ternary combinations, Molecules 26 (2021) 5212, https://doi.org/10.3390/molecules26175212.

[39] J. Grabska, K.B. Beć, C.G. Kirchler, Y. Ozaki, C.W. Huck, Distinct difference in sensitivity of NIR vs. IR Bands of Melamine to Inter-Molecular Interactions with Impact on Analytical Spectroscopy Explained by Anharmonic Quantum Mechanical Study, Molecules 24 (2019) 1402, https://doi.org/10.3390/molecules24071402.

[40] S.E. Krikorian, M. Mahpour, The Identification and Origin of N-H Overtone and Combination Bands in the Near-Infrared Spectra of Simple Primary and Secondary Amides, Pergamon Press, 1973.

[41] J.J. Workman, Interpretive spectroscopy for near infrared, Appl. Spectrosc. Rev. 31 (1996) 251–320, https://doi.org/10.1080/05704929608000571.

[42] K.B. Whetsel, Near-infrared spectrophotometry, Appl. Spectrosc. Rev. 2 (1968) 1–67, https://doi.org/10.1080/05704926808050886.

[43] L.G. Weyer, Near-infrared spectroscopy of organic substances, Appl. Spectrosc. Rev. 21 (1985) 1–43, https://doi.org/10.1080/05704928508060427.

[44] M. Schwanninger, J.C. Rodrigues, K. Fackler, A review of band assignments in near infrared spectra of wood and wood components, J. Near Infrared Spectrosc. 19 (2011) 287–308, https://doi.org/10.1255/jnirs.955.