

RESEARCH ARTICLE

Transfer of near-infrared spectroscopic prediction models for corn nutritional components: Comparison of different preprocessing and wavelength selection methods

Longxin Guo*

Haide College, Ocean University of China, Qingdao, Shandong, China.

Received: May 8, 2026; accepted: August 25, 2026.

Near-infrared (NIR) spectroscopy provides a rapid method for predicting corn nutritional components. However, spectral differences among instruments can reduce model accuracy after transfer. This study applied a public corn dataset measured with three instruments of m5, mp5, and mp6 to examine the effects of preprocessing, wavelength selection, and slope-bias correction (SBC) on the prediction of moisture, oil, protein, and starch. Partial least squares regression was used to build prediction models. Four preprocessing methods and four wavelength-selection strategies were compared. The m5 instrument was used as the master instrument for model transfer to mp5 and mp6. For single-instrument modeling, the first derivative gave the best results for moisture, oil, and protein. Starch performed the best when the first derivative was combined with standard normal variate (SNV). On m5, the best models achieved coefficients of determination for prediction (R^2P) of 0.989, 0.950, 0.966, and 0.969 for moisture, oil, protein, and starch, respectively. Wavelength selection improved some models, but the effect differed among the four components. Competitive adaptive reweighted sampling (CARS) increased the R^2P of the moisture model from 0.988 to 0.996 and reduced the root mean square error of prediction (RMSEP) from 0.035 to 0.026. The Union method increased the R^2P values of oil and protein from 0.932 to 0.958 and from 0.980 to 0.988, respectively. Only a small improvement was found for starch. Direct transfer of the m5 models to mp5 and mp6 produced clear prediction bias. SBC reduced this bias and improved the agreement between predicted and reference values. The effects of preprocessing and wavelength selection depended on the target component. Oil and protein benefited from combining preprocessing with wavelength selection. Starch relied mainly on CARS, while moisture showed no stable improvement across the two transfer directions. These results showed that SBC was important for cross-instrument prediction. Preprocessing and wavelength-selection methods should be selected separately for each corn component.

Keywords: near-infrared spectroscopy; corn; model transfer; preprocessing; wavelength selection.

*Corresponding author: Longxin Guo, Haide College, Ocean University of China, Qingdao, Shandong 266100, China. Email: glxdell@outlook.com.

Introduction

Corn is widely used as food, animal feed, and an industrial raw material [1]. It is also an important source of starch, oil, and protein [2]. The contents of these components affect grain storage, feed formulation, processing

performance, and product quality [3]. Reliable measurement of moisture, oil, protein, and starch is therefore important in corn breeding, storage, processing, and quality control [4, 5]. Traditional chemical methods are still widely used for corn quality analysis. Moisture is commonly measured by oven drying. Oil can be

determined by Soxhlet extraction, while protein is often measured by the Kjeldahl method. Chemical or enzymatic methods can be used for starch analysis [6]. These methods can provide accurate results, but they usually require sample preparation, chemical reagents, trained operators, and a relatively long testing time [7]. They are not well suited to rapid testing of large sample groups or online monitoring during production [8, 9].

Near-infrared (NIR) spectroscopy is a rapid and non-destructive method for material analysis. Its spectral information is related to chemical bonds such as C–H, O–H, and N–H. These bonds are closely linked to water, fat, protein, and carbohydrates. NIR spectroscopy has been applied in agricultural product evaluation, food safety testing, feed analysis, and process monitoring [10–12]. Previous studies have also used NIR spectra to predict several corn components. Xu *et al.* reported test-set R^2 values of 0.996, 0.990, 0.974, and 0.976 for moisture, oil, protein, and starch, respectively [13]. Zou *et al.* combined a convolutional neural network (CNN) with eXtreme Gradient Boosting (XGBoost) feature extraction with the test-set R^2 values ranged from 0.991 to 0.998 [14]. These results show that NIR spectroscopy can provide useful information for the rapid analysis of corn components. A model developed on one instrument may lose accuracy when it is applied to another instrument. Differences in optical design, detector response, calibration, and measurement conditions can change the recorded spectra [15], which can arise even when the same samples and wavelength range are used. In the public Corn dataset, m5 and mp5 refer to two FOSS NIRSystems 5000 spectrometers, while mp6 refers to a FOSS NIRSystems 6000 spectrometer [16]. Cataltas and Tutuncu used the three instruments to measure the same 80 corn samples and reported that a partial least squares regression (PLSR) model for moisture achieved an R^2 of 0.996 on m5, 0.827 on mp5, and 0.757 on mp6 [16]. This difference shows that cross-instrument use remains a practical problem for NIR models. Preprocessing

and wavelength selection are often used to improve model performance, but their effects after model transfer are not always clear.

The purpose of this study was to examine the effects of preprocessing and wavelength selection on single-instrument prediction and cross-instrument model transfer. PLSR models were developed for moisture, oil, protein, and starch. Four preprocessing methods and four wavelength sets were compared. The m5 models were then transferred to mp5 and mp6. Slope-bias correction was used to reduce the prediction bias caused by instrument differences. The study compared the different methods under the same data and modeling conditions and examined whether one strategy could be used for all four corn components. The results provided information for selecting preprocessing, wavelength-selection, and correction methods in cross-instrument corn analysis, which might support the wider use of NIR models when several instruments were involved.

Materials and methods

Data source and NIR spectra

The public corn dataset was obtained from the data repository of Eigenvector Research, Inc. (<https://eigenvector.com/resources/data-sets/>), which contained NIR spectra of 80 corn samples measured with three instruments labeled m5, mp5, and mp6. The spectra covered the wavelength range from 1,100 to 2,498 nm at intervals of 2 nm. Each spectrum therefore contained 700 wavelength variables. The same samples and wavelength points were used for the three instruments. Therefore, the three spectral matrices were the same size although their recorded spectral responses differed. The dataset also contained reference values for the same 80 samples. Moisture, oil, protein, and starch were used as the target components in this study with the approximate ranges of 9 – 11% for moisture, 3 – 4% for oil, 8 – 10% for protein, and 63 – 66% for starch. These values were used as the response variables in the

regression models and transfer analysis.

Construction of quantitative NIR models

During acquisition, NIR spectra could be affected by sample condition, instrument response, and measurement environment. These factors might cause baseline drift, scattering effects, and unwanted spectral variation. Four preprocessing methods were examined including no preprocessing (none), Savitzky–Golay first derivative (1st), standard normal variate (SNV), and first derivative followed by SNV (1st+SNV). The first derivative was calculated with the Savitzky–Golay filter [17]. A window size of 11 and a polynomial order of 2 were used. The SciPy implementation of the filter was performed (https://docs.scipy.org/doc/scipy/reference/generated/scipy.signal.savgol_filter.html). SNV was applied separately to each spectrum through mean centering and standardization [18]. PLSR models were built separately for moisture, oil, protein, and starch on each instrument. The samples were divided into calibration and prediction sets by the Kennard–Stone algorithm. Five-fold cross-validation was used to select the number of latent variables. The one-standard-error (1SE) rule was then applied to avoid unnecessary model complexity. The latent variable number with the lowest root mean square error of cross-validation (RMSECV) was identified. The final model used the smallest number of latent variables whose RMSECV was within one standard error of the minimum value. The maximum number of latent variables was set to 12. After the best preprocessing method was identified for each component, wavelength selection was performed on the m5 spectra. Four wavelength sets were compared including the full spectrum (none), competitive adaptive reweighted sampling (CARS), interval partial least squares (iPLS), and Union. CARS retained important wavelengths according to the variable importance in projection threshold of $VIP \geq 1$. In iPLS, the full spectrum was divided into several intervals, and the better-performing intervals were selected according to their cross-validation results. Union combined the wavelengths selected by CARS and iPLS. PLSR models were

built for all wavelength sets using the same Kennard–Stone sample division and cross-validation procedure. Model performance was compared using the prediction-set results. R^2P was used as the main selection criterion, while RMSEP was used as a supplementary measure of prediction error. Cross-instrument model transfer was also examined because systematic differences were present among the three instruments [19, 20]. The m5 instrument was used as the master instrument, while mp5 and mp6 were used as the target instruments. The PLSR models developed on m5 were first applied directly to the target-instrument spectra to obtain the Direct predictions. Slope-bias correction (SBC) was then used to correct these predictions [21]. Representative spike samples were selected from each target-instrument dataset and accounted for approximately 20% of the samples. A linear relationship was fitted between the reference values and direct predictions of the spike samples. This relationship was then applied to all predictions from the target instrument to obtain the SBC-corrected results.

Model evaluation metrics

To evaluate the predictive performance of the calibration and regression models under different preprocessing conditions, variable selection strategies, and cross-instrument transfer settings, error-based metrics and coefficients of determination were employed. The error metrics included the root mean square error of cross validation ($RMSE_{CV}$) and the root mean square error of prediction ($RMSE_P$). $RMSE_{CV}$ was used to assess performance during latent variable selection and model comparison, while $RMSE_P$ reflected the model's extrapolation error on the prediction set. These RMSE values were calculated as follows.

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \quad (1)$$

where N was the number of samples in the corresponding dataset. y_i and $\hat{y}_i$ were the reference value and predicted value of the i -th

sample, respectively. The coefficient of determination was used to describe the agreement between predicted and reference values and was calculated for the cross-validation, calibration, and prediction sets as R^2_{CV} , and R^2_P , respectively. In this study, R^2 was calculated as the squared Pearson correlation coefficient between the reference values and predicted values as below.

$$R^2 = \left(\frac{\sum_{i=1}^N (y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\sqrt{\sum_{i=1}^N (y_i - \bar{y})^2} \sqrt{\sum_{i=1}^N (\hat{y}_i - \bar{\hat{y}})^2}} \right)^2 \quad (2)$$

where N was the number of samples in the corresponding dataset. y_i and $\hat{y}_i$ were the reference and predicted values of the i -th sample. $\bar{y}$ and $\bar{\hat{y}}$ were the mean values of the reference and predicted data, respectively. R^2 ranged from 0 to 1, and a larger value indicated better agreement between model predictions and reference values. In the subsequent comparison of variable-selection strategies, R^2_P was taken as the primary criterion, while $RMSE_P$ was used as a supplementary indicator of prediction error magnitude.

Results and discussion

NIR spectral characteristics of corn collected by different instruments and principal component analysis

The overall characteristics of the corn NIR spectra collected by the three instruments (m5, mp5, and mp6), the statistical distribution of the four physicochemical reference values, and the relative positions of the spectra from different instruments in the low dimensional principal component space demonstrated that the mean spectra of the three datasets showed a consistent overall trend across the wavelength range, indicating that the three instruments provided generally consistent responses to the major absorption information in corn samples. However, noticeable intensity differences were still observed in several wavelength regions. In

particular, the overall absorbance level of m5 was relatively higher, whereas the spectral curves of mp5 and mp6 were more similar to each other (Figure 1a). Such relatively stable intensity differences were usually related to differences in optical path design, detector performance, or instrument calibration. These differences might lead to evident instrument-related discrepancies in cross-instrument applications and provided the basis for subsequent cross-instrument modeling comparison and model transfer analysis. The distribution of the physicochemical reference values further indicated that there were measurable differences in the target constituent contents among samples, which provided the basis for subsequent model development. Among the four indicators, starch content was clearly higher than moisture, oil, protein and exhibited a relatively wider range of variation. In contrast, the value ranges of moisture, oil, and protein were more concentrated although clear differences among samples were still observed (Figure 1b). Overall, the variation in these reference values across samples provided the necessary basis for establishing stable quantitative regression models and suggested that different constituents might differ in modeling difficulty and sensitivity. After mean centering the spectral matrix, principal component analysis was further performed. The score distribution obtained by combining data from all three instruments showed that the first two principal components explained most of the variation in the dataset ($PC1 \approx 95.98\%$, $PC2 \approx 3.73\%$), while also revealing clear differences among instruments. Samples from m5 were largely separated from those of mp5 and mp6 in the principal component space, indicating that the overall spectral distribution of m5 differed relatively consistently from that of the other two instruments (Figure 1c). Although the samples from mp5 and mp6 overlapped to some extent, differences in their distributions could still be observed. When only the mp5 and mp6 datasets were analyzed, the explanatory power of the first two principal components became even more concentrated ($PC1 \approx 98.87\%$, $PC2 \approx 0.91\%$), and the score distributions of the

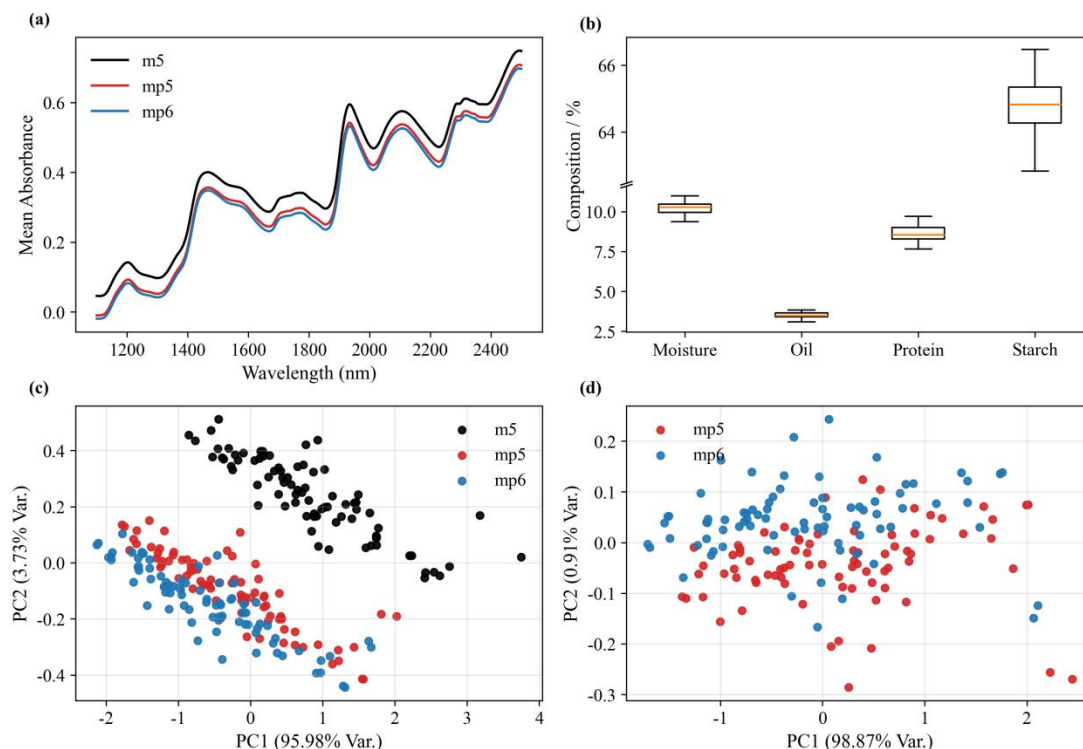


Figure 1. NIR spectral characteristics of corn collected by different instruments and principal component analysis.

two instruments largely overlapped, suggesting that their spectral characteristics were more similar. Nevertheless, a small number of samples still showed slight shifts along the main direction, which indicated that, in cross-instrument applications, calibration or model transfer methods might still be needed to reduce the remaining systematic differences (Figure 1d).

Quantitative NIR models for corn nutritional constituents

The preprocessing results followed a clear pattern across the four components. The first derivative gave the best overall results for moisture, oil, and protein. For starch, the best results were obtained by combining the first derivative with SNV. Model performance also differed among the three instruments. For moisture, the first derivative improved the m5 model from R^2P of 0.974 and RMSEP of 0.057 without preprocessing to R^2P of 0.989 and RMSEP of 0.040. The same preprocessing method gave lower prediction accuracy on mp5 and mp6 with their R^2P values of 0.865 and 0.827, RMSEP

values of 0.128 and 0.161, respectively. SNV alone and 1st+SNV produced weaker results for moisture. The first derivative also gave the best overall results for oil. The m5 model reached R^2P of 0.950 and RMSEP of 0.041. The corresponding R^2P values for mp5 and mp6 were 0.817 and 0.834. Their RMSEP values increased to 0.077 and 0.100. This difference showed that the oil models were more sensitive to the spectra collected by mp5 and mp6. Protein prediction followed a similar pattern. After first-derivative preprocessing, the m5 model reached R^2P of 0.966 and RMSEP of 0.099. The mp5 and mp6 models reached R^2P of 0.945 and 0.915 with RMSEP values of 0.148 and 0.143. The results showed that useful protein-related information was present in the spectra from all three instruments although m5 provided the most accurate predictions. Starch required a different preprocessing method. For 1st+SNV, the m5 and mp5 models reached R^2P values of 0.969 and 0.959 with RMSEP values of 0.166 and 0.180. The mp6 model was weaker with R^2P of 0.820 and RMSEP of 0.185. The reference values of starch

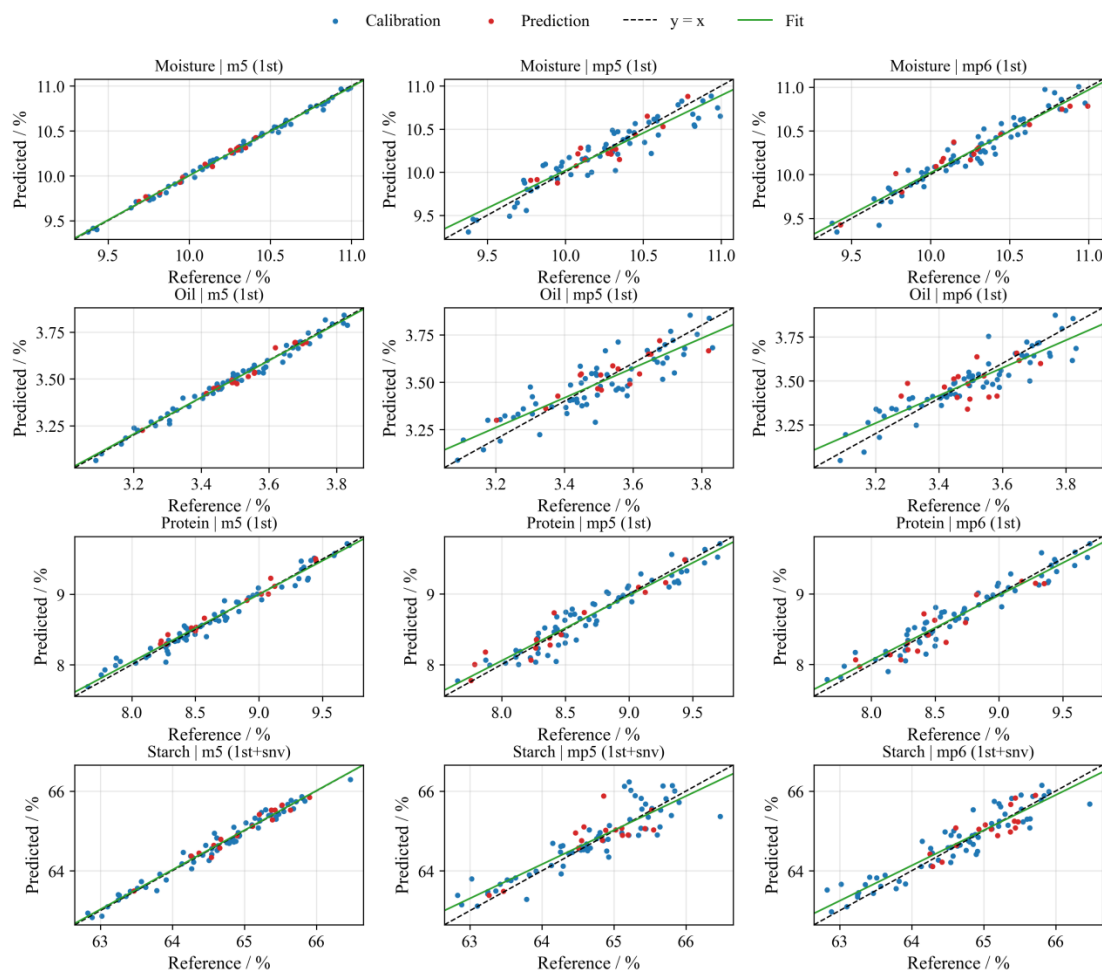


Figure 2. Calibration–prediction scatter plots of the PLS models for corn moisture, oil, protein, and starch based on the optimal preprocessing methods.

covered a narrow range of about 63 – 66%. The remaining error on mp6 was therefore important in relation to the sample range. This result suggested that 1st+SNV could not fully remove the instrument-related effects in the mp6 spectra. The calibration and prediction results obtained with the selected preprocessing method for each component showed that the points from m5 were closely distributed around the $y = x$ line. The calibration and prediction points also showed similar distributions. These patterns agree with the higher R^2_P and lower RMSEP values of the m5 models. The points from mp5 and mp6 were more dispersed. Some fitted lines also deviated from the $y = x$ line. The difference was most evident in the mp6 starch model. Its prediction points showed a wider

spread and weaker agreement with the reference values. The starch models on m5 and mp5 retained a clearer linear relationship (Figure 2). The results showed that the choice of preprocessing depended on the target component. The first derivative was more suitable for moisture, oil, and protein, while 1st+SNV was more suitable for starch. Instrument effects remained clear after preprocessing. M5 gave the best overall performance followed by mp5, while mp6 showed the largest decline for starch.

Quantitative models for corn nutritional components based on selected NIR wavelengths
Wavelength selection affected the four corn components in different ways. The number of

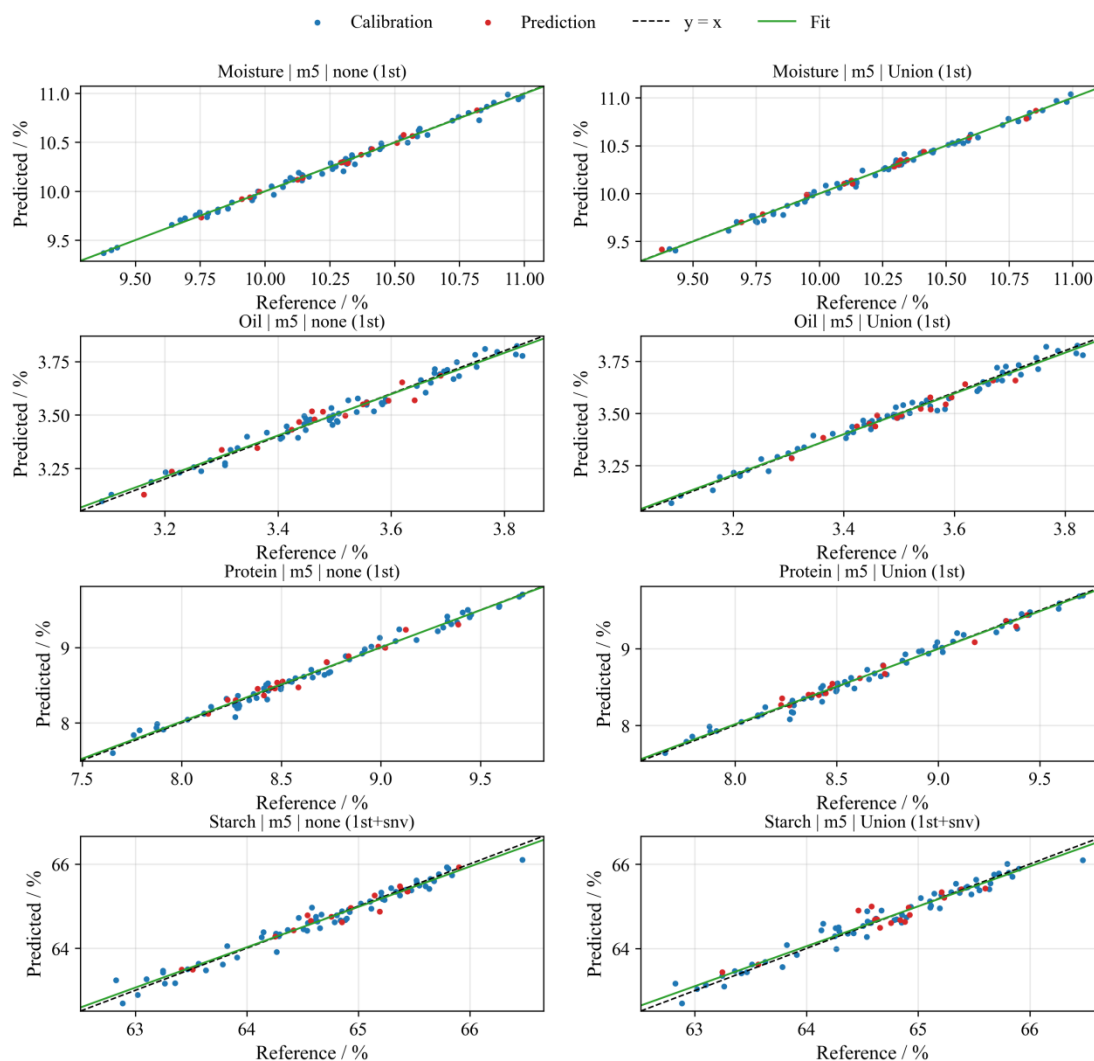


Figure 3. Comparison of the fitting results between full-spectrum models and feature-variable-selected models.

input variables was reduced in all selected models, but the change in prediction performance depended on the target component. The full-spectrum moisture model used 700 wavelengths and reached R^2P of 0.988 with RMSEP of 0.035. CARS reduced the number of wavelengths to 287 and increased R^2P to 0.996, while reduced RMSEP to 0.026, which was the clearest improvement among the moisture models. For oil, the full-spectrum model reached R^2P of 0.932 and RMSEP of 0.038. Union retained 345 wavelengths and improved R^2P to 0.958. RMSEP decreased to 0.030. CARS and iPLS also improved the prediction results, but Union gave the best overall performance. Protein showed a

clear reduction in prediction error after Union selection. The full-spectrum model reached R^2P of 0.980 and RMSEP of 0.075. With 345 Union wavelengths, R^2P increased to 0.988 and RMSEP decreased to 0.053. The result suggested that useful protein information was distributed across several spectral regions retained by the two selection methods. The effect of wavelength selection was limited for starch. The full-spectrum model reached R^2P of 0.919 and RMSEP of 0.193. CARS retained 220 wavelengths and produced a slightly higher R^2P of 0.920, but RMSEP increased to 0.208. The other selected wavelength sets did not improve the model. Wavelength reduction therefore provided little

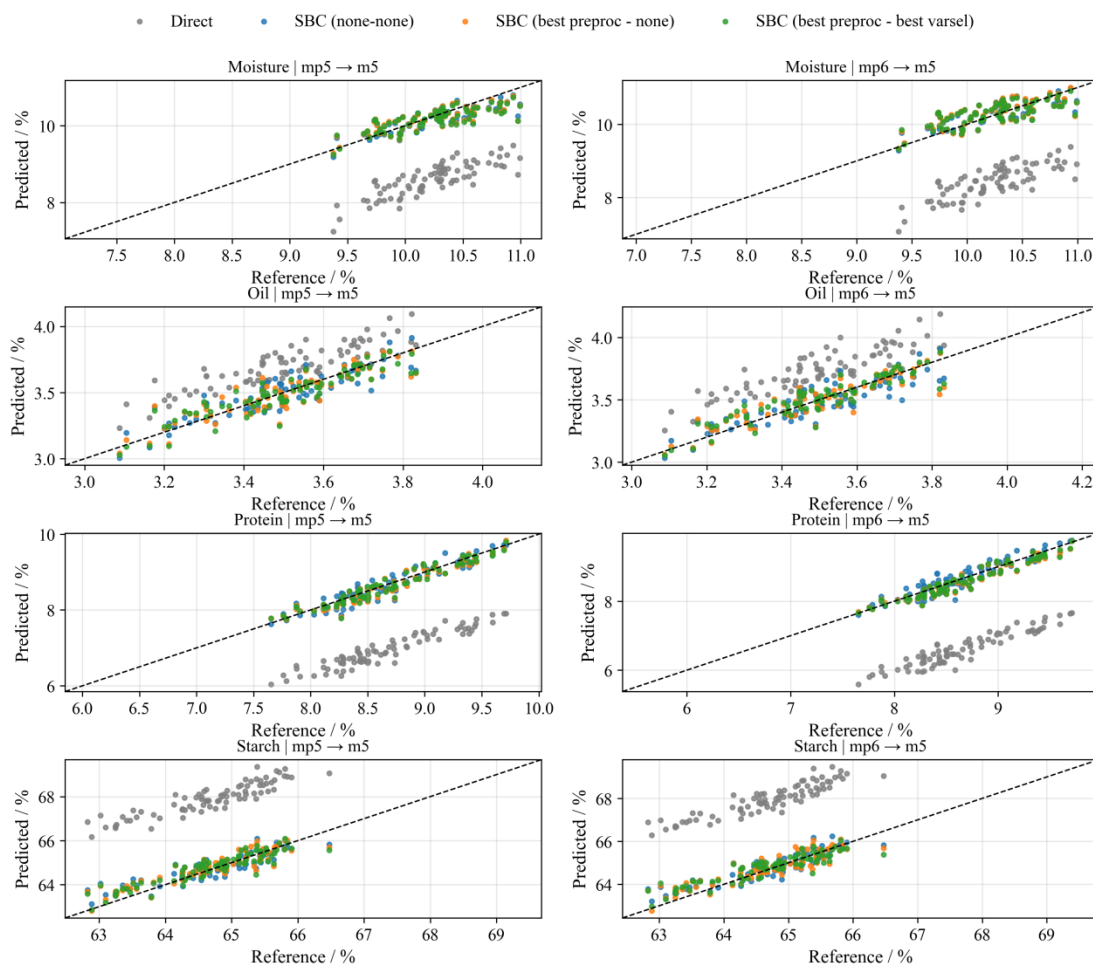


Figure 4. Comparison of model prediction fitting under different transfer strategies.

benefit for starch prediction under the current modeling conditions. The fitting results of the full-spectrum and selected-wavelength models were compared. The regression lines remained close to the $y = x$ line in most panels. Moisture, oil, and protein retained clear linear relationships after wavelength selection. The prediction points for oil and protein also became more concentrated. The difference was smaller for starch, and its points remained more dispersed (Figure 3). The results showed that one wavelength-selection method was not suitable for all four components. CARS gave the selected result for moisture and starch, while Union gave the selected result for oil and protein. These wavelength sets were used in the following model-transfer analysis.

Model transfer of quantitative models among different NIR instruments

The effect of preprocessing and wavelength selection on model transfer differed among the four components. For moisture on mp5, the combination of no preprocessing and CARS gave the highest R^2P of 0.935 and the lowest RMSEP of 0.075. The same combination did not improve the mp6 result. On mp6, first-derivative preprocessing without wavelength selection gave R^2P of 0.896 and RMSEP of 0.158. Combining the first derivative with CARS reduced the performance of R^2P to 0.867 and RMSEP to 0.196. Moisture therefore showed no common improvement pattern across the two target instruments. Oil showed a more consistent result. The combination of the first derivative and Union gave the best performance on both mp5

and mp6. For each target instrument, this combination reached R^2P of 0.821 and RMSEP of 0.076. The full-spectrum models without preprocessing reached R^2P values of only 0.775 on mp5 and 0.763 on mp6. Protein also benefited from the combined use of preprocessing and wavelength selection. The first derivative with Union reached R^2P of 0.928 and RMSEP of 0.139 on mp5. On mp6, the same combination gave the highest R^2P of 0.936 with RMSEP of 0.145. These results indicated that the selected preprocessing and wavelength set improved the agreement between predicted and reference protein values. For starch, CARS without preprocessing gave the best transfer results. The model reached R^2P of 0.848 and RMSEP of 0.323 on mp5. On mp6, it reached R^2P of 0.842 and RMSEP of 0.341. Adding 1st+SNV did not improve these results. Starch transfer therefore depended more on wavelength selection than on preprocessing. The comparison of the direct and SBC-corrected predictions demonstrated that the direct predictions showed clear shifts from the $y = x$ line for all four components. These shifts were presented on both target instruments. After SBC, the prediction points moved closer to the reference line. The change was especially clear for oil and protein. Some differences remained after correction. The moisture points were still relatively dispersed, which agreed with the lack of a stable preprocessing and wavelength-selection pattern. The starch results improved after SBC, but visible dispersion remained in the higher-value region. Oil and protein showed more compact distributions after correction (Figure 4). These findings showed that SBC reduced the main linear bias caused by instrument differences. The effects of preprocessing and wavelength selection were component specific. Oil and protein benefited from the first derivative combined with Union. Starch performed better with CARS alone. Moisture did not show a stable improvement across mp5 and mp6. A single preprocessing and wavelength-selection strategy was therefore not suitable for all four components.

Conclusions

The three instruments produced similar overall spectral patterns, but their differences were large enough to affect model prediction. Among them, m5 provided the most reliable single-instrument models and was suitable for use as the master instrument. The effects of preprocessing and wavelength selection depended on the corn component. The first derivative was more suitable for moisture, oil, and protein, while starch required first derivative combined with SNV in single-instrument modeling. Wavelength selection improved the models for moisture, oil, and protein, but had only a limited effect on starch. Direct model transfer caused clear prediction bias between instruments. SBC reduced this bias and improved the agreement between predicted and reference values. Additional preprocessing and wavelength selection did not provide the same benefit for every component. Oil and protein responded well to the combined use of preprocessing and wavelength selection. Starch depended more on CARS, while moisture showed no consistent pattern. Cross-instrument NIR analysis should therefore use a separate modeling and transfer strategy for each target component.

References

1. Maitra S, Singh V. 2021. Invited review on “maize in the 21st century”: Emerging trends of maize biorefineries in the 21st century: scientific and technological advancements in biofuel and bio-sustainable market. *J Cereal Sci.* 101:103272.
2. Yu J, Moon Y. 2022. Corn starch: Quality and quantity improvement for industrial uses. *Plants.* 11(1):92.
3. Obadi M, Qi Y, Xu B. 2022. High-amylose maize starch: Structure, properties, modifications and industrial applications. *Carbohydr Polym.* 299:120185.
4. Wang Z, Fan Y, Tian X, Long Y, Huang W, Chen L. 2025. Development of a rapid detection method for maize seed purity using a modular high-throughput near-infrared non-destructive testing system. *Infrared Phys Technol.* 148:105836.
5. Yu B, Chen X, Zhou C, Ding T, Wang Z, Zou C. 2022. Nutritional composition of maize grain associated with phosphorus and zinc fertilization. *J Food Compos Anal.* 114:104775.
6. Tomar M, Bhardwaj R, Singh P, Kaur S, Singh SP, Dahuja A, *et al.* 2025. From grain to gain: Bridging conventional methods with

- chemometric innovations in cereal quality analysis through near-infrared spectroscopy. *Food Control*. 178:111482.
7. Tian W, Li Y, Guzman C, Ibba MI, Tilley M, Wang D, *et al.* 2023. Quantification of food bioactives by NIR spectroscopy: Current insights, long-lasting challenges, and future trends. *J Food Compost Anal*. 124:105708.
 8. Porep JU, Kammerer DR, Carle R. 2015. On-line application of near infrared spectroscopy in food production. *Trends Food Sci Technol*. 46(2):211-230.
 9. Priya TSR, Manickavasagan A. 2021. Characterising corn grain using infrared imaging and spectroscopic techniques: A review. *J Food Meas Charact*. 15(4):3234-3249.
 10. Fodor M, Matkovits A, Benes EL, Jókai Z. 2024. The role of near-infrared spectroscopy in food quality assurance: A review of the past two decades. *Foods*. 13(21):3501.
 11. Bailly S, Orlando B, Brustel J, Bailly J, Levasseur-Garcia C. 2024. Rapid detection of aflatoxins in ground maize using near infrared spectroscopy. *Toxins*. 16(9):385.
 12. Pandiselvam R, Sruthi NU, Kumar A, Kothakota A, Thirumdas R, Ramesh S, *et al.* 2023. Recent applications of vibrational spectroscopic techniques in the grain industry. *Food Rev Int*. 39(1):209-239.
 13. Xu L, Liu J, Wang C, Li Z, Zhang D. 2023. Rapid determination of the main components of corn based on near-infrared spectroscopy and a BiPLS-PCA-ELM model. *Appl Opt*. 62(11):2756-2765.
 14. Zou X, Wang Q, Chen Y, Wang J, Xu S, Zhu Z, *et al.* 2025. Fusion of convolutional neural network with XGBoost feature extraction for predicting multi-constituents in corn using near infrared spectroscopy. *Food Chem*. 463:141053.
 15. Zhang W, Kasun LC, Wang QJ, Zheng Y, Lin Z. 2022. A review of machine learning for near-infrared spectroscopy. *Sensors*. 22(24):9764.
 16. Cataltas O, Tutuncu K. 2023. Detection of protein, starch, oil, and moisture content of corn kernels using one-dimensional convolutional autoencoder and near-infrared spectroscopy. *PeerJ Comput Sci*. 9:e1266.
 17. Savitzky A, Golay MJE. 1964. Smoothing and differentiation of data by simplified least squares procedures. *Anal Chem*. 36(8):1627-1639.
 18. Rinnan Å, van den Berg F, Engelsen SB. 2009. Review of the most common pre-processing techniques for near-infrared spectra. *TrAC Trends Anal Chem*. 28(10):1201-1222.
 19. Kamruzzaman M, Kalita D, Ahmed MT, ElMasry G, Makino Y. 2021. Effect of variable selection algorithms on model performance for predicting moisture content in biological materials using spectral data. *Anal Chim Acta*. 1202:339390.
 20. Workman JJ. 2018. A review of calibration transfer practices and instrument differences in spectroscopy. *Appl Spectrosc*. 72(3):340-365.
 21. Zhang X, Chen X, Xiong Z, Siesler HW, Liang L. 2022. Transfer of a calibration model for the prediction of lignin in pulpwood among four portable near infrared spectrometers. *J Near Infrared Spectrosc*. 30(4):208-218.