

Research article

Comparative Study of Spectrometer Sensors for Corn Moisture Content Prediction and Machine Learning Models

Harki Himawan¹, Muhammad Dzakky Alghifari¹, Rut Juniar Nainggolan¹, Mochamad Bagus Hermanto¹, Sandra¹, Nazmi Mat Nawi², Ken Abamba Omwange³ and Dimas Firmanda Al Riza^{1*}

¹Department of Biosystems Engineering, Faculty of Agricultural Technology, Universitas Brawijaya, Indonesia

²Department of Biological and Agricultural Engineering, Universiti Putra Malaysia, Malaysia

³Department of Biological and Agricultural Engineering, 95616, University of California, Davis, CA, USA

Received: 21 May 2025, Revised: 9 February 2026, Accepted: 13 March 2026, Published: 13 July 2026

Abstract

Determining the right harvest time is crucial to ensure seed quality, with moisture content as the primary parameter. This study evaluated the ability of portable near-visible infrared (Vis-NIR) spectrometers (AS7265X, C12880MA, and AS7421) combined with partial least squares regression (PLSR) and artificial neural network (ANN) models to accurately predict corn moisture content. A total of 250 samples of corn variety NK 7207 were analyzed using spectral reflectance data processed with multiple scatter correction (MSC), standard normal variate (SNV), and wavelength selection. The C12880MA sensor provided the best results, with the PLSR model achieving a testing accuracy of $R^2 = 0.91$ and the ANN model achieving $R^2 = 0.94$. The results show that Vis-NIR portable spectrometer can accurately predict moisture content with high reliability as a non-destructive, low-cost, and efficient device. This technology can help farmers determine the optimal harvest time, improve seed quality, and support agricultural productivity.

Keywords: machine learning; moisture content; non-destructive; spectroscopy

1. Introduction

Corn is one of the most important agricultural commodities in Indonesia, which plays an important role as food, animal feed, and industrial raw materials for various products such as corn flour, cereals, and sweeteners (Fatemi et al., 2022). World corn production in 2022 reached 1.17 billion tons, with the United States being the largest corn producer with 345.75 million tons. In 2023, the production of dry shelled corn with a moisture content of 14% in Indonesia reached 14.77 million tons, a decrease of 10.61% compared to the 2022 production of 16.53 million tons (BPS-Statistics Indonesia, 2023). This data shows that

*Corresponding author: E-mail: dimasfirmanda@ub.ac.id
<https://doi.org/10.55003/cast.2026.267701>

Copyright © 2026 by King Mongkut's Institute of Technology Ladkrabang, Thailand. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

there are challenges in the corn farming sector in Indonesia, which can be caused by genetic factors such as the seeds used. The use of superior quality seeds can enhance resistance to diseases and pests, drought or excessive rainfall, and thus potentially increase yields. Determining the right harvest time is also a crucial factor in ensuring seed viability, optimal growth capacity, and good storability.

Harvesting at the right time is an important factor in ensuring high-quality seeds in seed production. If harvesting is done too early, the seeds may have not reached full maturity and tend to have low vigor. Conversely, late harvesting increases the risk of damage from insects and microorganisms that accelerate the process of seed deterioration. In the post-harvest process, delays can also increase the likelihood of damage due to environmental conditions. Therefore, determining the ideal maturity level and harvest time for hybrid maize is essential to produce seeds with optimal quality and vigor in hybrid corn seedling programs (Aswin et al., 2023).

Moisture content is one of the important parameters in determining the harvesting time of corn. Typically, corn cobs are harvested when the moisture content reaches 20-35%. Previously, the measurement of moisture content was done by laboratory methods, but these methods require taking samples from the farm to be analyzed in the laboratory and thus the samples can be damaged. Although there are commercial tools based on spectroscopic sensors available, these tools tend to be expensive for Indonesian farmers. Therefore, it is necessary to develop technology to measure corn moisture content in a non-destructive, real-time, and low-cost manner.

In recent years, researchers have attempted to predict moisture content in grain crops as a quality assessment using spectroscopy/optical based instruments combined with machine learning. Lin et al. (2019) showed that near-infrared (NIR) spectroscopy could be used to predict the moisture content of rice with an R^2 of 0.936. Studies using visible-NIR spectroscopy combined with PLSR modelling to predict corn moisture content in various corn varieties had better performance than that with Vis/NIR spectra (Zhang & Guo, 2020). Based on these facts, measurement methods with a spectroscopy approach show potential as non-destructive tools for measurement of moisture content.

Currently, the fabrication and use of spectroscopy-based sensors is being massively experimented in the agriculture and food industry. This provides the opportunity to develop highly portable and relatively inexpensive detection systems that have the capabilities of conventional spectrometers. In general, spectroscopy-based sensors have different resolutions and wavelength responses. This gap is very crucial because the performance of spectroscopy sensors can vary significantly based on factors such as sensor resolution, wavelength response, and environmental conditions. Thus, researchers need to evaluate the ability of sensors to be used as test instruments. Although many studies have proven the ability of spectroscopic techniques to detect corn moisture content, exploration of the use of portable spectrometer systems is still relatively low. Therefore, there is a need for a comparative study of spectroscopy-based sensors that can predict corn moisture content.

Despite its important role, challenges in the corn farming sector, such as declining production, indicate a need for innovation in agricultural technology, particularly in seed quality and harvest time. The application of spectroscopy sensors combined with machine learning techniques to predict corn moisture content appears to offer considerable benefits. This non-destructive approach has the potential to replace previous laboratory-based methods that are costly and time-consuming. Therefore, this research was aimed at evaluating the performance of spectroscopy-based sensors in predicting corn moisture content using partial least square regression (PLSR) and artificial neural network (ANN) models.

2. Materials and Methods

2.1 Instruments and materials

This research was conducted at the Mechatronics and BIO-AI Laboratory, Universitas Brawijaya, Indonesia. The sensors used were AS7265X, AS7421 Hyperspectral NIR Sensor, and C12880MA Mini spectrometer. The AS7265X and C12800MA sensors were integrated with electronic components such as ESP32 Thing Plus, 3.7V Li-Po battery, and TP223 touch sensor. The corn variety used was NK 7207 produced by Syngenta.

2.2 Research method

The implementation of the study began with individual corn cultivation aimed at preparing corn samples and assembling spectroscopy sensor-based instruments. This study used 250 samples of corn variety NK 7207. Subsequently, spectral and moisture content data were acquired, followed by data preprocessing, model development, and model performance evaluation.

The AS7265X sensor system consists of three sensors: AS72651, AS72652, and AS72653. Each of these three sensor devices is equipped with 6 optical filters on independent devices whose spectral response is fixed in the range from 410 nm to 940 nm with a FWHM of 20 nm and form the 18-channel multispectral sensor chipset AS7265X. Furthermore, the C12880MA mini-spectrometer has a spectral response of 312-882 nm with 228 pixels. These three sensors are integrated with a microcontroller programmed using Arduino IDE software with the C++ programming language and a Bluetooth module in order to read spectral data and send it to a PC. All the main components are assembled in a housing made of polylactic acid (PLA) printed with an ANET A8 3D printer. Furthermore, the AS7421 hyperspectral NIR sensor has a 64-channel response at 705-1050 nm spectrum. The schematic of spectral data acquisition with the instrument can be seen in Figure 1.

Corn reflectance-based spectral data acquisition was performed using the three sensors described. A total of 250 corn samples were obtained from individual cultivation. In this study, a spectra measurement approach was used on the corn husks. Approach through corn husks to develop a non-destructive, low-cost, and practical method of measuring moisture content, considering that corn husks are the outer part of the plant that does not require damage to the corn kernels themselves. The spectral data was obtained from three points of the corn husk: base, center, and tip. After the light source interacts with the corn husk, the diffused light intensity from the corn husk was then captured by the sensor and recorded as I_{raw} . Each sample spectra was calculated as an average of three scans to obtain a more accurate I_{raw} value, resulting in 250 spectra datasets. The spectra data from the reference white card was used to obtain the reference spectra, which were recorded as reference. Next, the relative reflectance (R) spectra were calculated using equation 1.

$$R = \frac{I_{raw}}{I_{reference}} \quad (1)$$

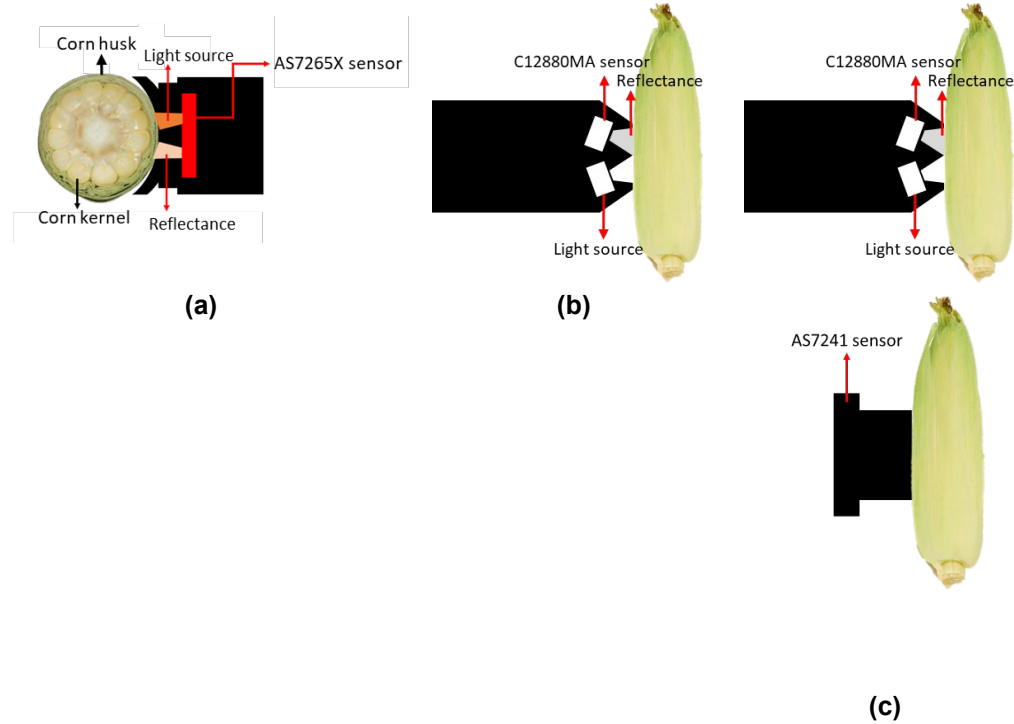


Figure 1. Schematic of spectral data acquisition using sensors (a) AS7265X, (b) C12880MA, and (c) AS7421

Data acquisition was performed in accordance with the actual conditions of farmers during harvest. After spectral data collection, a 5 g subsample from each corn sample was immediately weighed for the gravimetric moisture content determination using an analytical balance with a precision of 0.0001 g. Then, each seed sample was placed in an 8 cm diameter cup and dried for 4 h in a drying oven at 130°C in accordance with Indonesian National Standard 6944:2015. After removal from the oven, the dishes containing the samples were cooled in a desiccator for 30 minutes before being reweighed to determine moisture loss. The moisture content on a wet basis of each seed was calculated according to equation 2.

$$M = \frac{m_0 - m_1}{m_0 - m} \times 100\% \quad (2)$$

Where m_0 is the mass of each seed and cup before drying, m_1 is the mass of each seed and cup after drying, m is the mass of the cup and M is the moisture content of the seed on a wet basis.

The development of a prediction model using machine learning algorithms, PLSR and ANN, was carried out to predict the moisture content of corn. These algorithms were used to assess the consistency of the resulting model. Spectral data from corn husk was used as predictor data, while moisture content data was used as target data. A total of 250 datasets were analyzed using Python 3.10.12 in Google Collaboratory. The dataset was divided into 80% for training and 20% for testing. The initial dataset contained noise, so

several data pre-processing techniques were applied to ensure model accuracy. The first pre-processing step involved removing outliers from spectral data, as outliers can interfere with model performance. Additionally, two pre-processing techniques were applied to reduce the effects of scatter: multiplicative scatter correction (MSC), which adjusts the raw spectrum to the reference spectrum, and standard normal variate (SNV), which normalizes the spectrum by subtracting the mean and dividing it by standard deviation. These methods were used to reduce the variation in the spectrum caused by differences in the surface properties of corn kernels (Adnan et al., 2017).

To improve model performance, reduce complexity, and facilitate interpretation, a wavelength feature selection process was applied using principal component analysis (PCA). The aim of this process is to identify the most relevant and informative wavelengths from the entire dataset by reducing its dimensions while retaining the most significant features. PCA works by transforming the original spectral data into a set of uncorrelated variables called principal components, which capture the variance in the data. By selecting the principal components that explain the greatest variance, this process effectively reduces noise and focuses on the most critical features. The selection uses a threshold value based on the x-loadings from PCA, which represent the contribution of each wavelength to the principal components. While PLSR modeling was performed using the entire dataset after data pre-processing, ANN modeling was based on raw data, eliminating outliers, and data selected based on features generated from PCA. For the ANN model, the H2O library was used and employed the AutoML feature to automatically optimize hyperparameters, including the number of layers, number of nodes, activation function, and learning rate. The differences in these preprocessing steps have an impact on the performance of the models. PLSR, which uses dimension reduction through PCA, is superior for data with simple linear relationships. This process helps reduce complexity and eliminate noise in the data, allowing the model to be more easily understood. Meanwhile, ANN is capable of capturing more complex and non-linear relationships. Although ANN is capable of handling more complex data, this advantage comes with difficulties in interpreting the model, especially when it involves many layers and parameters. To evaluate the models, the following parameters were used: coefficient of determination (R^2), root mean square error (RMSE), and ratio of performance to deviation (RPD). In addition, to validate the robustness of the models, the k-fold cross-validation method was used, $k=10$ (Pal & Patel, 2020).

3. Results and Discussion

3.1 Corn moisture content measurement result

This study involved 250 samples of the corn variety NK 7207. The moisture content of the samples was measured using gravimetric methods between October 22 and November 27, 2023, when the corn was 3-4 months old. During this period, the moisture content of the corn dropped from 85.66% to 28%, reaching the optimal level for seed harvesting. It should be noted that this study only used samples of the corn variety NK 7207. The box plot of moisture content can be seen in Figure 2.

Changes in moisture content in corn are influenced by several factors, including the environment, the drying process, and the intrinsic properties of the corn itself. Environmental factors such as temperature and air humidity play an important role in the evaporation of moisture from corn kernels. Higher temperatures can increase the rate of water evaporation by providing additional energy to water molecules, thereby accelerating

the evaporation process from the surface of the corn. On the other hand, high air humidity acts as an inhibitor of evaporation, because saturated air cannot absorb more water vapor, thereby reducing the efficiency of evaporation. The drying process using natural methods such as sun drying or mechanical drying, also greatly affects the final moisture content of the corn. Freshly harvested corn has a relatively high moisture content, around 20-30%, which needs to be reduced to around 12-14% to ensure optimal quality and storability (Nurba et al., 2021).

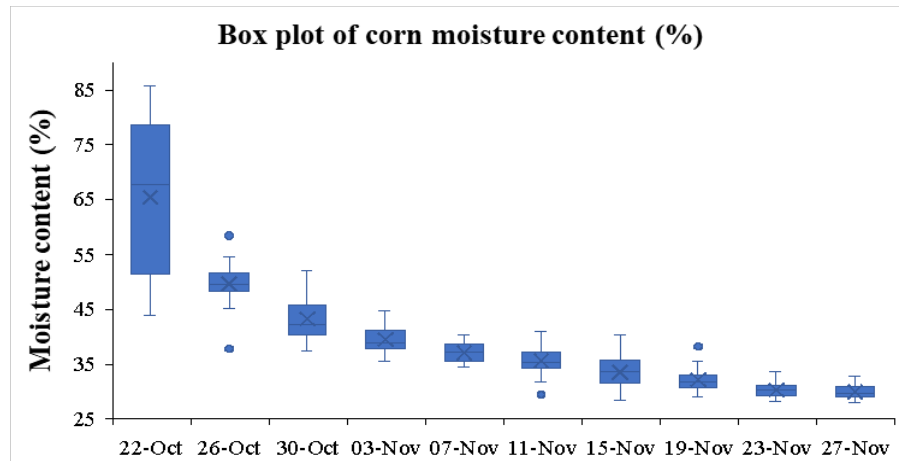


Figure 2. Box plot of moisture content

3.2 Spectra data comparison

The AS7265X sensor generated 18 spectral variables in the 410-940 nm wavelength range. The C12880MA sensor generated 288 spectral variables in the 312-882 nm wavelength range, and the AS7421 sensor generated 64 spectral variables in the 750-1050 nm wavelength range. The spectral data generated from the instruments is the reflectance spectrum of the corn husk. In this study, visible spectroscopy can be used to analyze color changes in corn husks. The color change that occurs depends on the wavelength of light reflected by the corn husk. Visible spectroscopy measures how much light at a particular wavelength is absorbed by a sample. The absorbance or reflectance of this light is directly related to the electron structure in the molecules and types of bonds present in the corn kernel (Karoui, 2018).

The color of corn kernels is produced by natural pigments such as carotenoids. As moisture content decreases, these pigments may undergo changes in terms of concentration, degradation, or oxidation. Research shows that the AS7265X and C12880MA sensors absorb light at 400-500 nm which indicates the absorption of carotenoid pigments (Ashenafi et al., 2023). Figure 3c shows the absorbance intensity of the sample at a wavelength of 970 nm. Heman and Hsieh (2016) explained that the absorbance at around 970 nm is due to the combined effect of the second O-H overtone found in carbohydrates and water. The absorbance at about 1200 nm is associated with the C-H stretch of the second overtone of carbohydrates, while the absorbance at about 1450 nm is mainly due to the water absorption band associated with the first O-H overtone stretch. The reflectance spectra can be seen in Figure 3.

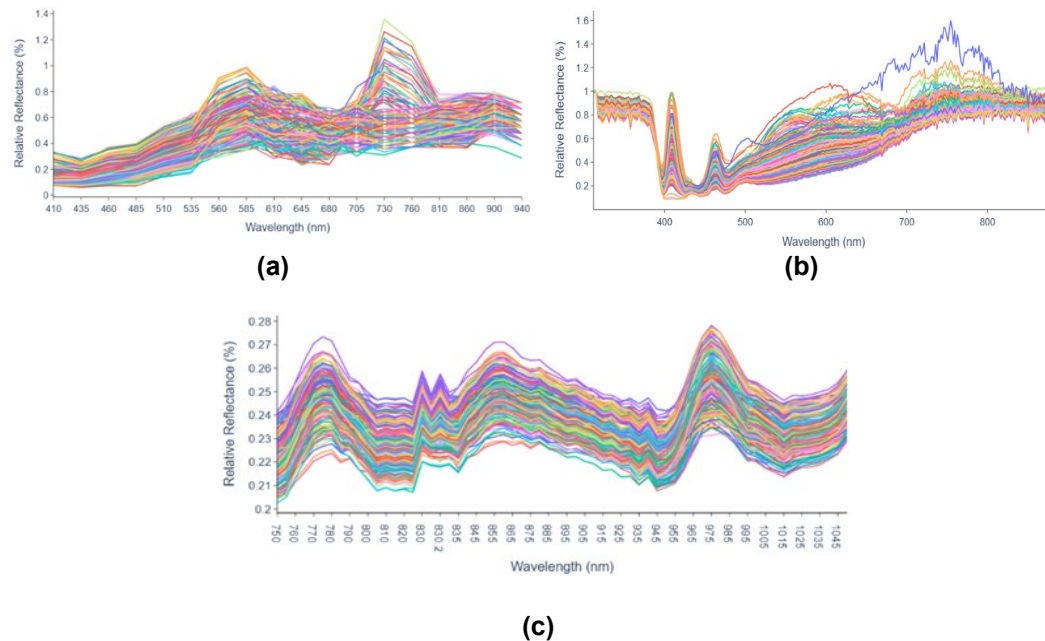


Figure 3. Reflectance spectra of corn husks, (a) AS7265X; (b) C12880MA; and (c) AS7421

3.3 Outlier detection results using principal component analysis (PCA)

Outlier detection is the process of identifying data that has patterns or values that are significantly different from the majority of data in a dataset. In the context of corn spectrum data, outliers can indicate measurement errors or significant differences that need to be analyzed further. In this study, PCA was used to reduce the dimensionality of the data while looking for unusual patterns. This technique finds linear combinations of the original variables called principal components, which represent the greatest variance in the data (Lepot et al., 2017). A cumulative explained variance threshold of >80% was applied, which in this case was achieved with three principal components (PCs) (Liu et al., 2018). Data located far from PC1, PC2, and PC3 in the projection can be identified as outliers, which can be seen in Figure 4. In this technique, data located far from the center of distribution on the first (PC1), second (PC2), and third (PC3) principal components are considered outliers. Data that is far from these three principal components may indicate inconsistencies or even errors in measurement. Outliers detected by PCA are at points that are significantly distant from the center of the main component distribution. Data that is too far from the average value of these main components can affect the model being built, making the model too sensitive to this data and potentially causing overfitting, where the model only works well on training data but cannot generalize well when applied to new data (Smiti, 2020).

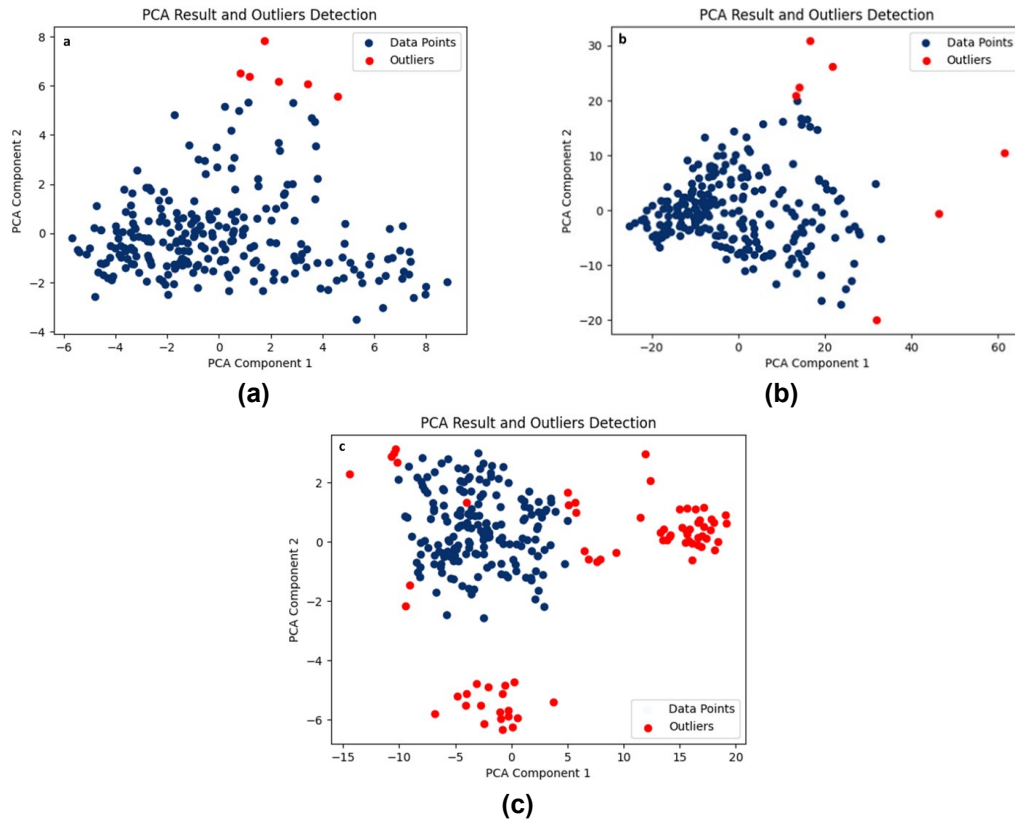


Figure 4. Outlier detection results using PCA of sensors (a) AS7265X, (b) C12880MA, and (c) AS7421

3.4 Wavelength feature selection result

The selection of wavelength features from each sensor that are sensitive to the target data (moisture content) uses the X-Loadings technique in PCA. These principal components represent the largest variance in the data. X-loadings show how much each wavelength contributes to the variance explained by the principal component. The larger the loading value, the greater the influence of that wavelength on the information contained by the principal component (Mehmood et al., 2012). The selection of wavelengths was based on the threshold loading score. Table 1 shows the results of the wavelength feature selection from each sensor.

Table 1. Wavelength feature selection result

Sensor Type	Selected Wavelength (nm)
AS7265X	510, 560, 585, 705, 730, 760
C12880MA	381, 384, 386, 389, 391, 394, 397, 399, 402, 404, 407, 409, 412, 415, 417, 420, 422, 425, 427, 430, 432, 435, 460, 463, 465, 468, 470, 701, 703, 709, 710, 718, 723
AS7421	950, 955, 960, 965, 970, 975, 980, 985

The results of the selection of wavelength features are in accordance with previous research (Charilaou & Battat, 2022), which report water absorption bands at approximately 760, 970, 1450, and 1940 nm, corresponding to O–H overtone and combination vibrations. The absorption near 970-975 nm is associated with the second overtone of the O–H stretching vibration, typically arising from free or weakly hydrogen-bonded O–H groups, resulting in a relatively stronger absorption at lower wavelengths. In contrast, absorption bands at higher wavelengths (e.g., 1450 and 1940 nm) are attributed to the first overtone and combination bands of O–H stretching, which are influenced by stronger hydrogen bonding interactions, such as those present in cellulose, hemicellulose, and lignin (Sundaram et al., 2015).

3.5 Comparison of PLSR prediction models of corn moisture content

In this research, in addition to using raw data, data pre-processing techniques were also used to improve the reflectance signal. The techniques used were outlier removal, multiplicative scatter correction (MSC), standard normal variate (SNV), and wavelength feature selection. Table 2 shows the results of the PLSR models generated from the three sensors. The best model for the AS7265X sensor was generated using MSC pre-processed data, with R^2 training and R^2 testing of 0.86. From the RPD evaluation metric, 2.67 was obtained, confirming the good accuracy of the model. Furthermore, the best model for the C12880MA sensor was obtained from the SNV pre-processed data, which yielded R^2 training and R^2 testing values of 0.91. Moreover, the resulting RPD value was more than 3, indicating that the model built had good accuracy and the potential to be applied in predicting corn moisture content (Abu-Khalaf & Hmidat, 2020). For the AS7421 sensor, the best model was obtained from the SNV pre-processed data where R^2 training = 0.89, R^2 testing = 0.88, and RPD = 2.90. Of the three models obtained, the results of signal reflectance correction using data pre-processing techniques such as MSC and SNV significantly improved model accuracy (Lazim et al., 2016). Scatter plot of measured versus predicted moisture content for the best model can be seen in Figure 5.

The C12880MA sensor demonstrated the best overall model performance compared to the other sensors, as shown in Table 2. This is evidenced by its consistently higher coefficient of determination (R^2) and lower prediction error. In particular, the SNV-preprocessed model achieved the highest R^2 values for both training (0.91) and testing (0.91), along with the lowest RMSE for testing (3.50%). Moreover, the RPD value reached 3.25, indicating excellent predictive capability and practical applicability. This superior performance can be attributed to the higher spectral resolution and sensitivity of the C12880MA sensor (Sonobe et al., 2022). Several studies have explored this sensor to evaluate the quality of bioproducts (Ulinuhta et al., 2021). In addition, this sensor is easily integrated with several electronic devices that support software-based usage (Sokołowski et al., 2021).

Based on the PLSR model generated from these three sensors, the Vis-NIR portable spectrometer demonstrated strong potential to be developed into an alternative method as a device to predict corn moisture content. The device should enable farmers to easily determine the right harvest time. This conclusion is supported by various studies that used NIR spectroscopy techniques as a non-destructive, easy, fast, low-cost, and accurate measurement techniques (Adnan et al., 2017; Makky et al., 2019).

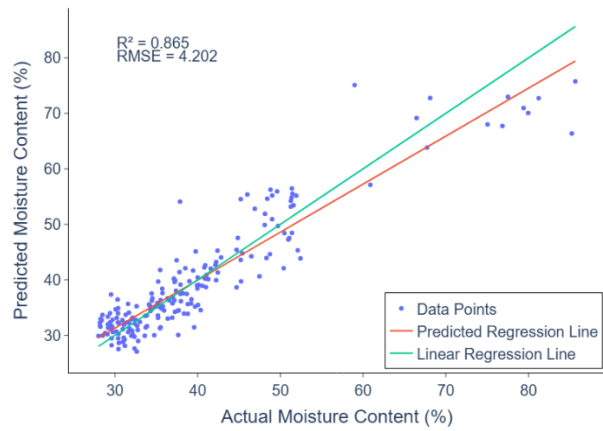
Table 2. Comparison of PLSR prediction model results of corn moisture content for each sensor

Type of Sensors	Dataset	R ² Training	R ² Testing	RMSE	RMSE	RPD Testing
				Training (%)	Testing (%)	
AS7265X	<i>Raw data</i>	0.85	0.82	4.46	4.66	2.37
	<i>Remove outlier</i>	0.83	0.83	4.17	4.26	2.42
	MSC	0.86	0.86	4.20	4.40	2.67
	SNV	0.86	0.84	4.20	4.75	2.50
	<i>Selected feature</i>	0.82	0.82	4.86	4.89	2.33
C12880MA	<i>Raw data</i>	0.90	0.85	3.56	4.60	2.58
	<i>Remove outlier</i>	0.90	0.84	3.64	4.57	2.51
	MSC	0.92	0.89	3.21	3.79	3.01
	SNV	0.91	0.91	3.42	3.50	3.25
	<i>Selected feature</i>	0.65	0.63	6.53	7.95	1.65
AS7421	<i>Raw data</i>	0.86	0.84	4.32	4.63	2.46
	<i>Remove outlier</i>	0.89	0.84	4.10	4.47	2.49
	MSC	0.88	0.87	3.90	4.34	2.74
	SNV	0.89	0.88	3.80	4.22	2.90
	<i>Selected feature</i>	0.68	0.67	6.53	6.42	1.74

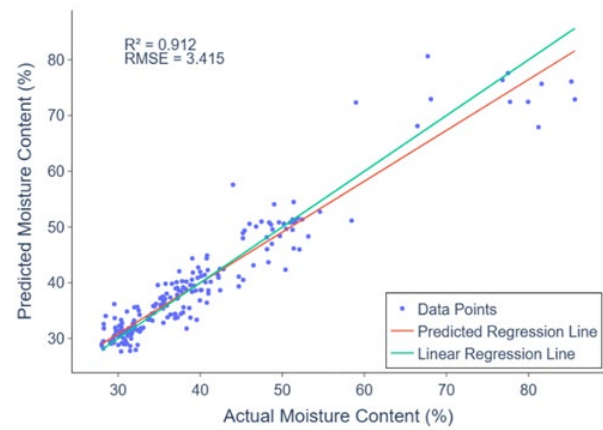
Note: *MSC: Multiplicative scatter correction, SNV: Standard normal variate, R²: Coefficient of determination, RMSE: Root mean square error: RPD: Ratio of performance to deviation

3.6 Comparison of ANN prediction models of corn moisture content

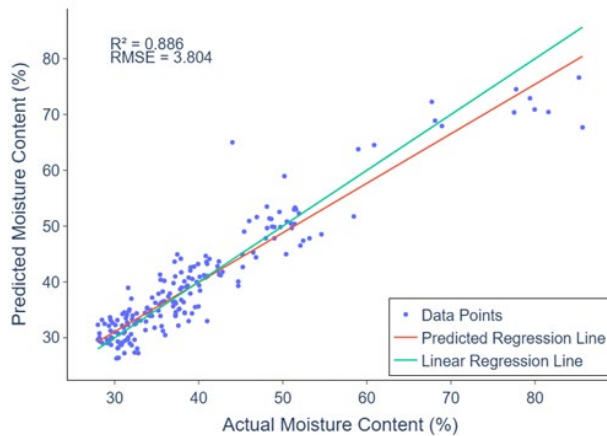
In addition to developing regression models such as the PLSR model, this research also used machine learning artificial neural network (ANN) models to see the consistency of the resulting models. The development of the ANN model involved the use of the H2O library, which facilitated an automate search to determine the optimal number of hidden layers and numbers of neuron in each layer. The AutoML features in H2O was capable of determining the most effective architecture by testing various configurations of neurons and layers based on the training data. This feature helps streamline the process of model optimization by evaluating numerous possible architectures and selecting the one that yields the best performance in terms of metrics like R² and RMSE.



(a)



(b)



(c)

Figure 5. Scatter plots of measured versus predicted moisture content for the best PLSR models; (a) AS7265X, (b) C12880MA, and (c) AS7421 sensors

The input data for the ANN structure consisted of spectral measurement data and the results of data pre-processing, while the output data was the value of corn moisture content. Table 3 shows the prediction models obtained using the ANN machine learning model. The best ANN model for the AS7265X sensor was generated from raw data and used a single hidden layer with 100 neurons (ANN structure 100), where R^2 training = 0.90, R^2 testing = 0.93. The best model generated from the C12880MA sensor was obtained using raw data, using two hidden layers with 100 neurons in each layer (ANN structure 100-100), giving R^2 training value = 0.93 and R^2 testing = 0.94. The AS7421 sensor produced the best model using raw data with an ANN structure of two hidden layers, each with 100 neurons, yielding R^2 training value = 0.80 and R^2 testing = 0.90. Scatter plot of measured versus predicted moisture content from the best model can be seen in Figure 6.

Table 3. Comparison of ANN prediction model results of corn moisture content for each sensor

Type of sensors	Dataset	ANN Structure	R ² Training	R ² Testing	RMSE Training (%)	RMSE Testing (%)
AS7265X	Raw data	100	0.9	0.93	3.84	2.52
	Remove outlier	100	0.92	0.7	3.02	3.2
	Selected feature	100-100	0.86	0.82	4.45	4.02
C12880MA	Raw data	100-100	0.93	0.94	3.28	2.38
	Remove outlier	50	0.8	0.71	5.36	4.67
	Selected feature	20-20	0.66	0.41	6.91	7.28
AS7421	Raw data	100-100	0.8	0.9	5.4	2.87
	Remove outlier	100-100	0.77	0.91	5.86	2.73
	Selected feature	20-20-20	0.5	0.32	8.47	7.8

Note: * R^2 : Coefficient of determination, RMSE: Root mean square error, RPD: Ratio of performance to deviation

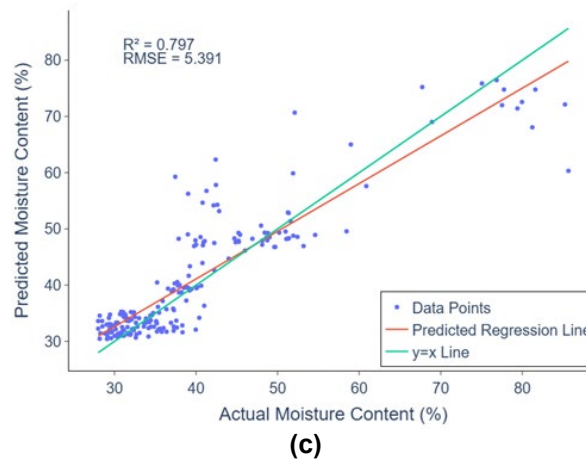
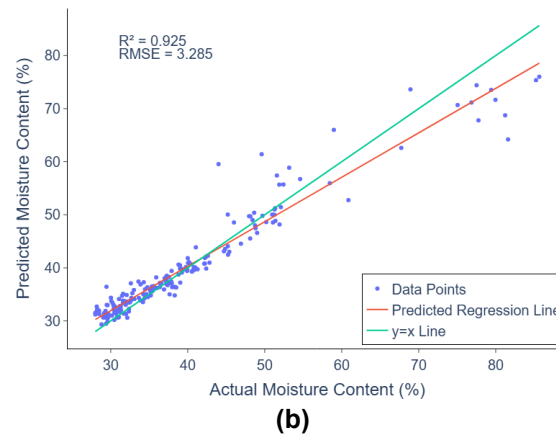
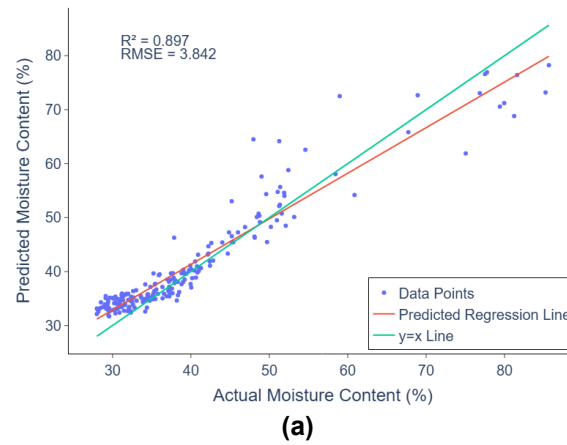


Figure 6. Scatter plots of measured versus predicted moisture content from the best ANN models; (a) AS7265X, (b) C12880MA, and (c) AS7421 sensors

The ANN structure used in this study involved various configurations of the number of hidden layers and the number of nodes in each layer. The model generated from the AS7265X sensor has one layer consisting of 100 nodes, while the C12880MA and AS7421 sensors had two layers consisting of 100 nodes in each layer. The use of two hidden layers with a relatively large number of nodes allows the model to capture the complexity of the data better and improves prediction accuracy, although this can increase the risk of overfitting if the training data is limited (Uzair & Jamil, 2020). The C12880MA sensor demonstrated good accuracy in the ANN model due to the combination of the two hidden-layer structure (100-100) and the sensor's ability to detect spectral data relevant to corn moisture content (Sandak et al., 2020). Although the ANN model generated from the AS7421 sensor uses two hidden layers, the accuracy of this model was lower than that of a single-layer model. This is due to overfitting caused by increased model complexity. The additional layer can cause the model to memorize the training data rather than effectively generalize unseen data, especially if the training dataset is limited. This can lead to lower performance on the test set, which is reflected in lower accuracy (Charilaou & Battat, 2022). These outlier data points can be caused by various factors, such as sensor noise, environmental conditions, or inherent limitations of the sensor in capturing certain characteristics of corn moisture content. The presence of outliers can significantly affect model performance and can skew the learning process, resulting in less accurate predictions (Smiti, 2020).

Based on the ANN model results obtained, these sensors show good ability to predict corn moisture content. This can be seen from the resulting PLSR and ANN models that showed an accuracy rate of 80-90%. Along with the development of portable spectrometer instruments that have been commercialized and supported by computational techniques, it clearly shows that these portable spectrometers can replace laboratory spectrophotometers in various measurements. This has advantages in the analysis of various samples.

3.7 Practical considerations for sensor deployment

In addition to predictive performance discussed in the previous sections, the real-world applicability of spectrometer sensors in agricultural environments depends on several practical considerations, including cost, portability, maintenance requirements, and ease of integration. The AS7265X, a multispectral sensor array, offers broad spectral coverage within a compact architecture that can be readily embedded into portable devices. Its relatively low cost and factory calibration reduce system complexity and minimize maintenance requirements, making it particularly suitable for cost-efficient, large-scale field monitoring applications such as plant health assessment and nutrient status estimation (Lopin et al., 2025). The C12880MA (Hamamatsu) represents a miniaturized spectrometer with high sensitivity across the 340-850 nm range. Its hermetically sealed design enhances durability under varying environmental conditions, including humidity, thereby supporting reliable in-field operation. Although its cost is generally higher than that of basic multispectral sensors, its compact form factor and robustness make it well-suited for portable analytical devices used in agricultural and food quality assessment. Maintenance requirements are relatively minimal and primarily involve ensuring stable illumination conditions and maintaining optical cleanliness. The AS7421 sensor, which operates in the near-infrared range (750-1050 nm) with a 64-channel photodiode array, provides higher spectral resolution for applications such as moisture content estimation and compositional analysis in crops. Its low power consumption and small footprint facilitate integration into

embedded sensing systems. However, the increased spectral complexity necessitates more advanced calibration strategies and data processing techniques, which may increase implementation effort compared to simpler sensor configurations. Overall, each sensor demonstrates distinct trade-offs between cost, spectral capability, portability, and operational complexity. These factors should be carefully considered when selecting appropriate sensing technologies for practical deployment in precision agriculture systems and other field-based applications.

4. Conclusions

This research shows that NIR spectroscopy-based sensors have significant potential as a non-destructive, rapid, and cost-effective method of predicting moisture content in corn. The study employed two modeling techniques, PLSR and ANN, to develop prediction models based on the spectral data obtained from the sensors. The results revealed that prediction models developed using partial least square regression (PLSR) and artificial neural network (ANN) achieved high accuracy with R^2 values above 0.8 to 0.9. Among the sensors tested, the C12880MA sensor provided the best performance, owing to its high resolution and improved sensitivity. Data preprocessing techniques, including MSC and SNV, were applied to enhance the accuracy of the models by reducing noise and correcting for variations in the data. The results confirm that Vis-NIR spectroscopy, combined with advanced data processing and modeling techniques, can serve as an effective alternative method for predicting corn moisture content. This approach supports farmers in making informed decisions regarding the optimal time for harvesting, ultimately contributing to more efficient and sustainable agricultural practices.

5. Acknowledgements

This work was funded by the Ministry of Education, Culture, Research, and Technology (MoECRT) of Indonesia, through the grant scheme of Penelitian Fundamental - Regular 2024 with grant master contract number 045/E5/PG/02.00.PL/2024 and derivative contract number 00309.45/UN10.A0501/B/PT.01.03.2/2024.

6. Authors' Contributions

Harki Himawan conducted measurements; analyzed data; wrote the original paper. Muhammad Dzakky Alghifari performed experiment; contributed analytic tools; wrote the paper. Rut Juniar Nainggolan provide data visualization; wrote the paper. Mochamad Bagus Hermanto contributed analytic tools, coordinated research. Sandra conducted research supervision; contributed to methodology. Nazmi Mat Nawi conducted supervision; wrote the paper. Ken Abamba Omwange conducted supervision; wrote the paper. Dimas Firmanda Al Riza designed research methodology; performed supervision; contributed analytic tools; analyzed data; coordinated research; wrote the paper.

ORCID

Mochamad Bagus Hermanto  <https://orcid.org/0000-0002-9366-0184>

Sandra  <https://orcid.org/0000-0001-9600-9349>

Nazmi Mat Nawi  <https://orcid.org/0000-0002-5916-5745>

Ken Abamba Omwange  <https://orcid.org/0000-0002-8268-1940>

Dimas Firmanda Al Riza  <https://orcid.org/0000-0001-6790-6729>

7. Conflicts of Interest

The authors declare no conflict of interest.

References

- Abu-Khalaf, N., & Hmidat, M. (2020). Visible/near infrared (VIS/NIR) spectroscopy as an optical sensor for evaluating olive oil quality. *Computers and Electronics in Agriculture*, 173, Article 105445. <https://doi.org/10.1016/j.compag.2020.105445>
- Adnan, A., von Hörsten, D., Pawelzik, E., & Mörlin, D. (2017). Rapid prediction of moisture content in intact green coffee beans using near infrared spectroscopy. *Foods*, 6(5), Article 38. <https://doi.org/10.3390/foods6050038>
- Ashenafi, E. L., Nyman, M. C., Shelley, J. T., & Mattson, N. S. (2023). Spectral properties and stability of selected carotenoid and chlorophyll compounds in different solvent systems. *Food Chemistry Advances*, 2, Article 100178. <https://doi.org/10.1016/j.focha.2022.100178>
- Aswin, C., Geetha, R., Sujatha, K., Vanniarajan, C., & Sivakumar, T. (2023). Effect of harvesting time on the seed quality of hybrid maize (*Zea mays* L.) COHM 8. *Agricultural Science Digest*, 43(5), 622-629. <https://doi.org/10.18805/ag.D-5748>
- BPS-Statistics Indonesia. (2023). *Corn and soybean production in Indonesia 2023*. BPS-Statistics Indonesia.
- Charilaou, P., & Battat, R. (2022). Machine learning models and over-fitting considerations. *World Journal of Gastroenterology*, 28(5), 605-607. <https://doi.org/10.3748/wjg.v28.i5.605>
- Fatemi, A., Singh, V., & Kamruzzaman, M. (2022). Identification of informative spectral ranges for predicting major chemical constituents in corn using NIR spectroscopy. *Food Chemistry*, 383, Article 132442. <https://doi.org/10.1016/j.foodchem.2022.132442>
- Heman, A., & Hsieh, C. L. (2016). Measurement of moisture content for rough rice by visible and near-infrared (NIR) spectroscopy. *Engineering in Agriculture, Environment and Food*, 9(3), 280-290. <https://doi.org/10.1016/j.eaef.2016.02.002>
- Karoui, R. (2018). Spectroscopic technique: Fluorescence and ultraviolet-visible (UV-Vis) spectroscopies. In D.-W. Sun (Ed.). *Modern techniques for food authentication* (2nd ed., pp. 219-252). Elsevier Inc. <https://doi.org/10.1016/b978-0-12-814264-6.00007-4>
- Lazim, S. S. R. M., Naw, N. M., Chen, G., Jensen, T., & Rasli, A. M. M. (2016). Influence of different pre-processing methods in predicting sugarcane quality from near-infrared (NIR) spectral data. *International Food Research Journal*, 23, S231-S236.
- Lepot, M., Aubin, J. B., Clemens, F. H. L. R., & Mašić, A. (2017). Outlier detection in UV/Vis spectrophotometric data. *Urban Water Journal*, 14(9), 908-921. <https://doi.org/10.1080/1573062X.2017.1280515>
- Lin, L., He, Y., Xiao, Z., Zhao, K., Dong, T., & Nie, P. (2019). Rapid-detection sensor for rice grain moisture based on NIR spectroscopy. *Applied Sciences*, 9(8), Article 1654. <https://doi.org/10.3390/app9081654>
- Liu, Y. J., Tran, T., Postma, G., Buydens, L. M. C., & Jansen, J. (2018). Estimating the number of components and detecting outliers using angle distribution of loading subspaces (ADLS) in PCA analysis. *Analytica Chimica Acta*, 1020, 17-29. <https://doi.org/10.1016/j.aca.2018.03.044>

- Lopin, P., Nawsang, P., Laywisadkul, S., & Lopin, K. V. (2025). Evaluation of low-cost multi-spectral sensors for measuring chlorophyll levels across diverse leaf types. *Sensors*, 25(7), Article 2198. <https://doi.org/10.3390/s25072198>
- Makky, M., Santosa, Putri, R. E., & Nakano, K. (2019). Determination of moisture content in rice using non-destructive short-wave near infrared spectroscopy. *AIP Conference Proceedings*, 2155, Article 5125518. <https://doi.org/10.1063/1.5125518>
- Mehmood, T., Liland, K. H., Snipen, L., & Sæbø, S. (2012). A review of variable selection methods in partial least squares regression. *Chemometrics and Intelligent Laboratory Systems*, 118, 62-69. <https://doi.org/10.1016/j.chemolab.2012.07.010>
- Nurba, D., Agustina, R., & Zahara, M. (2021). Study of changes in physical and chemical properties of corn during drying and storage process using a modified in-store dryer. *IOP Conference Series: Earth and Environmental Science*, 644(1), Article 012016. <https://doi.org/10.1088/1755-1315/644/1/012016>
- Pal, K., & Patel, B. V. (2020). Data classification with k-fold cross validation and holdout accuracy estimation methods with 5 different machine learning techniques. *Proceedings of the 4th international conference on computing methodologies and communication, ICCMC 2020*, 83-87. <https://doi.org/10.1109/ICCMC48092.2020.ICCMC-00016>
- Sandak, J., Sandak, A., Zitek, A., Hintestoisser, B., & Picchi, G. (2020). Development of low-cost portable spectrometers for detection of wood defects. *Sensors*, 20(2), Article 545. <https://doi.org/10.3390/s20020545>
- Smiti, A. (2020). A critical overview of outlier detection methods. *Computer Science Review*, 38, Article 100306. <https://doi.org/10.1016/j.cosrev.2020.100306>
- Sokołowski, P., Wróbel, M. S., & Karpieńko, K. (2021). Design and implementation of the driver system for a Hamamatsu C12880MA microspectrometer. *Photonics Letters of Poland*, 13(2), 34-36. <https://doi.org/10.4302/plp.v13i2.1103>
- Sonobe, R., Yamashita, H., Nofrizal, A. Y., Seki, H., Morita, A., & Ikka, T. (2022). Use of spectral reflectance from a compact spectrometer to assess chlorophyll content in *Zizania latifolia*. *Geocarto International*, 37(18), 5363-5375. <https://doi.org/10.1080/10106049.2021.1914747>
- Sundaram, J., Mani, S., Kandala, C. V. K., & Holser, R. A. (2015). Application of NIR reflectance spectroscopy on rapid determination of moisture content of wood pellets. *American Journal of Analytical Chemistry*, 6(12), 923-932. <https://doi.org/10.4236/ajac.2015.612088>
- Ulinuha, A. R. I., Bahtiar, Z. A., Nauri, A. R., Rhamadan, I., Wulansari, R. C., & Iqbal, Z. (2021). The performance of the C12880MA MEMS sensor for classification of the roasting level of coffee bean. *IOP Conference Series: Earth and Environmental Science*, 924, Article 012021. <https://doi.org/10.1088/1755-1315/924/1/012021>
- Uzair, M., & Jamil, N. (2020). Effects of hidden layers on the efficiency of neural networks. *2020 IEEE 23rd international multi-topic conference (INMIC)* (pp. 1-6). IEEE. <https://doi.org/10.1109/INMIC50486.2020.9318195>
- Zhang, Y., & Guo, W. (2020). Moisture content detection of maize seed based on visible/near-infrared and near-infrared hyperspectral imaging technology. *International Journal of Food Science and Technology*, 55(2), 631-640. <https://doi.org/10.1111/ijfs.14317>