





Article

NIR Instruments and Prediction Methods for Rapid Access to Grain Protein Content in Multiple Cereals

Keerthi Chadalavada ^{1,2,†}, Krithika Anbazhagan ^{1,†}, Adama Ndour ³, Sunita Choudhary ¹, William Palmer ⁴, Jamie R. Flynn ⁴, Srikanth Mallayee ¹, Sharada Pothu ⁵, Kodukula Venkata Subrahmanya Vara Prasad ⁵, Padmakumar Varijakshapanikar ⁵, Chris S. Jones ⁶ and Jana Kholová ^{1,7,*}

- ¹ Crop Physiology & Modeling, International Crops Research Institute for Semi-Arid Tropics, Patancheru, Hyderabad 502 324, India; keerthichadalawada@gmail.com (K.C.); a.krithika@cgiar.org (K.A.); s.choudhary@cgiar.org (S.C.); srikanthmallayee@gmail.com (S.M.)
- ² Department of Botany, Bharathidasan University, Tiruchirappalli 620 024, India
- ³ Crop Physiology & Modeling, International Crops Research Institute for Semi-Arid Tropics, Bamako BP 320, Mali; a.ndour@cgiar.org
- ⁴ Hone, Newcastle, NSW 2300, Australia; william@honeag.com (W.P.); jamie@honeag.com (J.R.F.)
- ⁵ South Asia Regional Center, International Livestock Research Institute, Patancheru 502 324, India; p.sharada@cgiar.org (S.P.); k.v.prasad@cgiar.org (K.V.S.V.P.); v.padmakumar@cgiar.org (P.V.)
- ⁶ Feed and Forage Development, International Livestock Research Institute, Addis Ababa P.O. Box 5689, Ethiopia; c.s.jones@cgiar.org
- ⁷ Department of Information Technologies, Faculty of Economics and Management, Czech University of Life Sciences Prague, 165 00 Prague, Czech Republic
- * Correspondence: j.kholova@cgiar.org
- † These authors contributed equally to this work.



Citation: Chadalavada, K.; Anbazhagan, K.; Ndour, A.; Choudhary, S.; Palmer, W.; Flynn, J.R.; Mallayee, S.; Pothu, S.; Prasad, K.V.S.V.; Varijakshapanikar, P.; et al. NIR Instruments and Prediction Methods for Rapid Access to Grain Protein Content in Multiple Cereals. *Sensors* **2022**, *22*, 3710. <https://doi.org/10.3390/s22103710>

Academic Editors: Cecilia Jimenez, Maria Luz Rodriguez-Mendez and Stella Vallejos Vargas

Received: 24 February 2022

Accepted: 29 April 2022

Published: 13 May 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Achieving global goals for sustainable nutrition, health, and wellbeing will depend on delivering enhanced diets to humankind. This will require instantaneous access to information on food-source quality at key points of agri-food systems. Although laboratory analysis and benchtop NIR spectrometers are regularly used to quantify grain quality, these do not suit all end users, for example, stakeholders in decentralized agri-food chains that are typical in emerging economies. Therefore, we explored benchtop and portable NIR instruments, and the methods that might aid these particular end uses. For this purpose, we generated NIR spectra for 328 grain samples from multiple cereals (finger millet, foxtail millet, maize, pearl millet, and sorghum) with a standard benchtop NIR spectrometer (DS2500, FOSS) and a novel portable NIR-based instrument (HL-EVT5, Hone). We explored classical deterministic methods (via winISI, FOSS), novel machine learning (ML)-driven methods (via Hone Create, Hone), and a convolutional neural network (CNN)-based method for building the calibrations to predict grain protein out of the NIR spectra. All of the tested methods enabled us to build relevant calibrations out of both types of spectra (i.e., $R^2 \geq 0.90$, $RMSE \leq 0.91$, $RPD \geq 3.08$). Generally, the calibration methods integrating the ML techniques tended to enhance the prediction capacity of the model. We also documented that the prediction of grain protein content based on the NIR spectra generated using the novel portable instrument (HL-EVT5, Hone) was highly relevant for quantitative protein predictions ($R^2 = 0.91$, $RMSE = 0.97$, $RPD = 3.48$). Thus, the presented findings lay the foundations for the expanded use of NIR spectroscopy in agricultural research, development, and trade.

Keywords: cereals; protein; near-infrared spectroscopy (NIRS); prediction methods; winISI; Hone Create; Convolution Neural Network (CNN)

1. Introduction

Near-infrared spectroscopy (NIRS) is a non-destructive method that is widely used to predict the organic compounds of grain materials based on electromagnetic wave interactions. This technology offers time- and cost-effective options to analyze grain quality

parameters [1–4]. While several companies offer the standard benchtop NIR spectrometers (such as the FOSS-DS2500 flour analyzer, [5], Bruker’s Tango FT-NIR spectrometer [6], Pertent-IM9520 [7]), within the last decade, the market has offered several options for portable NIR instruments as well [8–10]. Examples include the MicroNIR OnSite-W from VIAVI Solutions [11], the DLP NIRScan™ Nano EVM spectrophotometer from Texas Instruments’ DMD™ [12], the MEMS spectrometer from Fraunhofer [13], and the Hone Lab Red from Hone [14]. While many benchtop NIR instruments are already used for standard applications across the agri-food sector, portable instruments are not regularly used [15–17]. This is because accurate NIRS-based predictions require, among others, quality instrumentation to generate spectra, as well as robust prediction methods, and both can be problematic for some of the portable NIRS instruments [18–23].

There are many software packages enabling the prediction of material composition from the NIR spectra, which build on classical deterministic methods such as principal component analysis (PCA), partial least squares regression (PLSR), and multiple linear regression (MLR) [20,23–33]. Recent software programs offer machine learning (ML)-based methods—such as random forest, support vector machine regression, or stacked ensembles. However, more complex ML-based methods still require user configuration or custom software builds [34–41]. ML-based methods, in particular, are gaining a lot of attention, as these may offer specific advantages for applications where feature prediction is required from imperfect spectra or spectra derived from a range of materials. This is because ML-algorithms have the intrinsic capacity to identify common patterns in diverse data information, from which the required algorithms are built [38].

In the case of the cereal-based food and feed industry [25,42,43], benchtop NIRS systems with calibrations based on single-cereal species are routinely utilized—rice [26,28,30,32,36,40,44–48], sorghum [43,49,50], wheat [27,51–55], corn [31,53,56–60], and barley [53,61–63]. These typically use a large number of single-species samples from different environments representing a relevant range of target trait variability [3,52,64–69]. With rising global attention on food quality, minor cereals (such as sorghum, fonio, teff, and millets) are being explored and promoted [70–73]. These minor cereals are important sources of human and livestock diets that significantly influence their nutritional status and health [74–76], especially in the tropics [77–80]. However, for the industrial use of these cereals, rapid access to their grain components is required, and this is currently limited. At the same time, the trait variability within the minor species may be another significant bottleneck constraining the development of robust calibrations. Furthermore, even if large variable datasets were available for these minor crops, the development and maintenance of separate calibrations for each of these cereal species may prove to be time- and cost-inefficient. Therefore, a multi-cereal species calibration could be a convenient alternative. Most of the reported multi-crop calibrations focus on forage and feed analysis [81–87]. Very few reports have documented robust calibration models for grain across multiple species using the classical statistical modelling methods. A few authors have argued that ML-based prediction approaches could improve the reliability of multi-species calibrations [88].

So far, there are not many examples where NIRS-based systems have been used for the quality assessment of minor cereal grains [89–92]. In our case, we were interested in evaluating some of the novel instruments and methods. While the standard benchtop NIRS instruments are routinely used and certified for the assessment of grain quality (e.g., wheat in Australia) [93,94], the relevance of the upcoming generation of portable NIR instruments is not commonly standardized or documented [95]. Therefore, we aimed to evaluate the efficiency of several NIRS technology options (instruments and software) for the prediction of grain protein content in multiple cereal species. The emerging NIRS technology (portable instruments and ML-based prediction methods) were included in the study to test whether these could suit the specific needs of de-centralized agri-food value chains. The specific objectives of the study were to:

- (i) Compare the NIR spectra produced by benchtop (DS2500, FOSS) and portable (HL-EVT5, Hone) instruments, and their suitability for predicting grain protein in multiple cereal species;
- (ii) Assess the suitability of different model-building methods (via winISI, FOSS; Hone Create, Hone; and a customized CNN-based method) to predict protein content in multiple cereals using two types of NIR spectra;
- (iii) Ascertain the predictions made using multiple instrument–method combinations (i.e., FOSS-DS2500–winISI; FOSS-DS2500–Hone Create; FOSS-DS2500–CNN; HL-EVT5–Hone Create; and HL-EVT5–CNN) and discuss the suitability of their applications (for example, in decentralized breeding programs and markets).

2. Materials and Methods

The overall methodology used in the study is summarized in Figure 1. Briefly, 328 grain samples from five cereals were used for building NIR-based multiple cereal prediction models for estimating grain protein content. For this two NIR instruments and three calibration model building methods were explored. Also, the relevance of these instrument–method combinations were assessed using linear regression and goodness of fit parameters. Each of the steps are described in the sections below.

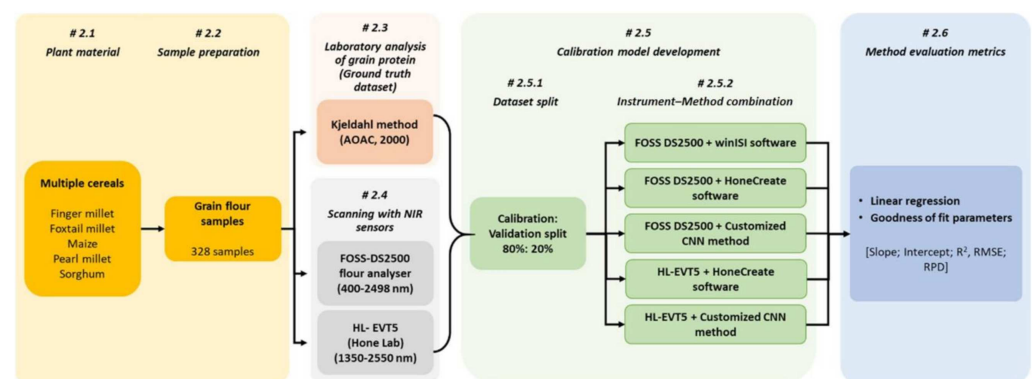


Figure 1. Graphical overview of the methodology visualizing the process used for testing the NIR instruments and methods for prediction of protein content in multiple cereal grains.

2.1. Plant Material

A total of 328 grain samples from 5 cereal species—154 genotypes of sorghum (*Sorghum bicolor* (L.) Moench), 125 genotypes of pearl millet (*Cenchrus americanus* (L.) Morrone), 20 genotypes of finger millet (*Eleusine coracana* Gaertn.), 19 genotypes of foxtail millet (*Setaria italica* (L.) P. Beauvois), and 10 genotypes of maize samples (*Zea mays* L.; details in Supplementary Table S1) were used in the study. The maize cultivars were obtained from the maize improvement program of the International Maize and Wheat Improvement Center (CIMMYT), and the remaining material were from the genebank repository of ICRISAT (Patancheru, India) [96] and the ICRISAT crop improvement programs. The subset of 154 sorghum samples included 4 races (bicolor, caudatum, durra, and guinea) originating from Burkina Faso, Cameroon, Ethiopia, India, Lesotho, Mali, Nigeria, and USA [97–100]. The subset of 125 pearl millet samples used in the study comprised 100 lines from the pearl millet inbred germplasm association panel (PMiGAP) [101] and 25 elite cultivars of Asian and African origin. Samples of 20 finger millet genotypes originating from India, Kenya, Malawi, Senegal, Uganda, and Zimbabwe, and 19 foxtail millet genotypes originating from China, India, Iran, Pakistan, Russia, and USA were used in the study [96].

2.2. Sample Collection and Preparation

The crops were raised on alfisol soil with recommended management practices [102] under irrigated conditions at the ICRISAT campus (Patancheru, India, 17.53° N, 78.27° E, 545 m.a.s.l) during the post-rainy season (October 2018–January 2019). The panicles from

physiologically mature plants were harvested and manually threshed. Grains of each genotype were pooled, cleaned, and ground to flour of <1 mm particle size, using a CM 290 Cemotech™ laboratory grinder (FOSS, Hillerød, Denmark). The flour samples were then stored in 50 mL conical polypropylene Falcon tubes at 4 °C until laboratory analysis (see Section 2.3) and scanning with NIR instruments (see Section 2.4).

2.3. Laboratory Analysis of Grain Protein Content (“Ground Truth” Dataset)

The flour samples were dried at 130 °C for 2 h in an oven and cooled to room temperature prior to chemical analysis. Standard AOAC (2000) protocols [103] were followed to estimate moisture (AOAC 925.10) and total nitrogen content (N%; Kjeldahl method, AOAC 2001.11) in each sample (i.e., for each genotype separately; see Section 2.1). The total protein content was then calculated using the generically agreed consensus for N content conversion into protein [104,105]; i.e., by multiplying the nitrogen content with a protein conversion factor of 6.25 (Equation (1)).

$$\text{Protein} = \text{N\%} \times 6.25 \quad (1)$$

All of the values were reported on a dry matter basis, i.e., weight of the component per total dry weight of the sample (% (g·100 g⁻¹)) (Table 1; Supplementary Table S1).

Table 1. Results of laboratory estimation of protein content in multiple cereal samples used in the study. The table indicates the range and average grain protein content % (g·100 g⁻¹), along with the number of samples used per species.

Species	Number of Samples	Range of Protein (% (g·100 g ⁻¹))	Average of Protein (% (g·100 g ⁻¹))
Finger millet	20	5.99–9.59	7.93
Foxtail millet	19	9.08–13.42	11.50
Maize	10	8.53–9.97	9.14
Pearl millet	125	9.69–21.51	15.78
Sorghum	154	8.68–18.38	13.09
Multiple cereals	328	5.99–21.51	13.59

2.4. Scanning Samples with Two NIR Instruments

Prior to scanning, the samples were dried at 50 °C for 16 h and cooled to room temperature. The samples were then scanned using a benchtop NIR spectrometer DS2500 flour analyzer from FOSS (FOSS-DS2500; FOSS Electric A/S, Hillerød, Denmark) [5] and Hone Lab’s portable NIR instrument HL-EVT5 (Hone Lab-Engineering Validation Test Model 5; Hone, Newcastle, NSW, Australia) [106].

Benchtop NIR Instrument: For obtaining the spectral sample signature from the FOSS-DS2500, each flour sample was transferred to the standard circular ring cup (inside diameter ~6 cm, FOSS sample cup) and scanned three times at room temperature (~26 °C). The sample was mixed before each scan. The NIR spectral absorbance, with a range of 400–2498 nm, was recorded as the logarithm of reciprocal reflectance (1/R) with 2 nm intervals, using the WinISI spectral analytical software (v4.4, InfraSoft International LLC, PA, USA).

Portable NIR instrument: To obtain the sample spectral signature from a portable HL-EVT5, the dried flour sample was spread on a glass petri plate with a minimum 5 mm thickness. The instrument was then placed on the layer of flour and scanned at a room temperature of ~26 °C. The instrument was operated via a Bluetooth-connected Hone Create mobile application (v25.2.2 Hone, Newcastle, NSW, Australia; retrieved from play.google.com) (accessed on 8 February 2022). Each sample was scanned at three different points of the sample spread on the petri plate. The mobile application was programmed to record two scans at each position, resulting in six scans per sample. NIR spectra with a range of 1350–2550 nm and a resolution of 16 nm at a wavelength of 1550 nm (NeoSpectra-

Micro optical engine, Si-Ware Systems, CA, USA), were extracted from the Hone Create platform [107].

2.5. Calibration Model Development

2.5.1. Definition of Calibration and Validation Datasets

The spectral data of 328 samples extracted from the FOSS-DS2500 and the HL-EVT5 were associated with the respective laboratory protein estimates (see Section 2.3). The spectral data from the HL-EVT5 were then split into calibration and validation datasets (80%:20%, respectively). Several methods were considered based on the literature reviews [95,108–111], but random split was finally made using the Hone Create Platform, as random selection minimizes user bias and is acceptable for large datasets with a normal or uniform distribution. In this case, Hone Create ensures that the minimum and maximum values of protein were designated to the calibration set, and that each species was represented in both the calibration and validation sets, with a ratio of 80:20. Subsequently, the calibration dataset with 262 samples (80% of the total dataset) was used to develop the calibration model and the validation dataset (20% of the total dataset), with 66 samples being used to evaluate the prediction potential of the model (details in Section 2.6). The exact same split of calibration and validation samples was made for the FOSS-DS2500 spectral data. In this way, we formed the basis for evaluating the efficiency of several instruments and the methods of building calibration models—i.e., the WinISI spectral analytical software (v4.4, InfraSoft International LLC, PA, USA), the cloud-based Hone Create software (v25.2.2 Hone, Newcastle, NSW, Australia; retrieved from play.google.com) (accessed on 8 February 2022), and the customized convolution neural network algorithm-based method (TensorFlow/Keras API) [112,113].

2.5.2. Prediction Method Development Using Established Software and a Custom-Made Pipeline: Instrument–Method Combinations

The WinISI analytical software is designed to assess FOSS instrument-generated data in the proprietary data format (.nir). Therefore, the HL-EVT5 instrument data could not be evaluated using the WinISI software. However, the Hone Create Platform enables the users to load any type of data, as long as it is in the prescribed .csv format. Similarly, the customized CNN-based method allows users to import spectral data from any instrument (<https://github.com/adamavip/nirs-protein-prediction>) (accessed on 20 March 2022). Consequently, it was feasible to treat both spectra types, FOSS-DS2500 and HL-EVT5, using the Hone Create Software, and the customized CNN-based model-building method, while the WinISI could be used only to treat FOSS-DS2500-generated spectra.

FOSS-DS2500 NIR Spectra Processed using WinISI Software:

The WinISI software (v4.4) offers several mathematical spectra pre-processing steps: standard normal variate (SNV, range tested), baseline shift, NIR trajectory derivative, and smoothing. After spectral pre-processing, calibrations can be built using several deterministic methods—principal component regression (PCR), partial least squares regression (PLSR), and modified partial least squares (MPLS)—in combination with pre-treatment methods [114–116]. Iterations between the methods can be performed manually, and the prediction potential of the models built can be tested using the validation dataset (described in Section 2.5.1). Accordingly, we performed several manual iterations between the available methods. The calibration models achieving the best metrics, i.e., the slope and intercept of linear regression, coefficient of determination (R^2), root mean squared error (RMSE), and the relative prediction deviation (RPD) for the calibration and validation datasets were then reported.

FOSS-DS2500 and HL-EVT5 NIR Spectra Processed using Hone Create Software:

The automated Hone Create software [107] applies a matrix of pre-processing options similar to the WinISI software: baseline correction, area normalization, smoothing, derivative, SNV, or combinations of these techniques. Hone Create automatically iterates

and selects the best-performing pre-processing method(s) based on the regression (PLS) or classification (C4.5) models. Once processed, a range of machine learning techniques are automatically tested and compared, including distributed random forest (DRF), generalized linear model (GLM), gradient boosting machine (GBM), extreme gradient boosting (XGBoost), and stacked ensembles. Currently, the best-performing calibration model is selected based on the root mean squared error (RMSE) and the coefficient of determination (R^2) of the calibration set. “Holdover validation” metrics (i.e., the independent validation set, described in Section 2.5.1.) are automatically processed for the user, with interactive results being displayed to allow the user to interrogate the dataset and to perform further model iterations as needed.

In this study, to compare the prediction potential of instrument–model combinations, the spectra of calibration datasets from the FOSS-DS2500 and the HL-EVT5 instruments (described in Section 2.5.1) were treated separately. The spectral data from the FOSS-DS2500 and the HL-EVT5 in Excel format (.csv) were uploaded into the Hone Create platform. The pipeline was set to automatically iterate the spectrum pre-processing methods before building the calibration model. The best-performing pre-processing method(s) were selected and subsequently used to transform the dataset prior to automatically building the calibration model, using the supervised AutoML framework of Hone Create. Once the optimal calibration model was identified for the calibration set, Hone Create automatically applied the same pre-processing method(s) to the validation set, ran the data against the model, and displayed the analogous metrics, which were then compared with other tested methods (details in Section 2.6).

FOSS-DS2500 and HL-EVT5 NIR Spectra Processed using a Customized CNN-Based Algorithm:

So far, more complex ML-algorithms such as convolutional neural networks (CNNs) are difficult to reliably automate within the software interface for regular use by non-experts. Therefore, the methodology involving convolutional neural networks (CNNs) [112] for building multivariate regression calibration models was explored separately. The CNN method was built on the publicly available open-source TensorFlow/Keras API [113]. This CNN was composed of three convolutional layers, three pooling layers, and three fully connected layers. Each convolution layer had 24, 48, and 96 filters, with kernel sizes set to 10, 15, and 25, respectively. All stride parameters were attributed to two. The network was then organized between the convolutional layer and pooling to realize the extraction and mapping of local features from the input NIR dataset. Several fully connected layers were then consecutively arranged, and the regression of targets was performed using a sigmoid function. Batch normalization was added after every convolutional layer to prevent an internal covariate shift and to speed up convergence. The ADAM50 function, a gradient descent algorithm, was set to minimize the loss function with an initial learning rate of 3×10^{-4} , which enabled the reverse adjustment of weights from the network, using a backpropagation algorithm, reducing the mean squared error of the model after each training iteration [117]. A max-pooling layer, with a kernel filter size set to two, was connected to each layer of activation function. A dropout of 0.02 was then used to deactivate 2% of the network neurons. Finally, the output of the last dropout layer was flattened to represent the high-dimensional features of the input dataset. The extracted high-dimensional features were fed into a multi-layer perceptron (MLP) to execute the final regression task. There were hidden layers in the MLP, with 512 and 128 neurons, successively. A regularization term (index = 10^{-7}) was added to every hidden layer to minimize overfitting, followed by batch normalization. The model was trained with a training batch size of 64, using Google Colaboratory (NVIDIA K80s GPU, 12.72 of RAM, and 358.27 GB of hard disk for one runtime), an open-source service provided by Google.

The spectral datasets from the FOSS-DS2500 and the HL-EVT5 were pre-treated using the Savitzky–Golay smoothing filter [114], with a window size of 15 and a polynomial order set to 2, as was performed in similar studies [88,118]. The transformed data, before being fed to the CNN, were then normalized using min/max normalization of the first derivatives so that values ranged between 0 and 1 [37]. Subsequently, the CNN training

structure was constructed to predict protein quantity (% (g·100 g⁻¹)) from the spectral data. For this, the calibration set model was trained using a five-fold cross-validation approach to determine the optimal number of epochs and the effectiveness of certain hyperparameters, such as activation functions, neuron counts, and layer counts. The model (with the same architecture as that of the cross-validation) was retrained with the selected hyperparameters on the entire training dataset and tested with the validation dataset (described in Section 2.5.1). Iterations between the pre-treatment and the normalization methods were performed, and the best-performing model was selected based on the common metrics of both the calibration and validation datasets (details in Section 2.6).

2.6. Prediction Method Evaluation

To compare the predictive potentials of the different instrument–method combinations, we used the statistical metrics describing the linear interdependency between the ground-truth (grain protein content estimated using the laboratory method, see Section 2.3) and the best method for predicting protein content from the NIR spectra (separately for the calibration and validation sets, see Section 2.5.1). To assess the developed instrument–method combinations (Section 2.5), five parameters were used—the slope and intercept of the linear regression; coefficient of determination (R^2 , Equation (2)); root mean squared error (RMSE, Equation (3)); and the ratio of prediction to deviation (RPD, Equation (4)) [119–121],

$$R^2 = 1 - \frac{\sum(\hat{y}_i - y_i)^2}{\sum(\hat{y}_i - \bar{y})^2} \quad (2)$$

$$RMSE = \sqrt{\frac{\sum(\hat{y}_i - y_i)^2}{n}} \quad (3)$$

where n is the number of samples; y_i is the ground-truth (see Section 2.3) value of sample i ; \hat{y}_i is the model-predicted value of sample i ; \bar{y} is the mean of the ground-truth values; and SD is the standard deviation of the ground-truth values.

$$RPD = \frac{SD}{RMSE} \quad (4)$$

We adopted the previously-reported classification based on the RPD values [119], wherein an RPD value <1.5 indicates that the calibration is not reliable; a value between 1.5 and 2.0 indicates the capacity of a model to distinguish between high and low values; a value between 2.0 and 2.5 signifies the model’s capacity to “approximate” quantitative prediction; a value between 2.5 and 3.0 suggests “good” quantitative prediction; and a value > 3.0 indicates “excellent” quantitative prediction.

3. Results

3.1. Diversity of Grain Protein Content in Five Cereal Species

The laboratory analysis of protein content obtained from 328 grain samples across five cereal species ranged from 5.99% to 21.51% (Table 1, Supplementary Table S1). The range of protein content in multiple cereals was considerably larger compared to the protein content variability within any of the individual species tested (Table 1, Figure 2). Among the five cereals tested, the mean protein content in pearl millet grains was the highest (21.51%), while the mean protein content was the lowest in finger millet grains (5.99%; Table 1, Figure 2).

Based on the protein content, the Hone Create software randomly splits the samples of each species into calibration and validation datasets (80%:20%, respectively) (Figure 3). The range, average, standard deviation, and distribution of protein content across the calibration and validation datasets were comparable (Supplementary Tables S2 and S3).

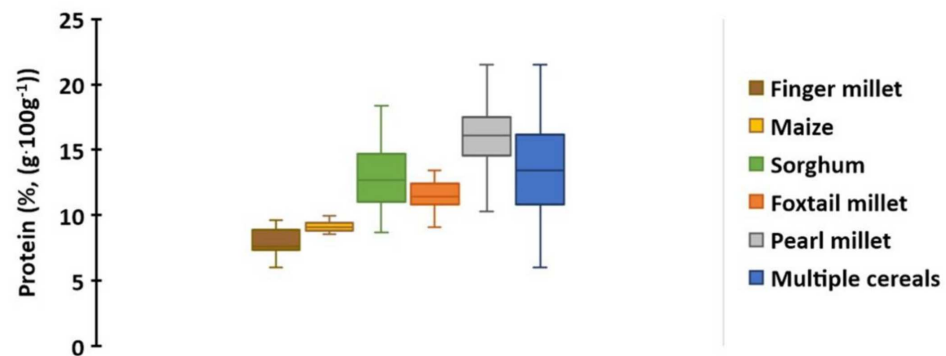


Figure 2. Box plots depicting variation and distribution of protein content (% (g·100 g⁻¹)) in the grains of five cereals, as estimated through laboratory analyses. Legend: Each box represents one crop species; different crops are distinguished by color (finger millet = brown; maize = yellow; sorghum = green; foxtail millet = orange; pearl millet = grey; and the entire set of 328 multiple cereals = blue); solid line within the box (–) represents the mean of each crop.

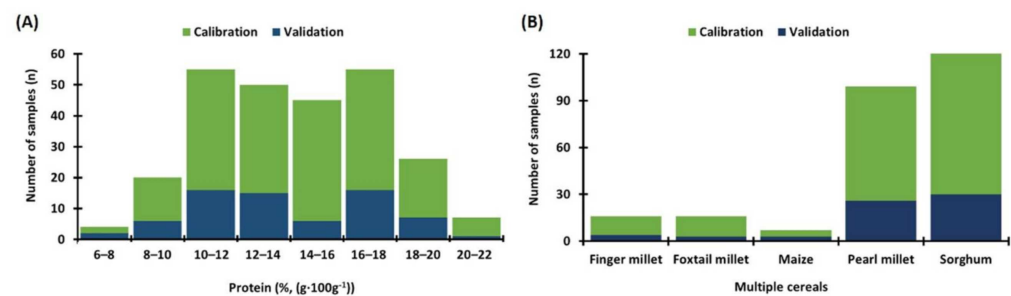


Figure 3. Histograms depicting the distribution of (A) the protein content (% (g·100 g⁻¹)) in samples, and (B) the number of samples used within each of the crop species belonging to the calibration (80%) and validation (20%) datasets.

3.2. NIR Spectrum Obtained from the Benchtop FOSS-DS2500 and the Portable HL-EVT5

The NIR absorbance spectra of 328 samples were recorded using two NIR instruments—the benchtop FOSS-DS2500 (400–2498 nm) and the portable HL-EVT5 (1350–2550 nm). The spectrum profile generated from each instrument was very similar within the range of 1352–2498 nm (Figure 4). This indicated that the technology used to generate the NIR spectral signatures captured the biochemical signature of the biological samples very similarly.

Overall, the FOSS-DS2500 signal was dominated by 13 groups of prominent peaks (Figure 5A) and 5 peaks for the HL-EVT5 (Figure 5B). The protein content is known to be linked to several spectral bands: (i) a range of 950–1050 nm as the N–H second stretch overtone, (ii) around 1500 nm as the N–H stretching first overtone, and (iii) the N–H bend second overtone, and the C = O stretch–N–H in-plane bending–C–N stretch combination bands are further associated with a range of 2150–2200 nm [1,2,17]. Therefore, both of the instruments should be sufficient to capture some, if not all, of these critical NIR spectral ranges to predict protein content.

3.3. NIR Spectrum Generated Using the FOSS-DS2500 and Processed via WinISI Software

The WinISI software enabled manual iterations between the spectra pre-processing steps (SNV, detrend, NIR trajectory derivative, and smoothing) and several deterministic calibration model algorithms (PLS, MPLS, and PCR). For our dataset, the best calibration method that attained the highest accuracy metrics was obtained using spectra pre-processed using combinations of scatter-correction algorithms, SNV&D with a mathematical pre-treatment setting of “2,4,4,1” (i.e., 2 = second derivative treatment; 4 = a gap of four wavelength points over which the derivative was calculated; 4 = first smoothing using the Savitzky–Golay algorithm at four data points; and 1 = no secondary smoothing),

combined with the modified partial least squares (MPLS) regression. This calibration model achieved RMSE values of 0.91 and an R^2 of 0.90. The validation dataset possessed RMSE values of 1.09 and an R^2 of 0.86 (Table 2; Figure 6). The RPD values for the calibration and validation datasets were 3.56 and 3.08, respectively (Table 2).

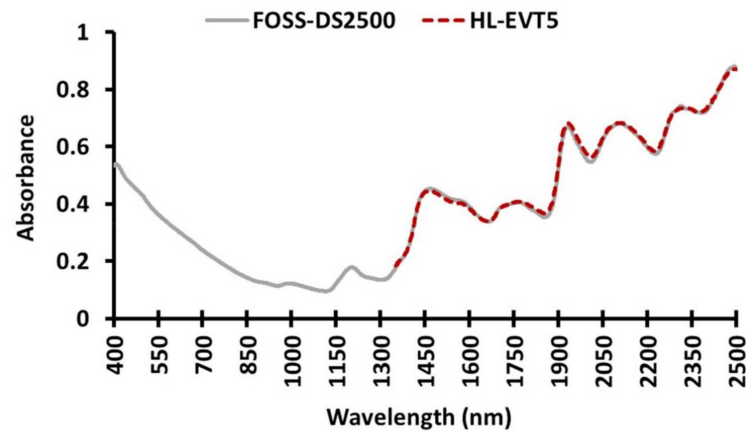


Figure 4. Mean of the near-infrared (NIR) spectra of all grain samples extracted from the benchtop FOSS-DS2500 (400–2498 nm; solid line (–) in grey colour) and the portable HL-EVT5 (1350–2550 nm; dashed line (–) in red colour) instruments.

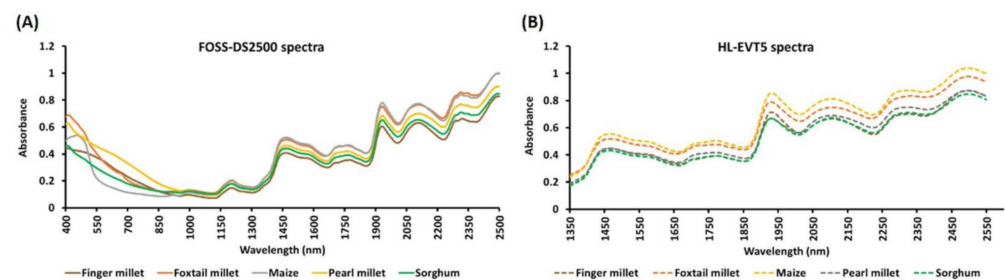


Figure 5. Means of the near-infrared (NIR) spectra of the grain samples of five cereal species produced using (A) FOSS-DS2500, 400–2498 nm; solid line (–), and (B) HL-EVT5, 1350–2550 nm; dashed line (–) instruments. Different crops are distinguished by color (Legend: finger millet = brown; foxtail millet = orange; maize = yellow; pearl millet = grey; sorghum = green).

Table 2. Comparative metrics of NIR spectroscopy calibration (80%) and validation (20%) models developed using combinations of two different instruments (FOSS-DS2500 and HL-EVT5) and three model-building methods (WinISI software, Hone Create software, CNN-based customized pipeline) for protein content estimation in grains of multiple cereal species. Legend: R^2 = coefficient of determination; RMSE = Root Mean Squared Errors, RPD = ratio of prediction to deviation, CNN = convolutional neural networks.

Instrument	Method	Set	Slope	Intercept	R^2	RMSE	RPD
FOSS-DS2500	WinISI software	Calibration	0.87	1.74	0.90	0.91	3.56
		Validation	0.82	2.38	0.86	1.09	3.08
	Hone Create software	Calibration	0.95	0.64	0.96	0.66	4.93
		Validation	0.89	1.44	0.90	1.00	3.38
	CNN-based customized pipeline	Calibration	0.98	0.29	0.99	0.33	9.85
HL-EVT5	Hone Create software	Validation	0.88	1.61	0.89	1.03	3.26
		Calibration	0.97	0.43	0.98	0.42	7.79
		Validation	0.90	1.35	0.91	0.97	3.48
	CNN- based customized pipeline	Calibration	0.98	0.28	0.98	0.46	7.00
		Validation	0.87	1.70	0.87	1.10	3.06

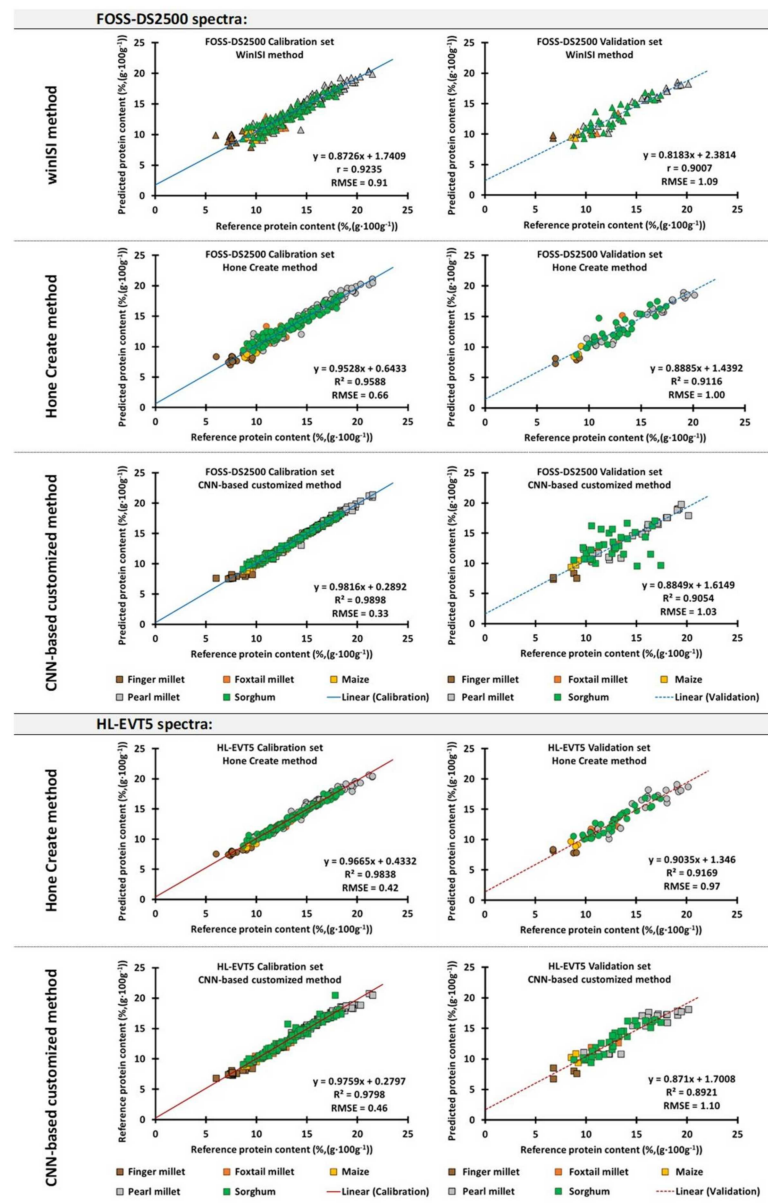


Figure 6. Matrix of scatter plots showing protein predicted for the calibration and validation datasets of FOSS-DS2500 and HL-EVT5 via methods available in (I) WinISI software, (II) Hone Create software, and (III) CNN-based customized method. Detailed metrics for comparison with other methods are shown in Table 2.

3.4. NIR Spectrum Generated Using the FOSS-DS2500 and HL-EVT5, Processed via Hone Create Software

The Hone Create software is able to perform several combinations of pre-processing methods specific for the generated data and iterate these with a range of deterministic and ML-based calibration methods. The pipeline returns a “Top 10” calibration model leaderboard, with the best-performing calibrations (based on the R^2 and RMSE of the calibration set).

In the case of the longer spectrum signatures generated using the benchtop FOSS-DS2500 instrument (Figure 6), Hone Create’s best model, achieved with the spectrum transformed using the first-order derivative (pre-processing) combined with the stacked ensemble method, determined the protein content with RMSE values of 0.66 and 1.00, and R^2 values of 0.96 and 0.90 for the calibration and validation datasets, respectively. This

model had an RPD value of 3.38 for the validation and 4.93 for the calibration dataset (Table 2; Supplementary Table S4).

For the portable NIR-instrument, HL-EVT5, we observed that the best prediction method was achieved with a spectrum that was preprocessed via area normalization (wherein the magnitude of each value in the spectra is adjusted so that the sum of every absolute magnitude equals 1), followed by spectrum merging steps using smoothing (a Savitsky–Golay filter with a period of 5 and a polynomial order of 2) and a first-order derivative combined with stacked ensemble models (Figure 6). The model showed an R^2 of 0.98, an RMSE of 0.42, and a RPD of 7.79 for the calibration dataset. The corresponding metrics for the validation set were an R^2 of 0.91, RMSE of 0.97, and an RPD of 3.48 (Table 2; Supplementary Table S4).

3.5. NIR Spectrum Generated Using FOSS-DS2500 and HL-EVT5, Processed via CNN-Based Customized Pipeline

The CNN models were also experimented with, for predicting the protein content in cereal grains. The model was built using spectral data preprocessed using the Savitzky–Golay filter, followed by min/max normalization of the first derivatives and a customized deep learning CNN algorithm for model building. For the FOSS-DS2500-generated spectra, the CNN model achieved an R^2 of 0.99 and an RMSE of 0.33 in the calibration set, and an R^2 of 0.89 and an RMSE of 1.03 in the validation set (Figure 6, Table 2). Consequently, for the HL-EVT5 instrument samples, an RMSE of 0.46 and 1.10 and an R^2 of 0.98 and 0.87 were obtained for the calibration and validation datasets, respectively (Figure 6, Table 2). The RPD values for the FOSS-DS2500 and HL-EVT5 instrument-derived validation datasets were 3.26 and 3.06, respectively (Table 2).

3.6. Instrument–Method Combination Comparisons for Protein Content Predictions in Cereal Grains

Two instruments (FOSS-DS2500 and HL-EVT5) were used to generate the NIR spectra (Figure 4) of the ground grain samples, and three analytical methods (WinISI software, Hone Create software, and the CNN-based customized algorithm) were used to build the prediction model for grain protein content (%), ($\text{g} \cdot 100 \text{g}^{-1}$). Five metrics were used to assess the performance of the different methods (slope and intercept of the linear regression, R^2 , RMSE, and RPD; Section 2.6).

Overall, as shown in Table 2, the NIR spectral signals generated from both instruments (FOSS-DS2500 and HL-EVT5) yielded reliable models (the lowest achieved $R^2 \geq 0.86$ and highest RMSE ≤ 1.10 in the validation set). The validation dataset RPD of all of the instrument–method combinations for estimating protein were greater than 3.06 (Table 2). This suggests that all of the generated methods were well suited for providing quantitative protein estimates.

Nevertheless, the prediction methods for the NIR spectra from FOSS-DS2500 using the ML methods achieved a notably higher RPD (≥ 3.26), a marginally higher R^2 (≥ 0.89), and a slightly lower RMSE (≤ 1.03), compared to the deterministic models created through the WinISI software (RPD = 3.08, $R^2 = 0.86$, and RMSE = 1.09) (Table 2). The best-performing instrument–method combination, based on the validation set metrics, was achieved with the HL-EVT5 and Hone Create software.

4. Discussion

4.1. Importance of NIR Spectroscopy for Rapid Cereal Grain Quality Assessment

The cereals investigated within this study significantly influence the food and nutritional security of farming communities across the tropics [77–79,122]. The protein content in these cereals is one of the key parameters determining nutritional grain values in human diets and its suitability for food industries [15–18,25]. Currently, NIRS-based methods are used for the rapid prediction of organic grain components [15–18,40–48,56–59,63–66,89,90,92–94]. However, there are not many examples for minor cereal grains (such as millets in the presented study [89,90,123]). We argue that the rapid assessment of minor cereal qualities could be

an entry point for their integration into mainstream food-value chains. Therefore, in this work, we assessed the suitability of some of the standard and novel NIR instruments, and of some of the standard and novel model-building methods for the estimation of protein content in minor cereals.

4.2. Expected Data Properties as Prerequisites to Building Reliable Prediction Models

Important considerations for building reliable prediction models from the NIR spectra have been comprehensively summarized in [24,110,120,121,123]. In our study, we focused on the following aspects: (i) the number of samples, range of variation and the distribution of the protein content values; (ii) the properties of the spectra generated using two sensors (FOSS-DS2500 and HL-EVT5); and (iv) the methods for building the algorithms. The number, range, and distribution of target trait variability (i.e., protein) are critical prerequisites for the development of reliable calibration models from NIR-based signals [108,120]. Generally, a higher number of precisely generated ground-truth data points increases the probability of building reliable calibrations. While deterministic methods require ~100 ground-truth points, larger sample numbers (>300) are generally required as a basis for ML-based algorithms and industrial calibrations [39,108,124,125]. In our study, we used 328 ground-truth samples, which is an adequate sample size that suits the purpose of the presented study. Additionally, the inclusion of five cereal species extended the range of protein content variability compared to the range found in any of the individual species. Also, both of the instruments tested in this study (FOSS-DS2500, with a spectral range of 400–2498 nm, and HL-EVT5, with a spectral range of 1350–2550 nm) encompassed the spectrum ranges relevant to protein determination in grain samples (Section 3.2) [2,26,28,48,54,65,69,84,91], and sufficiently justified the next step, i.e., building the algorithms to infer the protein content from these two types of spectra (Sections 2.5 and 4.3).

4.3. Prediction Methods for the Estimation of Protein Content in Multiple Cereal Grains and Their Accuracy Metrics

For the predictions of organic grain composition from NIR spectral reflectance, deterministic methods have been widely used (e.g., MLR, PCR, and PLS regression [15–20,25]). These methods were mostly specific to a single species [26–31,44–50,56–63]. The adoption of ML methods in NIR spectroscopy research has quite recently begun, together with the increasing availability of computing power and efficient learning algorithms [34,37,118]. Therefore, in our work, we wanted to test whether the ML algorithms would provide any advantage, in terms of accuracy, to the prediction methods. For this, we tested the portable NIR instrument (HL-EVT5) and the benchtop NIR instrument (FOSS-DS2500) in combination with two software (standard FOSS-made WinISI and novel Hone-made Hone Create), and one externally built CNN-based algorithm. The winISI software leaves the user to manually test the combinations of several preprocessing methods, which proved to be time consuming. In comparison, the Hone Create software enables the automatic evaluation of a similar range of pre-processing methods and models in a relatively shorter amount of time. The one potential constraint of the current Hone Create pipeline is that it selects the “best” models based on the metrics of the calibration dataset. This process might prioritize models that are more specific for the presented dataset, with less generic-prediction capacity (i.e., “over-fitting models”; one sign of overfitting could be where the calibration model metrics are vastly higher compared to the metrics of the validation dataset [88]). This was also one of the reasons for why we tested the alternative process, i.e., the custom-designed CNN pipeline, where we integrated the element where the goodness of the model was evaluated based on the metrics of the validation set. Another reason for building a separate pipeline was that, at this moment, more complex ML-methods such as CNN are difficult to automate. The code is now available on the GitHub platform (<https://github.com/adamavip/nirs-protein-prediction>, accessed on 23 February 2022), and its particular elements can be utilized to enhance existing software products and to develop other pipelines.

For longer NIR spectral signatures generated using FOSS-DS2500, we compared all three model-building methods (WinISI software, Hone Create software, and custom-made algorithms involving CNN). In this case, it was notable that the prediction methods integrating ML-based models (Section 3.4), particularly the stacked ensemble model via Hone Create ($R^2 = 0.90$, RMSE = 1.0, RPD = 3.38) and the custom-designed CNN ($R^2 = 0.89$, RMSE = 1.03, RPD = 3.26), achieved better comparative metrics than the deterministic method through WinISI ($R^2 = 0.86$, RMSE = 1.09, RPD = 3.08). Overall, all five of the presented calibration models developed for protein assessment can be used for the high-quality quantitative estimation of protein over a range of cereal grains. In addition, the study suggests that the robustness of these calibrations can be further improved by including more diverse samples to further widen the range of trait and spectral variability. Similar studies have been carried out to evaluate the grain protein contents of individual major cereals [26,28,48,54,65]. These studies achieved the accuracy metrics of prediction methods (typically, $R^2 \geq 0.86$; RPD > 3.0) that were comparable to all of the prediction methods presented in this study.

5. Conclusions

The integration of minor cereals in mainstream diets will be an important step towards the improvement of human nutrition, as these cereals generally have higher nutritional values compared to the major cereals such as wheat, rice, or maize. We argue that rapid assessment of the minor cereal qualities could be an entry point for their integration into mainstream food-value chains. This will probably require the transition of standard NIR spectroscopic instrumentation for grain quality analysis from the benchtop to the portable form. Such a transition appears to be critical for its effective utilization in decentralized systems where these minor cereals are typically produced. The extended utilization of minor cereal grains by food industries might, in turn, become an important means for the improvement of human nutrition, particularly in the tropics.

The motivation of this study was to assess whether emerging technological approaches (portable NIR instruments and ML-based methods) would enable the accurate assessment of grain composition (protein content) in multiple minor cereal staples that are typical of such agri-food systems. We demonstrated that the NIR spectra generated from a novel portable NIR instrument (HL-EVT5) were sufficient for the reliable quantification of grain protein content in multiple cereal species. The results also show that the integration of ML-based algorithms in modeling processes enhances model accuracy compared to classical deterministic methods. Finally, we highlighted that some of these advanced data-modeling methods are available for non-experts through new software packages (e.g., Hone Create software). We argue that these novel technologies, which are becoming more accessible across all markets, have the power to streamline the production and trade of nutrition-dense food sources (such as millets) into human diets.

Supplementary Materials: The following supporting information can be downloaded at <https://www.mdpi.com/article/10.3390/s22103710/s1>. Table S1: List of 328 grain samples used in the study along with the crop species, genotype ID, and protein content (% (g·100 g⁻¹)), obtained from laboratory analysis. Table S2: Descriptive statistics presenting the variability and range of protein content in the calibration and validation sets used in the study. Legend: SD = standard deviation; CV% = coefficient of variation. Table S3: Descriptive statistics (minimum (min), maximum (max), average (avg), standard deviation (SD), and standard error (SE) of the protein content (% (g·100 g⁻¹)) in each of the cereal species of the calibration and validation sets used in the study. Table S4: Comparative metrics of NIR spectroscopy calibration (80%) and validation (20%) models developed using combinations of two different instruments (FOSS-DS2500 and HL-EVT5) and Hone Create software for protein content estimation in grains of multiple cereal species. Legend: R^2 = coefficient of determination; RMSE = Root Mean Squared Errors, RPD = ratio of prediction to deviation; best prediction model for each of the instruments is highlighted in bold. Video S1: Operating HoneTest app using HL-EVT5: <https://fernlab.herokuapp.com/evt5.html> (accessed on 20 March 2022).

Author Contributions: Conceptualization, J.K., S.C. and K.A.; Methodology, K.C., S.C., J.K., K.A. and A.N.; Software, A.N., J.R.F., W.P. and K.V.S.V.P.; Validation, K.A., S.C., A.N., J.R.F. and P.V.; formal analysis, K.C., K.A., A.N., S.M. and S.P.; Investigation, K.A., S.C., A.N., J.K. and J.R.F.; Resources, J.K. and P.V.; Data curation, K.C., K.A., S.M. and S.P.; Writing—Original draft preparation, K.C., K.A. and A.N.; Writing—Review and editing, K.A., J.K., C.S.J., W.P., J.R.F. and K.V.S.V.P.; Visualization, K.A.; Supervision, J.K.; Project administration, J.K. and P.V.; Funding acquisition, J.K. and S.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the CGIAR Research Program grant for Grain Legumes and Dryland Cereals–ICRISAT (GLDC-ICRISAT; 2018–2022); CGIAR’s Crop to End Hunger initiative–ICRISAT (a multi-funder initiative led by USAID and including the Gates Foundation; DFID, UK; GiZ, Germany; and ACIAR, Australia; CtEH-ICRISAT; 2019–2021); and the Global Challenge Research Fund (GCRF)/Biotechnology and Biological Sciences Research Council (BBSRC)-funded project, Transforming India’s Green Revolution by Research and Empowerment for Sustainable food Supplies (TIGR²ESS; BB/P027970/1; 2018–22).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data related to the laboratory analysis presented in this study is openly available in the Supplementary Materials section of the manuscript. The source code of the CNN-based customized pipeline for estimating protein content in multiple grain cereals using NIRS and machine learning has been published in the Github repository <https://github.com/adamavip/nirs-protein-prediction> (accessed on 9 April 2022). The spectral data and the model developed using the software winISI and Hone Create presented in this study are not publicly available due to their proprietary nature.

Acknowledgments: The authors would like to thank B.D. Ranjitha Kumari and T. Senthil Kumar from Bharathidasan University, for supporting the PhD student Keerthi Chadalavada. The authors thank the team at ICRISAT and CIMMYT for sharing the genetic material with us, Intertek AgriTech and International Livestock Research Institute (ILRI) for the laboratory analyses, and the Hone team for their expert input. Special acknowledgment goes to Felicity Fraser, Sivasakthi Kaliamoorthy, Rekha Baddam, and Srikanth Mallayee for their support with experimentation; Premalatha Teegalnagaram and Mallesh Rahini for their support in laboratory activities; and William Nelson for his inputs on improving the manuscript.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Agelet, L.E.; Hurburgh, C.R., Jr. A tutorial on near infrared spectroscopy and its calibration. *Crit. Rev. Anal. Chem.* **2010**, *40*, 246–260. [CrossRef]
2. Workman, J.; Weyer, L. *Practical Guide and Spectral Atlas for Interpretive Near-Infrared Spectroscopy*, 2nd ed.; CRC Press: Boca Raton, FL, USA, 2012. [CrossRef]
3. Villamuelas, M.; Serrano, E.; Espunyes, J.; Fernández, N.; López-Olvera, J.R.; Garel, M.; Santos, J.; Parra-Aguado, M.Á.; Ramanzin, M.; Fernández-Aguilar, X.; et al. Predicting herbivore faecal nitrogen using a multispecies near-infrared reflectance spectroscopy calibration. *PLoS ONE* **2017**, *12*, e0176635. [CrossRef] [PubMed]
4. Rukundo, I.R.; Danao, M.G.C.; Mitchell, R.B.; Masterson, S.D.; Weller, C.L. Comparing the use of portable and benchtop NIR spectrometers in predicting nutritional value of forage. *Appl. Eng. Agric.* **2021**, *37*, 171–181. [CrossRef]
5. FOSS-DS2500 Flour Analyzer from FOSS. Available online: <https://www.dksh.com/global-en/products/ins/foss-flour-analyzer-nirs-ds2500> (accessed on 7 January 2021).
6. Bruker-Tango FT-NIR Spectrometer from Bruker. Available online: <https://www.bruker.com/en/products-and-solutions/infrared-and-raman/ft-nir-spectrometers/tango-ft-nir-spectrometer.html> (accessed on 7 January 2021).
7. Perten-IM9520 Flour Analyzer from PerkinElmer. Available online: <https://www.calibrecontrol.com/main-product-list/perten-im9520-flour-analyser> (accessed on 7 January 2021).
8. Sorak, D.; Herberholz, L.; Iwascek, S.; Altinpinar, S.; Pfeifer, F.; Siesler, H.W. New developments and applications of portable raman, mid-infrared, and near-infrared spectrometers. *Appl. Spectrosc. Rev.* **2012**, *47*, 83–115. [CrossRef]
9. Sarikaş, A.; Başar, M.D. An electronic portable device design to spectroscopically assess fruit quality. *Turk. J. Electr. Eng. Comput. Sci.* **2017**, *25*, 4063–4076. [CrossRef]

10. Crocombe, R.A. Portable Spectroscopy. *Appl. Spectrosc.* **2018**, *72*, 1701–1751. [CrossRef]
11. MicroNIR OnSite-W from VIAVI Solutions. Available online: <https://www.viavisolutions.com/en-us/osp/products/micronir-onsite-w> (accessed on 7 January 2021).
12. DLP NIRScanTM Nano EVM Spectrometer from Texas Instruments. Available online: <https://www.ti.com/tool/DLPNIRNANOEVMS> (accessed on 7 January 2021).
13. MEMS Spectrometer from Fraunhofer. Available online: <https://www.ipms.fraunhofer.de/en/Components-and-Systems/Components-and-Systems-Sensors/Optical-Sensors/MEMS-based-spectroscopy.html> (accessed on 7 January 2021).
14. Hone Lab Red from Hone. Available online: <https://www.honeag.com/hone-lab> (accessed on 7 January 2021).
15. Osborne, B.G. Near infrared spectroscopy in food analysis. In *Encyclopedia of Analytical Chemistry: Applications, Theory and Instrumentation*; John Wiley & Sons Ltd: New York, NY, USA, 2006. [CrossRef]
16. Singh, C.B.; Paliwal, J.; Jayas, D.S.; White, N.D. Near-infrared spectroscopy: Applications in the grain industry. In Proceedings of the CSBE/SCGAB Annual Conference, Edmonton, Alberta, 16–19 July 2006.
17. dos Santos, C.A.T.; Lopo, M.; Páscoa, R.N.M.J.; Lopes, J.A. A Review on the Applications of portable near-infrared spectrometers in the agro-food industry. *Appl. Spectrosc.* **2013**, *67*, 215–1233. [CrossRef]
18. Williams, P.C. Application of near infrared reflectance spectroscopy to analysis of cereal grains and oilseeds. *Cereal Chem.* **1975**, *52*, 561–576.
19. Norris, K.H.; Barnes, R.F.; Moore, J.E.; Shenk, J.S. Predicting forage quality by infrared reflectance spectroscopy. *J. Anim. Sci.* **1976**, *43*, 889–897. [CrossRef]
20. Estienne, F.; Pasti, L.; Centner, V.; Walczak, B.; Despagne, F.; Rimbaud, D.J.; De Noord, O.E.; Massart, D.L. A comparison of multivariate calibration techniques applied to experimental NIR data sets: Part II. Predictive ability under extrapolation conditions. *Chemometr. Intell. Lab. Syst.* **2001**, *58*, 195–211. [CrossRef]
21. Esbensen, K.H.; Julius, L.P. Representative sampling, data quality, validation—A necessary trinity in chemometrics. In *Comprehensive Chemometrics*; Brown, S., Tauler, R., Walczak, R., Eds.; Elsevier: Oxford, UK, 2009; Volume 4, pp. 1–20.
22. Agelet, L.E.; Hurburgh, C.R., Jr. Limitations and current applications of near infrared spectroscopy for single seed analysis. *Talanta* **2014**, *121*, 288–299. [CrossRef] [PubMed]
23. Chang, H.; Zhu, L.; Lou, X.; Meng, X.; Guo, Y.; Wang, Z. A new local modelling approach based on predicted errors for near-infrared spectral analysis. *J. Anal. Methods Chem.* **2016**, *2016*, 5416506. [CrossRef] [PubMed]
24. Cheewapramong, P. Use of Near-Infrared Spectroscopy for Qualitative and Quantitative Analyses of Grains and Cereal Products. Ph.D. Thesis, University of Nebraska-Lincoln, Lincoln, NE, USA, 2007.
25. Downey, G. NIR and chemometrics in the service of the food industry. *NIR News* **2007**, *18*, 10–11. [CrossRef]
26. Chen, J.Y.; Miao, Y.; Sato, S.; Zhang, H. Near infrared spectroscopy for determination of the protein composition of rice flour. *Food Sci. Technol. Res.* **2008**, *14*, 132–138. [CrossRef]
27. Kahriman, F.; Egesel, C.Ö. Development of a calibration model to estimate quality traits in wheat flour using NIR (Near Infrared Reflectance) spectroscopy. *Res. J. Agric. Sci.* **2011**, *43*, 392–400.
28. Bagchi, T.B.; Sharma, S.; Chattopadhyay, K. Development of NIRS models to predict protein and amylose content of brown rice and proximate compositions of rice bran. *Food Chem.* **2016**, *191*, 21–27. [CrossRef]
29. Lyu, N.; Chen, J.; Pan, T.; Yao, L.; Han, Y.; Yu, J. Near-infrared spectroscopy combined with equidistant combination partial least squares applied to multi-index analysis of corn. *Infrared Phys. Techn.* **2016**, *76*, 648–654. [CrossRef]
30. Sampaio, P.S.; Soares, A.; Castanho, A.; Almeida, A.S.; Oliveira, J.; Brites, C. Optimization of rice amylose determination by NIR-spectroscopy using PLS chemometrics algorithms. *Food Chem.* **2018**, *242*, 196–204. [CrossRef]
31. Tomas, E.; Bayram, I. Establishing near infrared spectroscopy (NIR) calibration for starch analysis in corn grain. *Kocatepe Vet. J.* **2018**, *12*, 7–14.
32. Chen, J.; Li, M.; Pan, T.; Pang, L.; Yao, L.; Zhang, J. Rapid and non-destructive analysis for the identification of multi-grain rice seeds with near-infrared spectroscopy. *Spectrochim. Acta A Mol. Biomol. Spectrosc.* **2019**, *219*, 179–185. [CrossRef]
33. Kahriman, F.; Liland, K.H. SelectWave: A graphical user interface for wavelength selection and spectral data analysis. *Chemom. Intell. Lab. Syst.* **2021**, *212*, 104275. [CrossRef]
34. Lee, S.; Choi, H.; Cha, K.; Kim, M.-K.; Kim, J.-S.; Youn, C.H.; Lee, S.-H.; Chung, H. Random Forest as a non-parametric algorithm for near-infrared (NIR) spectroscopic discrimination for geographical origin of agricultural samples. *Bull. Korean Chem. Soc.* **2012**, *33*, 4267–4270. [CrossRef]
35. Kong, W.; Zhang, C.; Liu, F.; Nie, P.; He, Y. Rice seed cultivar identification using near-infrared hyperspectral imaging and multivariate data analysis. *Sensors* **2013**, *13*, 8916–8927. [CrossRef] [PubMed]
36. Chen, H.; Tan, C.; Lin, Z. Authenticity detection of black rice by near-infrared spectroscopy and support vector data description. *Int. J. Anal. Chem.* **2018**, *2018*, 8032831. [CrossRef] [PubMed]
37. Cui, C.; Fearn, T. Modern practical convolutional neural networks for multivariate regression: Applications to NIR calibration. *Chemometr. Intell. Lab. Syst.* **2018**, *182*, 9–20. [CrossRef]
38. Das, B.; Nair, B.; Reddy, V.K.; Venkatesh, P. Evaluation of multiple linear, neural network and penalised regression models for prediction of rice yield based on weather parameters for west coast of India. *Int. J. Biometeorol.* **2018**, *62*, 1809–1822. [CrossRef]
39. Le, T.H.; Chen, H.; Babar, M.A. Deep learning for source code modeling and generation: Models, applications, and challenges. *ACM Comput. Surv.* **2020**, *53*, 62–100. [CrossRef]

40. Sampaio, P.S.; Castanho, A.; Almeida, A.S.; Oliveira, J.; Brites, C. Identification of rice flour types with near-infrared spectroscopy associated with PLS-DA and SVM methods. *Eur. Food Res. Technol.* **2020**, *246*, 527–537. [[CrossRef](#)]
41. Kabir, M.H.; Guindo, M.L.; Chen, R.; Liu, F. Geographic origin discrimination of millet using vis-NIR spectroscopy combined with machine learning techniques. *Foods* **2021**, *11*, 2767. [[CrossRef](#)]
42. Kahriman, F.; Egesel, C.Ö. Using near infrared (NIR) spectroscopy in the analysis of cereal products: The example of maize. In *Recent Researches in Science and Landscape Management*; Efe, R., Zencirkiran, M., Curebal, İ., Eds.; Cambridge Scholars Publishing: Newcastle, NSW, Australia, 2018; pp. 507–521.
43. Ejaz, I.; He, S.; Li, W.; Hu, N.; Tang, C.; Li, S.; Li, M.; Diallo, B.; Xie, G.; Yu, K. Sorghum grains grading for food, feed, and fuel using NIR spectroscopy. *Front. Plant Sci.* **2021**, *12*, 720022. [[CrossRef](#)]
44. Osborne, B.G.; Mertens, B.; Thompson, M.; Fearn, T. The authentication of Basmati rice using near infrared spectroscopy. *J. Near Infrared Spectrosc.* **1993**, *1*, 77–83. [[CrossRef](#)]
45. Wang, H.L.; Wan, X.Y.; Bi, J.C.; Wang, J.K.; Jiang, L.; Chen, L.M.; Zhai, H.Q.; Wan, J.M. Quantitative analysis of fat content in rice by near-infrared spectroscopy technique. *Cereal Chem.* **2006**, *83*, 402–406. [[CrossRef](#)]
46. Barnaby, J.Y.; Huggins, T.D.; Lee, H.; McClung, A.M.; Pinson, S.R.M.; Oh, M.; Bauchan, G.R.; Tarpley, L.; Lee, K.J.; Kim, M.S.; et al. Vis/NIR hyperspectral imaging production environment, and physicochemical grain properties in rice. *Sci. Rep.* **2020**, *10*, 1–13. [[CrossRef](#)] [[PubMed](#)]
47. Burestan, F.N.; Sayyah, A.H.; Taghinezhad, E. Prediction of some quality properties of rice and its flour by near-infrared spectroscopy (NIRS) analysis. *Food Sci. Nutr.* **2020**, *9*, 1099–1105. [[CrossRef](#)]
48. Fazeli, N.; Amir, B.; Afkari, H.; Mahdi, S. Prediction of amylose content, protein content, breakdown, and setback viscosity of Kadus rice and its flour by near-infrared spectroscopy (NIRS) analysis. *J. Food Process. Preserv.* **2020**, *45*, e15069.
49. De Alencar Figueiredo, L.F.; Davrieux, F.; Fliedel, G.; Rami, J.F.; Chantereau, J.; Deu, M.; Courtois, B.; Mestres, C. Development of NIRS equations for food grain quality traits through exploitation of a core collection of cultivated sorghum. *J. Agric. Food Chem.* **2006**, *54*, 8501–8509. [[CrossRef](#)]
50. Alfieri, M.; Cabassi, G.; Habyarimana, E.; Quaranta, F.; Balconi, C.; Redaelli, R. Discrimination and prediction of polyphenolic compounds and total antioxidant capacity in sorghum grains. *JNIRS* **2019**, *27*, 46–53. [[CrossRef](#)]
51. Pojić, M.M.; Mastilović, J.S. Near infrared spectroscopy—advanced analytical tool in wheat breeding, trade, and processing. *Food Bioprocess Technol.* **2013**, *6*, 330–352. [[CrossRef](#)]
52. Kahriman, F.; Egesel, C.Ö. Comparison of spectral and molecular analyses for classification of long term stored wheat samples. *Guang Pu Xue Yu Guang Pu Fen Xi Guang Pu* **2016**, *36*, 1266–1272.
53. Levasseur-Garcia, C. Updated overview of infrared spectroscopy methods for detecting mycotoxins on cereals (Corn, Wheat, and Barley). *Toxins* **2018**, *10*, 38. [[CrossRef](#)]
54. Baeten, V.; Pierna, J.F.; Vermeulen, P.; Lecler, B.; Minet, O.; Zio, D.; Dardenne, P. Performance comparison of bench-top, hyperspectral imaging and pocket near infrared spectrometers: The example of protein quantification in wheat flour. In *Proceedings of the 18th International Conference on Near Infrared Spectroscopy*, Copenhagen, Denmark, 11–15 June 2017; Engelsen, S.B., Sørensen, K.M., van den Berg, F., Eds.; IM Publications Open: Chichester, UK, 2019; pp. 151–155.
55. Johnson, J.B. An overview of near-infrared spectroscopy (NIRS) for the detection of insect pests in stored grains. *J. Stored Prod. Res.* **2020**, *86*, 101558. [[CrossRef](#)]
56. Egesel, C.Ö.; Kahriman, F. Determination of quality parameters in maize grain by NIR reflectance spectroscopy. *Tarim Bilim. Derg.* **2012**, *18*, 31–42. [[CrossRef](#)]
57. Egesel, C.Ö.; Kahriman, F.; Ekinci, N.; Kavdir, İ.; Büyükcan, M.B. Analysis of fatty acids in kernel, flour, and oil samples of maize by NIR spectroscopy using conventional regression. *Cereal Chem.* **2016**, *93*, 487–492. [[CrossRef](#)]
58. Kahriman, F.; Onac, I.; Turk, F.; Öner, F.; Egesel, C.Ö. Determination of carotenoid and tocopherol content in maize flour and oil samples using near-infrared spectroscopy. *Spectrosc. Lett.* **2019**, *52*, 473–481. [[CrossRef](#)]
59. Kahriman, F.; Onac, I.; Oner, F.; Mert, F.; Egesel, C.Ö. Analysis of secondary biochemical components in maize flour samples by NIR (Near Infrared Reflectance) Spectroscopy. *J. Food Meas. Charact.* **2020**, *14*, 2320–2332. [[CrossRef](#)]
60. Serment, M.; Kahriman, F. Ability of near infrared spectroscopy and chemometrics to measure the phytic acid content in maize flour. *Spectrosc. Lett.* **2021**, *54*, 520–527. [[CrossRef](#)]
61. Abeshu, Y. Developing Calibration Model for Prediction of Malt Barley and Teff Genotypes Quality Traits Using Near Infrared Spectroscopy (NIRS). Ph.D. Thesis, Addis Ababa University, Addis Ababa, Ethiopia, 2019.
62. Abeshu, Y. Development of NIRS re-calibration model for ethiopian barley (*Hordeum vulgare*) lines traits to determine their brewing potential. *J. Agric. Food Inf.* **2021**, *1*, 100238. [[CrossRef](#)]
63. Albanell, E.; Martínez, M.; De Marchi, M.; Manuelian, C.L. Prediction of bioactive compounds in barley by near-infrared reflectance spectroscopy (NIRS). *J. Food Compos. Anal.* **2021**, *97*, 103763. [[CrossRef](#)]
64. Stubbs, T.L.; Kennedy, A.C.; Fortuna, A.M. Using NIRS to predict fiber and nutrient content of dryland cereal cultivars. *J. Agric. Food Chem.* **2010**, *58*, 398–403. [[CrossRef](#)]
65. Rosales, A.; Galicia, L.; Oviedo, E.; Islas, C.; Palacios-Rojas, N. Near-infrared reflectance spectroscopy (NIRS) for protein, tryptophan, and lysine evaluation in quality protein maize (QPM) breeding programs. *J. Agric. Food Chem.* **2011**, *59*, 10781–10786. [[CrossRef](#)]

66. Piaskowski, J.L.; Brown, D.; Campbell, K.G. Near-infrared calibration of soluble stem carbohydrates for predicting drought tolerance in spring wheat. *Agron. J.* **2016**, *108*, 285–293. [[CrossRef](#)]
67. Norman, H.C.; Hulm, E.; Humphries, A.W.; Hughes, S.J.; Vercoe, P.E. Broad near-infrared spectroscopy calibrations can predict the nutritional value of >100 forage species within the Australian feedbase. *Anim. Prod. Sci.* **2020**, *60*, 1111–1122. [[CrossRef](#)]
68. Zerihun, M.; Fox, G.; Nega, A.; Seyoum, A.; Minuye, M.; Jordan, D.; Taddese, T.; Assefa, A. Near-Infrared Reflectance Spectroscopy (NIRS) for Tannin, Starch and Amylase Determination in Sorghum Breeding Programs. *J. Food Nutr. Sci.* **2020**, *7*, 45–50. [[CrossRef](#)]
69. Carreira, E.; Serrano, J.; Shahidian, S.; Nogales-Bueno, J.; Rato, A.E. Real-time quantification of crude protein and neutral detergent fibre in pastures under montado ecosystem using the portable NIR spectrometer. *Appl. Sci.* **2021**, *11*, 10638. [[CrossRef](#)]
70. Smartfood-International Year of Millets. Available online: <https://www.smartfood.org/international-year-of-millets-2023/millet> (accessed on 8 February 2022).
71. Sustainable Development Goal 3. Available online: <https://in.one.un.org/page/sustainable-development-goals/sdg-3-2/> (accessed on 8 February 2022).
72. Mainstreaming Millets. Available online: <https://pib.gov.in/PressReleasePage.aspx?PRID=1783716> (accessed on 8 February 2022).
73. Li, X.; Siddique, K.H. Future smart food: Harnessing the potential of neglected and underutilized species for Zero Hunger. *Matern. Child Nutr.* **2020**, *16*, 13008. [[CrossRef](#)] [[PubMed](#)]
74. McKeivith, B. Nutritional aspects of cereals. *Nutr. Bull.* **2004**, *29*, 111–142. [[CrossRef](#)]
75. Girish, C.; Meena, R.K.; Mahima, D.; Mamta, K. Nutritional properties of minor millets: Neglected cereals with potentials to combat malnutrition. *Curr. Sci.* **2014**, *107*, 1109–1111.
76. Diao, X. Production and genetic improvement of minor cereals in China. *Crop. J.* **2017**, *5*, 103–114. [[CrossRef](#)]
77. Dodevska, M.S.; Djordjevic, B.I.; Sobajic, S.S.; Miletic, I.D.; Djordjevic, P.B.; Dimitrijevic-Sreckovic, V.S. Characterization of dietary fibre components in cereals and legumes used in Serbian diet. *Food Chem.* **2013**, *141*, 1624–1629. [[CrossRef](#)]
78. Belesova, K.; Gasparrini, A.; Sié, A.; Sauerborn, R.; Wilkinson, P. Household cereal crop harvest and children’s nutritional status in rural Burkina Faso. *Environ. Health* **2017**, *16*, 1. [[CrossRef](#)]
79. Rankoana, S.A. The use of indigenous knowledge in subsistence farming: Implications for sustainable agricultural production in dikgale community in Limpopo Province, South Africa. *Towar. Sustain. Agric. Farming Pract. Water Use* **2017**, *63*, 63–72.
80. Yang, Z.; Han, L.; Li, Q. Discriminant analysis of meat and bone meal content in ruminant feed based on NIRS. *Trans. Chin. Soc. Agric. Eng.* **2009**, *40*, 124–128.
81. Garcia, J.; Cozzolino, D. Use of near infrared reflectance (NIR) spectroscopy to predict chemical composition of forages in broad-based calibration models. *Agric. Téc.* **2006**, *66*, 41–47. [[CrossRef](#)]
82. Black, J.L.; Hughes, R.J.; Nielsen, S.G.; Tredrea, A.M.; Flinn, P.C. Near infrared reflectance analysis of grains to estimate nutritional value for chickens. In Proceedings of the 20th Australian Poultry Science Symposium, Sydney, NSW, Australia, 9–11 February 2009; Poultry Research Foundation: Brownlow Hill, NSW, Australia, 2009; pp. 31–34.
83. Tahir, M.; Shim, M.Y.; Ward, N.E.; Smith, C.; Foster, E.; Guney, A.C.; Pesti, G.M. Phytate and other nutrient components of feed ingredients for poultry. *Poult. Sci.* **2012**, *91*, 928–935. [[CrossRef](#)] [[PubMed](#)]
84. Atalay, H.; Kahrman, F.; Alatürk, F. Estimation of dry matter, crude protein and starch values in mixed feeds by near-infrared reflectance (NIR). *J. İst. Vet. Sci.* **2020**, *4*, 125–130. [[CrossRef](#)]
85. Atalay, H.; Kahrman, F. Estimating roughage quality with near infrared reflectance (NIR) spectroscopy and chemometric techniques. *Kocatepe Vet. J.* **2020**, *13*, 234–240. [[CrossRef](#)]
86. Kahrman, F.; Atalay, H. Estimation of relative feed value, relative forage quality and net energy lactation values of some roughage samples by using near infrared reflectance spectroscopy. *J. İst. Vet. Sci.* **2020**, *4*, 109–118. [[CrossRef](#)]
87. Cherney, J.H.; Digman, M.F.; Cherney, D.J. Portable NIRS for forage evaluation. *Comput. Electron. Agric.* **2021**, *190*, 106469. [[CrossRef](#)]
88. Assadzadeh, S.; Walker, C.K.; McDonald, L.S.; Maharjan, P.; Panozzo, J.F. Multi-task deep learning of near infrared spectra for improved grain quality trait predictions. *J. Near Infrared Spectrosc.* **2020**, *28*, 275–286. [[CrossRef](#)]
89. Lee, Y.Y.; Kim, J.B.; Lee, S.Y.; Lee, H.S.; Gwag, J.G.; Kim, C.K.; Lee, Y.B. Application of near-infrared reflectance spectroscopy to rapid determination of seed fatty acids in foxtail millet (*Setaria italica* (L.) P. Beauv) germplasm. *Korean J. Breed. Sci.* **2010**, *42*, 448–454.
90. Lee, Y.-Y.; Kim, J.-B.; Lee, H.-S.; Jeon, Y.-A.; Lee, S.-Y.; Kim, C.-K. Evaluation of millet (*Panicum miliaceum* subsp. *miliaceum*) germplasm for seed fatty acids using near-infrared reflectance spectroscopy. *Korean J. Crop Sci.* **2012**, *57*, 29–34. [[CrossRef](#)]
91. Yang, X.-S.; Wang, L.-L.; Zhou, X.-R.; Shuang, S.-M.; Zhu, Z.-H.; Li, N.; Li, Y.; Liu, F.; Liu, S.-C.; Lu, P.; et al. Determination of protein, fat, starch, and amino acids in foxtail millet *Setaria italica* (L.) Beauv. by Fourier transform near-infrared reflectance spectroscopy. *Food Sci. Biotechnol.* **2013**, *22*, 1495–1500. [[CrossRef](#)]
92. Bhardwaj, R.; Yadav, S.; Suneja, P. NIRS based food quality assessment approaches for cereals, oilseeds, pulses, fruits and vegetables. In Proceedings of the 7th Indo-Global Summit and Expo on Food & Beverages, New Delhi, India, 8–10 October 2015.
93. Wheat Trading Standards in Australia. Available online: https://www.graintrade.org.au/commodity_standards (accessed on 8 February 2022).
94. Wheat quality and Markets in Queensland, Department of Agriculture and Fisheries, Queensland. Available online: https://www.daf.qld.gov.au/__data/assets/pdf_file/0006/53799/Wheat-FactSheet-Quality-Markets-Qld.pdf (accessed on 8 February 2022).

95. Huck, C.W. New Trend in Instrumentation of NIR Spectroscopy—Miniaturization. In *Near-Infrared Spectroscopy*; Ozaki, Y., Huck, C., Tsuchikawa, S., Engelsens, S.B., Eds.; Springer: Singapore, 2021; pp. 193–210. [CrossRef]
96. Genebank of ICRISAT. Available online: <https://www.genebank.icrisat.org> (accessed on 22 December 2021).
97. Upadhyaya, H.D.; Pundir, R.P.S.; Dwivedi, S.L.; Gowda, C.L.L.; Reddy, V.G.; Singh, S. Developing a mini core collection of sorghum for diversified utilization of germplasm. *Crop Sci.* **2009**, *49*, 1769–1780. [CrossRef]
98. Jordan, D.R.; Mace, E.S.; Cruickshank, A.W.; Hunt, C.H.; Henzell, R.G. Exploring and exploiting genetic variation from unadapted sorghum germplasm in a breeding program. *Crop Sci.* **2011**, *51*, 1444–1457. [CrossRef]
99. Deshpande, S.; Rakshit, S.; Manasa, K.G.; Pandey, S.; Gupta, R. Genomic Approaches for Abiotic Stress Tolerance in Sorghum. In *The Sorghum Genome. Compendium of Plant Genomes*; Springer: Berlin/Heidelberg, Germany, 2016; pp. 169–187.
100. Kassahun, B.; Bidinger, F.R.; Hash, C.T.; Kuruvinashetti, M.S. Stay-green expression in early generation sorghum [*Sorghum bicolor* (L.) Moench] QTL introgression lines. *Euphytica* **2010**, *172*, 351–362. [CrossRef]
101. Sehgal, D.; Skot, L.; Singh, R.; Srivastava, R.K.; Das, S.P.; Taunk, J.; Sharma, P.C.; Pal, R.; Raj, B.; Hash, C.T.; et al. Exploring potential of pearl millet germplasm association panel for association mapping of drought tolerance traits. *PLoS ONE* **2015**, *10*, e0122165. [CrossRef] [PubMed]
102. ICAR. Handbook of Agriculture. In *Directorate of Publications and Information on Agriculture*; ICAR Publication: New Delhi, India, 2011.
103. Association of Official Analytical Chemists (AOAC) International. *Official Methods of Analysis*, 17th ed.; Association of Official Analytical Chemists (AOAC) International: Gaithersburg, MD, USA, 2000.
104. Samireddypalle, A.; Boukar, O.; Grings, E.; Fatokun, C.A.; Kodukula, P.; Devulapalli, R.; Okike, I.; Blümmel, M. Cowpea and groundnut haulms fodder trading and its lessons for multidimensional cowpea improvement for mixed crop livestock systems in West Africa. *Front. Plant Sci.* **2017**, *8*, 30. [CrossRef] [PubMed]
105. Jayawardana, S.A.S.; Samarasekera, J.K.R.R.; Hettiarachchi, G.H.C.M.; Gooneratne, J.; Mazumdar, S.D.; Banerjee, R. Dietary fibers, starch fractions and nutritional composition of finger millet varieties cultivated in Sri Lanka. *J. Food Compos. Anal.* **2019**, *82*, 103249. [CrossRef]
106. Hone Lab Video. Available online: https://www.youtube.com/watch?v=c7f_p3p-SVg (accessed on 8 February 2022).
107. Hone Create Platform. Available online: <https://www.honecreate.com> (accessed on 8 February 2022).
108. Williams, P. Calibration development and evaluation methods B. Set-up and evaluation. *NIR News* **2013**, *24*, 20–24. [CrossRef]
109. Galvao, R.; Araujo, M.; Jose, G.; Pontes, M.; Silva, E.; Saldanha, T. A method for calibration and validation subset partitioning. *Talanta* **2005**, *67*, 736–740. [CrossRef]
110. Kempes, B.J.; Saeys, W.; Mertens, K.; Darius, P.; De Baerdemaeker, J.G.; De Ketelaere, B. The importance of choosing the right validation strategy in inverse modelling. *JNIRS* **2010**, *18*, 231–237. [CrossRef]
111. Au, J.; Youngentob, K.N.; Foley, W.J.; Moore, B.D.; Fearn, T. Sample selection, calibration and validation of models developed from a large dataset of near infrared spectra of tree leaves. *JNIRS* **2020**, *28*, 186–203. [CrossRef]
112. Kamilaris, A.; Prenafeta-Boldú, F.X. Deep learning in agriculture: A survey. *Comput. Electron. Agric.* **2018**, *147*, 70–90. [CrossRef]
113. Fandango, A. *Mastering TensorFlow 1. x: Advanced Machine Learning and Deep Learning Concepts Using TensorFlow 1. x and Keras*; Packt Publishing Ltd.: Birmingham, UK, 2018.
114. Savitzky, A.; Golay, M.J. Smoothing and differentiation of data by simplified least squares procedures. *Anal. Chem.* **1964**, *36*, 1627–1639. [CrossRef]
115. Barnes, R.J.; Dhanoa, M.S.; Lister, S.J. Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra. *Appl. Spectrosc.* **1989**, *43*, 772–777. [CrossRef]
116. Hopkins, D.M. Using data pretreatments effectively. In *Proceedings of the International Diffuse Reflectance Conference*, Chambersburg, PA, USA, 3–8 August 2008.
117. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv* **2014**, arXiv:1412.6980.
118. Zhang, X.; Lin, T.; Xu, J.; Luo, X.; Ying, Y. DeepSpectra: An end-to-end deep learning approach for quantitative spectral analysis. *Anal. Chim. Acta.* **2019**, *1058*, 48–57. [CrossRef]
119. Williams, P. The RPD statistic: A tutorial note. *NIR News* **2014**, *25*, 22–26. [CrossRef]
120. Williams, P.; Dardenne, P.; Flinn, P. Tutorial: Items to be included in a report on a near infrared spectroscopy project. *J. Near Infrared Spectrosc.* **2017**, *25*, 85–90. [CrossRef]
121. Williams, P.; Manley, M.; Antonicyzn, J. *Near-Infrared Technology-Getting the Best Out of Light*; Sun Press Imprint: Stellenbosch, South Africa, 2019; p. 301.
122. Kumar, A.; Tomer, V.; Kaur, A.; Kumar, V.; Gupta, K. Millets: A solution to agrarian and nutritional challenges. *Agric. Food Secur.* **2018**, *7*, 31. [CrossRef]
123. Cozzolino, D. An overview of the use of infrared spectroscopy and chemometrics in authenticity and traceability of cereals. *Food Res. Int.* **2014**, *60*, 262–265. [CrossRef]
124. Jordan, M.I.; Mitchell, T.M. Machine learning: Trends, perspectives, and prospects. *Science* **2015**, *349*, 255–260. [CrossRef] [PubMed]
125. Huang, Z.; Sha, S.; Rong, Z.; Chen, J.; He, Q.; Khan, D.M.; Zhu, S. Feasibility study of near infrared spectroscopy with variable selection for non-destructive determination of quality parameters in shell-intact cottonseed. *Ind. Crops Prod.* **2013**, *43*, 654–660. [CrossRef]