International Journal of Basic and Applied Science 15 (1) (2026) 01-20

Published by: IOCSCIENCE

International Journal of Basic and Applied Science

Journal homepage: www.ijobas.pelnus.ac.id

Metaheuristic optimized fuzzy ensemble for maize seed quality prediction using Vis/NIR Spectroscopy

Ridwan Raafi'udin [1], Ali Khumaidi [2], Indra Permana Solihin [1], Erik Mulyana [3]

[1] Faculty of Computer Science, Universitas Pembangunan Nasional Veteran Jakarta, 12450, Indonesia

[2]Doctoral Program in Informatics, Universitas Nusa Mandiri, Jakarta, 13620, Indonesia

[3] Department of Agronomy and Horticulture, Faculty of Agriculture, Universitas IPB, Bogor, 16680, Indonesia

Article Info

Article history:

Received Mar 03, 2026
Revised Apr 10, 2026
Accepted Jun 23, 2026

Keywords:

Fuzzy Ensemble Learning;
Maize Seed Quality Prediction;
Metaheuristic Feature Selection;
Spectral Transformation;
Vis/NIR Spectroscopy.

ABSTRACT

Maize (Zea mays) seed quality assessment is essential for supporting agricultural productivity and sustainable seed management. This study proposes a non-destructive machine learning framework for predicting maize seed quality using portable Visible/Near-Infrared (Vis/NIR) spectroscopy. The framework integrates NIPPY-based spectral preprocessing, metaheuristic wavelength selection, and fuzzy ensemble learning to handle spectral noise, multicollinearity, and nonlinear relationships in small-sample spectral data. Informative wavelengths were selected using Genetic Algorithm (GA) and Particle Swarm Optimization (PSO). Two fuzzy ensemble models were developed: a Fuzzy Residual-Corrected Ensemble that refines predictions through residual-based fuzzy correction, and an RF+XGB Fuzzy Ensemble that combines Random Forest and XGBoost outputs using confidence-based fuzzy weighting. The models were evaluated for Moisture Content (MC), Germination Rate (GR), and Electrical Conductivity (EC) using repeated cross-validation, variability measures, and statistical validation. The proposed fuzzy ensemble models achieved R^2 values ranging from 0.8249 to 0.8689 and showed performance comparable to the strongest Random Forest baseline. Statistical comparison indicated that the main contribution of the fuzzy ensemble framework lies not in large gains in mean accuracy, but in prediction stability, residual correction, and uncertainty-aware modeling. SHAP-based explainability further identified physiologically meaningful wavelength regions, including visible pigment-related bands and near-infrared moisture-related bands. The dataset consists of 800 maize seed samples from four varieties under laboratory conditions, which limits generalization to field environments. Future work will focus on multi-location validation, domain adaptation, and real-time implementation. Overall, the proposed framework provides a statistically validated and interpretable approach for portable Vis/NIR-based maize seed quality prediction.

Corresponding Author:

Ridwan Raafi'udin,
Faculty of Computer Science,
Universitas Pembangunan Nasional Veteran Jakarta,
Jl. R.S Fatmawati No. 1, Cilandak, Jakarta Selatan, 12450, Indonesia
Email: raafiudin@upnvj.ac.id

1. INTRODUCTION

Maize (Zea mays) is a strategic agricultural commodity that supports global food security [1]. High-quality maize seeds are a crucial factor in determining plant productivity and successful harvests [2].

The quality of maize seeds is influenced by various physiological factors such as viability, vigor, and germination rate, which determine the seeds' ability to develop into healthy and productive plants [3]. Therefore, evaluating seed quality at the early germination stage is essential to ensure optimal agricultural outcomes. Conventional seed quality tests are typically performed using destructive methods such as germination tests, staining with tetrazolium chloride (TTC), and moisture content measurements [4]. However, these methods have significant limitations in terms of analysis time, their destructive nature, and their incompatibility with modern precision agriculture systems [5].

Near-Infrared (NIR) spectroscopy has emerged as a promising non-destructive solution with high accuracy [6], [7], [8]. However, its implementation in developing countries is hindered by the high cost of equipment and operational complexity. Visible/Near Infrared (Vis/NIR) spectroscopy offers a more economical alternative with sufficient analytical capabilities [9], [10]. Despite this, the application of Vis/NIR faces technical challenges, including spectral noise, scattering effects, and sample variability, which impact the accuracy of predictive models. These challenges are particularly important in portable spectroscopy systems, where limited spectral resolution, environmental variation, and small-sample conditions may increase uncertainty in model prediction. Although several preprocessing and feature selection strategies have been proposed, there is still no universally optimal strategy for handling spectral noise and wavelength redundancy across different agricultural targets [11]. From a signal processing perspective, Vis/NIR data is characterized by high multicollinearity, overlapping absorption bands, and low signal-to-noise ratios, which significantly reduce model generalization capability, particularly in small-sample agricultural datasets [12], [13]. Some previous studies have addressed these problems through spectral transformations and feature selection. Spectral transformation techniques have been shown to improve model performance by 11-14% [14]. However, optimization of the transformation process and feature selection remains an underexplored area, particularly in terms of integration with advanced machine learning algorithms. Therefore, Vis/NIR-based seed quality prediction requires not only accurate predictive models, but also validation strategies that explicitly report uncertainty, stability, and reproducibility.

In the domain of machine learning, ensemble approaches have shown strong potential in handling the complexity of spectral data [15]. However, conventional ensemble methods such as stacking and voting tend to be deterministic and are less effective in handling uncertainty [16]. Moreover, most classical ensemble learning approaches assume stationary data distributions and do not explicitly model epistemic uncertainty, which is critical in noisy spectral environments. Fuzzy logic offers a robust mathematical framework for modeling uncertainty through the concept of fuzzy sets and membership functions [17]. In the context of Vis/NIR spectroscopy, this capability is relevant because spectral responses may vary due to seed heterogeneity, measurement noise, and overlapping biochemical absorption patterns.

Several studies have previously explored the integration of fuzzy logic with machine learning. Khumaidi et al. [18] successfully applied fuzzy logic for classifying mango ripeness using NIR data, but it was limited to a single-model approach. Ouifak and Idri [19] reviewed adaptive neuro-fuzzy inference systems and evolving neuro-fuzzy systems, providing the best balance between performance and interpretability; however, they did not address computational complexity and scalability issues. A study by Shi et al. [20] proposed a robust fuzzy ensemble framework based on spectral learning, but it focused more on clustering rather than prediction. In addition, explainable machine learning methods such as SHAP have increasingly been used to interpret model behavior by estimating feature-level contributions, yet their application for wavelength-level interpretation in portable Vis/NIR seed quality prediction remains limited [21], [22].

Despite these advancements, existing literature still exhibits several critical limitations. First, most studies do not integrate spectral preprocessing, feature selection, and ensemble learning within a unified, leakage-controlled, and statistically validated evaluation framework, leading to fragmented and potentially non-reproducible modeling pipelines. Second, fuzzy-based ensemble methods are often implemented as post-hoc combination strategies rather than uncertainty-aware adaptive systems that respond to local spectral variability. Third, many existing studies lack rigorous statistical

validation and stability analysis under small-sample spectral conditions, which raises concerns regarding model generalization reliability [23], [24]. Fourth, model interpretation in Vis/NIR studies is often limited to reporting selected wavelengths or feature importance values without linking them sufficiently to physiologically meaningful spectral regions such as pigment-related visible bands and moisture-related near-infrared bands [25].

Based on the literature review, this study identifies three interrelated research gaps: (1) the absence of an end-to-end spectral learning pipeline that jointly integrates preprocessing, metaheuristic feature selection, and ensemble modeling under a leakage-controlled validation strategy, (2) the limited development of uncertainty-aware fuzzy ensemble frameworks that explicitly adapt model contribution based on local spectral confidence, and (3) the lack of statistically robust and reproducible machine learning frameworks for small-sample Vis/NIR agricultural datasets that simultaneously address stability, interpretability, and spectral variability. This study does not claim to introduce a fundamentally new machine learning algorithm. Instead, its scientific contribution lies in developing and validating an uncertainty-aware and interpretable spectral learning framework for portable Vis/NIR-based maize seed quality prediction.

To address these gaps, this study proposes a unified uncertainty-aware spectral learning framework for maize seed quality prediction using Vis/NIR spectroscopy. Unlike conventional stacking or voting ensembles that rely on static aggregation rules, the proposed framework introduces adaptive fuzzy ensemble mechanisms that dynamically adjust model contributions based on local spectral uncertainty. Specifically, two complementary architectures are developed: (i) a Fuzzy Residual-Corrected Ensemble that explicitly models structured prediction errors as uncertainty signals, and (ii) an RF+XGB Fuzzy Ensemble that performs confidence-driven adaptive weighting between heterogeneous learners. Additionally, a modern deep learning baseline (1D-CNN) is incorporated to ensure a comprehensive benchmarking framework across classical machine learning, ensemble learning, and deep learning paradigms. The framework is evaluated using repeated cross-validation, variability measures, confidence intervals, and non-parametric statistical testing to avoid overinterpreting marginal performance differences. Furthermore, SHAP-based explainability is incorporated to connect wavelength-level model behavior with physiologically meaningful spectral regions. This design enables a transition from deterministic ensemble learning to statistically validated, uncertainty-aware, and interpretable predictive modeling in spectral analysis.

The contributions of this study are threefold. First, it develops a unified spectral machine learning framework that integrates preprocessing, metaheuristic feature selection, and fuzzy ensemble modeling under a leakage-controlled nested validation protocol. Second, it introduces uncertainty-aware fuzzy ensemble architectures that extend conventional stacking by incorporating adaptive confidence-based weighting and residual modeling. Third, it provides a comprehensive experimental evaluation across multiple preprocessing scenarios and machine learning paradigms, including classical, ensemble, and deep learning baselines for maize seed quality prediction. In addition, the study strengthens model evaluation by reporting mean ± standard deviation, 95% confidence intervals, Friedman testing, Wilcoxon signed-rank testing, Nemenyi post-hoc analysis, and effect size interpretation. Finally, the study expands interpretability through SHAP-based wavelength attribution, linking important spectral regions to pigment-related visible responses and moisture-related near-infrared absorption.

2. RESEARCH METHOD

The research framework proposed in this study is shown in Figure 1. In general, the workflow consists of four main stages: (1) spectral data acquisition using a portable Vis/NIR device, (2) spectral preprocessing including NIPPY transformation and feature selection using PCA, GA, and PSO, (3) predictive modeling using baseline algorithms, conventional ensembles, and two proposed fuzzy ensemble models, and (4) performance evaluation using cross-validation and statistical metrics. To address model reliability and avoid overly optimistic performance estimation, the overall workflow was implemented under a leakage-controlled repeated cross-validation framework. Model

performance was evaluated not only using point estimates of R^2, RMSE, and MAE, but also using variability measures, confidence intervals, and non-parametric statistical validation. In addition, SHAP-based explainability was incorporated to interpret wavelength-level contributions of the predictive model.

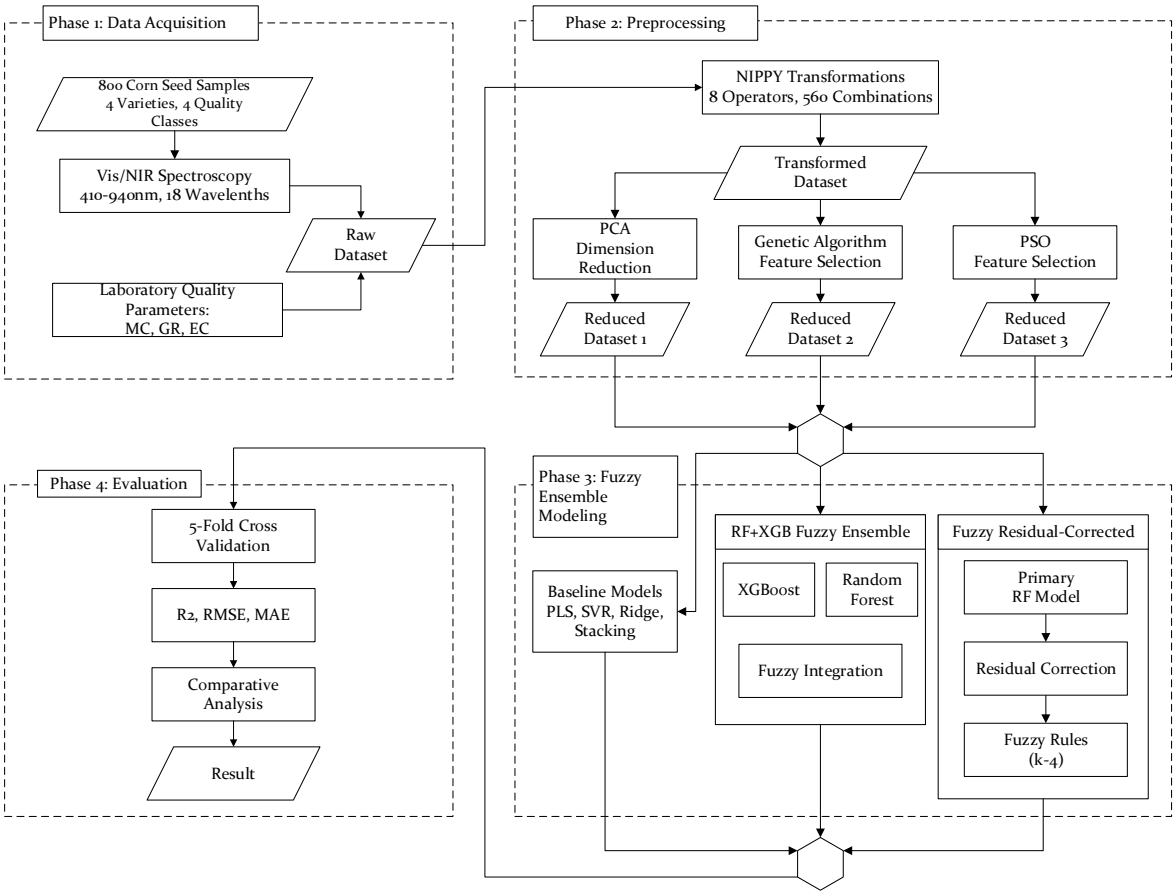

Figure 1. Research framework

2.1. Equipment and Spectral Data Acquisition

Spectral data were collected using a portable Vis/NIR spectrometer developed as a compact system for non-destructive seed analysis (Figure 2). The device integrates a multi-channel optical sensor, an embedded processing unit, and a display module in a single portable configuration. The primary sensing component is the AS7265x Smart Spectral Sensor, which provides 18 wavelength channels within the 410-940 nm range covering the visible to near-infrared region relevant for seed optical characterization. The sensor communicates with the control unit, a Raspberry Pi Zero W, through the I^2C protocol, enabling stable data acquisition and automatic scanning control [26].

To minimize ambient light interference, the seeds were measured inside a matte-black reflectance chamber with a fixed sensor-to-sample distance of approximately 1 cm. A transparent glass layer was placed between the sensor and the sample to ensure consistent measurement geometry. The internal illumination system, combined with the embedded controller, performs white-black calibration and converts measured intensity values into reflectance spectra stored in digital format. The device operates using a rechargeable battery, enabling field deployment without external power supply (Figure 3).

Figure 2. Design of portable Vis/NIR device and sensor-sample placement configuration

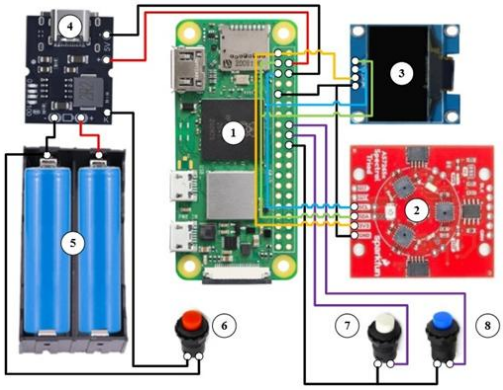

Figure 3. Circuit connection scheme (1) raspberry pi zero w, (2) AS7265x sensor, (3) oled 0.96 inch, (4) charging module, (5) battery, (6) power button, (7) calibration button, (8) scanning button

Maize seeds were categorized into four quality classes: Super, Premium, Fair, and Poor, with 200 samples per category. Before measurement, samples were conditioned at 25-30°C and 60-75% relative humidity for 24 hours to reduce environmental variability. Spectral measurements were performed in the wavelength range of 410-940 nm. Each measurement consisted of 10-15 seeds with five replications from different positions to reduce spatial variability. Calibration using a white reference panel was performed after every 20 scans to maintain reflectance stability. To minimize the risk of data leakage, spectral replicates associated with the same sample unit were kept within the same validation partition, ensuring that repeated measurements from the same sample were not simultaneously used for training and testing. The spectral acquisition procedure is illustrated in Figure 4.

Figure 4. Spectral data measurement procedure for maize seeds using Vis/NIR

2.2. Spectral Transformation

Spectral data preprocessing is a crucial step in spectroscopic analysis, as it enhances signal quality and extracts meaningful information from raw data [27], [28]. In this study, spectral transformation was performed using the NIPPY module, which offers a range of transformation operators to address various challenges in Vis/NIR spectral data [29]. The NIPPY implementation includes eight transformation operators grouped into five main functional categories based on their operating mechanisms (Table 1). Selection was carried out by evaluating 560 operator combinations using the

Random Forest algorithm and 5-fold cross-validation. All preprocessing selection procedures were embedded within the training folds of the validation process to prevent information from the test fold from influencing transformation selection. This procedure ensures that preprocessing optimization is performed in a leakage-controlled manner.

Table 1. Nippy Spectral Tranformation Operators

Category	Operator	Main Function	Parameter Configuration
Baseline Correction	BASELINE	Corrects baseline shift due to measurement variations	-
Scatter Correction	SNV, RNV, MSC, EMSC	Reduces light scattering effects	RNV: IQR 75-25, 90-10
Normalization	NORML	Standardizes the intensity scale	-
Smoothing & Derivatives	SAVGOL	Noise reduction and spectral feature extraction	Window: 5,7,11; Order: 3; Derivative: 0,1,2
Detrending	DETREND	Removes linear trends	Breakpoint: 0

2.3. Metaheuristic Feature Selection

High-dimensional spectral data often contain redundant information that may reduce predictive performance. Therefore, feature selection was implemented to identify the most informative wavelength variables [30], [31]. In this study, three feature extraction/selection strategies were investigated: Principal Component Analysis (PCA), Genetic Algorithm (GA), and Particle Swarm Optimization (PSO). PCA was used as a baseline dimensionality reduction technique by retaining eight principal components explaining approximately 94% of cumulative variance. Although PCA is effective in compressing spectral information, it transforms original variables into latent space representations, which reduces physical interpretability of wavelengths and limits direct spectral explanation[32].

In contrast, wrapper-based feature selection methods such as GA and PSO operate directly on original wavelength variables, allowing explicit selection of physically meaningful spectral regions. GA is an evolutionary optimization method introduced by Holland [33], inspired by natural selection principles. GA is widely used for combinatorial optimization problems, including feature selection in high-dimensional datasets. In this study, GA was applied to select optimal wavelength subsets by minimizing prediction error using Random Forest as the evaluation model. Random Forest is a robust ensemble learning method capable of handling nonlinear and high-dimensional data, as introduced by Breiman [34]. GA iteratively improves solutions through selection, crossover, and mutation, enabling efficient exploration of the spectral feature space. In this study, GA was configured with population size = 50 and generations = 50 to ensure sufficient exploration of the high-dimensional spectral search space and convergence stability.

PSO introduced by Kennedy and Eberhart [35], is a population-based stochastic optimization technique inspired by swarm intelligence. PSO updates candidate solutions based on both individual experience and global swarm knowledge. In this study, PSO was used under the same evaluation setting as GA to ensure fair comparison. PSO is particularly effective in exploring complex nonlinear search spaces such as spectral wavelength selection problems. PSO parameters were set to n_particles = 50, iterations = 50, w = 0.7, c_1 = 1.2, and c_2 = 1.2 to ensure balanced exploration and exploitation in the spectral feature space.

The ensemble learning framework used in this study is grounded in stacked generalization theory proposed by Wolpert [36], where combining multiple models can improve generalization performance and robustness. Furthermore, fuzzy modeling principles introduced by Zadeh [37] provide a mathematical framework to handle uncertainty in spectral data by allowing gradual membership rather than binary classification. In addition, gradient boosting principles introduced by Friedman [38] form the theoretical basis of boosting-based models used in comparative evaluation.

To ensure unbiased model evaluation, all feature selection procedures (GA and PSO) were conducted strictly within the training folds of a nested cross-validation framework. Each fold performs independent feature selection, and the selected wavelength subsets are applied only to the corresponding validation fold. This procedure ensures that no information from test data is used

during feature selection, thereby preventing data leakage and preserving the validity of performance estimation.

2.4. Modeling

Predictive modeling was conducted using three categories of machine learning approaches: baseline models, a conventional stacking ensemble, and the proposed fuzzy ensemble architectures. Five baseline algorithms were implemented, including Partial Least Squares (PLS), Support Vector Regression (SVR), Ridge Regression, Random Forest (RF), and XGBoost. Hyperparameters were optimized individually for each target variable: Moisture Content (MC), Germination Rate (GR), and Electrical Conductivity (EC). PLS was configured with six latent components to address multicollinearity in spectral data, while SVR, Ridge, RF, and XGBoost were tuned to balance model bias and variance [39], [40]. All hyperparameter tuning procedures were performed within the inner loop of a nested cross-validation framework, while the outer loop was used exclusively for model evaluation to ensure unbiased performance estimation.

In addition to classical machine learning and ensemble models, a one-dimensional Convolutional Neural Network (1D-CNN) was implemented as a deep learning baseline. The 1D-CNN model was used to provide a modern benchmark for spectral representation learning and to evaluate whether automatic feature extraction could improve prediction performance under small-sample Vis/NIR conditions. The architecture consisted of one-dimensional convolutional layers, pooling layers, and fully connected layers for regression output. The 1D-CNN was evaluated using the same validation protocol to ensure fair comparison with other algorithms.

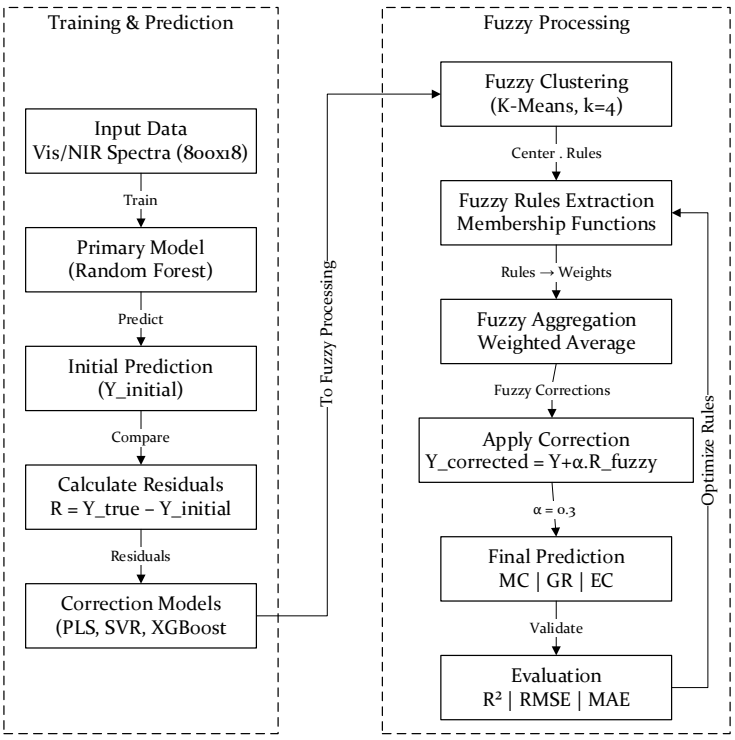

Figure 5. Fuzzy residual-corrected ensemble architecture

A stacking ensemble was implemented as a conventional ensemble approach for benchmarking [41], [42]. The architecture consists of PLS, SVR, RF, and XGBoost as base learners, while Ridge Regression acts as the meta-learner. Out-of-fold predictions generated through 5-fold cross-validation were used to train the meta-learner, ensuring that all stacking operations are performed exclusively within training folds to maintain a leakage-free evaluation protocol. The stacking formulation is defined as $\hat{y}$

= $g(f_1(x), f_2(x), f_3(x), f_4(x))$, where f_1-f_4 represent base learners and $g(\cdot)$ denotes the meta-learner function.

To improve predictive robustness under spectral noise and uncertainty, two fuzzy ensemble architectures were proposed. Unlike conventional stacking methods that rely on deterministic aggregation, the proposed fuzzy ensembles introduce uncertainty-aware adaptive weighting mechanisms that adjust model contributions based on local spectral confidence and prediction uncertainty. The fuzzy framework is formulated using Gaussian membership functions $\mu(x) = \exp(-(x - c)^2 / 2\sigma^2)$, where c represents cluster centers obtained via K-means clustering and σ controls the spread of uncertainty distribution. A representative set of fuzzy IF-THEN rules was constructed based on residual clustering behavior and model confidence patterns, rather than being used as fixed deterministic aggregation rules. The fuzzy inference engine follows a Mamdani-type structure with centroid-based defuzzification to obtain crisp outputs. The optimization objective is defined as minimization of prediction error measured by RMSE during cross-validation. Computational complexity arises mainly from nested cross-validation, ensemble training, and metaheuristic optimization, which increases computational cost but remains feasible for offline spectral modeling tasks.

The first model, the Fuzzy Residual-Corrected Ensemble, incorporates a fuzzy logic-based correction layer to explicitly model systematic residual errors produced by the primary predictor. The residual is defined as $e = y - \hat{y}\_RF$, where prediction errors are treated as structured uncertainty rather than random noise. The final corrected output is obtained as $\hat{y} = \hat{y}\_RF + F(e)$, where $F(e)$ represents the fuzzy inference output derived from Gaussian membership and rule aggregation. The architecture includes a Random Forest primary predictor and correction models (PLS, SVR, and XGBoost), as illustrated in Figure 5.

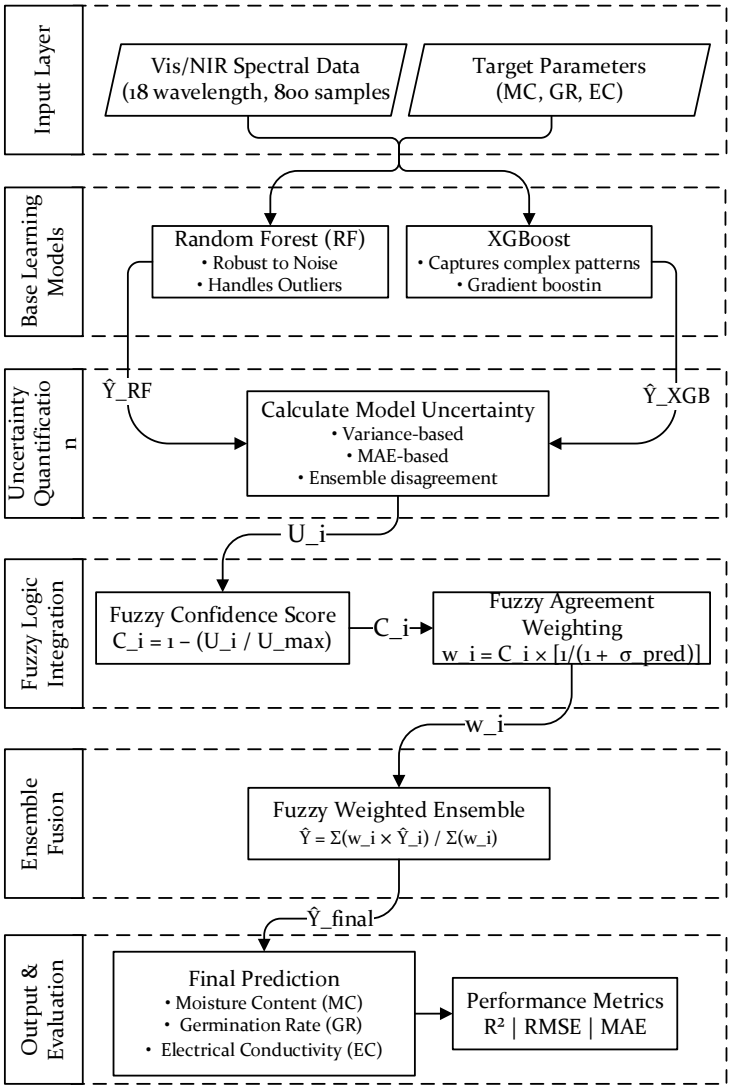

Figure 6. RF+XGB fuzzy ensemble architecture

The second model, the RF+XGB Fuzzy Ensemble, integrates Random Forest and XGBoost through adaptive fuzzy weighting mechanisms as illustrated in Figure 6. RF acts as a noise-robust stabilizer, while XGBoost captures complex nonlinear relationships in spectral data. The final prediction is computed as $\hat{y} = w\_RF\ \hat{y}\_RF + w\_XGB\ \hat{y}\_XGB$, with the constraint $w\_RF + w\_XGB = 1$. The fuzzy weights are defined using normalized Gaussian-based confidence scores $w\_i = \mu\_i(x) / (\mu\_RF(x) + \mu\_XGB(x))$, enabling dynamic adaptation of model contributions based on local spectral uncertainty conditions.

2.5. Performance Evaluation, Statistical Validation, and Explainability Analysis

To address statistical reliability, all model performances were evaluated using repeated 5-fold cross-validation. The evaluation metrics included coefficient of determination (R^2), root mean square error (RMSE), and mean absolute error (MAE). For each model and target variable, the results were reported as mean ± standard deviation (SD), and 95% confidence intervals (CI) were calculated across repeated validation folds. This reporting strategy was used to quantify performance variability and to avoid interpreting single-point accuracy values as definitive evidence of model superiority.

To statistically compare algorithm performance, non-parametric statistical tests were conducted using repeated cross-validation outputs. The Friedman test was applied to evaluate overall

differences among multiple algorithms, while the Nemenyi post-hoc test was used to identify pairwise rank differences among competing models. In addition, the Wilcoxon signed-rank test was used for paired comparisons between Random Forest, as the strongest baseline model, and the proposed fuzzy ensemble models. Effect size measures, including Kendall's W for Friedman analysis and rank-biserial correlation for Wilcoxon comparisons, were used to interpret practical significance beyond p-values. This statistical analysis follows established recommendations for comparing machine learning models across repeated validation results [43], [44], [45].

SHAP (SHapley Additive exPlanations) analysis was conducted to interpret wavelength-level contributions to model prediction. In this study, SHAP analysis was computed using the Random Forest primary learner, which represents the strongest and most stable predictive component within the proposed fuzzy ensemble framework. Mean absolute SHAP values were calculated for each wavelength and target variable across repeated validation folds. The resulting wavelength importance profile was used to identify physiologically meaningful spectral regions associated with maize seed quality, particularly the pigment-related visible region around 410-500 nm and the moisture-related near-infrared region around 850-940 nm [46].

3. RESULTS AND DISCUSSIONS

3.1. Statistical Characteristics of Maize Seed Data and Vis/NIR Spectra

Descriptive statistical analysis of 800 maize seed samples indicates substantial variability across the three quality parameters (Table 2). Moisture Content (MC) ranged from 9.00-15.00% with a mean of 12.49 ± 1.61%, consistent with the typical moisture range used for cereal seed storage and vigor evaluation (11-15%) [47]. The distribution shows relatively low dispersion with mild asymmetry, indicating a stable sampling range suitable for regression-based modeling. Germination Rate (GR) exhibited a wider distribution (50-99%, CV = 17.22%) compared with certified seed lots that typically show narrow germination ranges above 80-90% [48]. This wider variability is intentionally beneficial for machine learning modeling, as it increases the representativeness of both high-vigor and low-vigor seed conditions, thereby improving model generalization capability under heterogeneous data distributions. Electrical Conductivity (EC) showed the largest variability (5.10-69.90 µS/cm, CV = 56.11%), reflecting its sensitivity as a physiological indicator of membrane integrity degradation. The high coefficient of variation confirms EC as a highly nonlinear response variable, which justifies the use of nonlinear models such as ensemble learning and fuzzy-based approaches. Small physiological differences among seeds can induce substantial EC variations, making it a strong proxy for seed deterioration and vigor assessment [49].

Table 2. Descriptive Statistics of Maize Seed Quality Parameters

Parameter	Min	Max	Mean ± SD	Median	Q25	Q75	CV (%)
MC (%)	9.00	15.00	12.49 ± 1.61	12.60	11.40	13.80	12.89
GR (%)	50.00	99.00	83.17 ± 14.32	87.00	73.00	96.00	17.22
EC (µS/cm)	5.10	69.90	30.18 ± 16.94	26.50	17.20	41.15	56.11

The dataset also demonstrates balanced representation across seed quality categories (Super, Premium, Fair, and Poor) and maize varieties (Maxxi, MPM, SR9, and Konsumsi), ensuring diversity in spectral and physiological characteristics. This balanced distribution reduces potential sampling bias and supports robust supervised learning across multiple modeling strategies, including ensemble learning and deep neural networks. Spectral analysis within the 410-940 nm region revealed reflectance characteristics related to maize seed physiological properties. The visible region is associated with pigment absorption and surface optical characteristics, while the early near-infrared region reflects biochemical responses related to C-H, N-H, and O-H bond vibrations. The late near-infrared region, particularly 850-940 nm, is frequently associated with moisture-related O-H overtone absorption in grain materials [50]. These spectral characteristics provide the physical basis for the SHAP-based

explainability analysis discussed in Section 3.4, where wavelength-level model contributions are interpreted in relation to pigment-related and moisture-related spectral responses.

Figure 7 visualizes the spectral patterns across seed quality categories. In general, high-quality seeds such as Super and Premium exhibit stronger reflectance patterns in selected visible and near-infrared regions than lower-quality seeds. This suggests that spectral responses are associated with physiological and biochemical differences among seed quality groups.

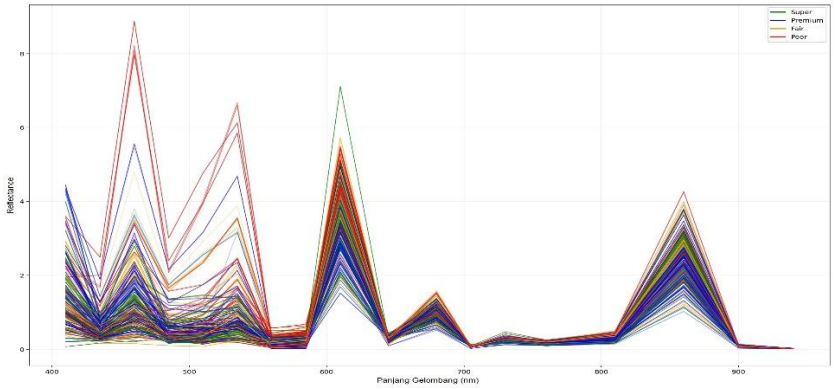

Figure 7. Spectral visualization based on quality categories (super, premium, fair, poor).

3.2. Predictive Performance Based on Preprocessing Scenarios

Table 3 summarizes the predictive performance obtained for each maize seed quality parameter under eight preprocessing scenarios combining spectral transformation and feature selection. In response to the reviewer's recommendation on statistical reporting, the results are now presented as mean ± standard deviation (SD) with 95% confidence intervals (CI) obtained from repeated cross-validation. This reporting format provides a clearer indication of performance variability and avoids overinterpreting single-point R^2 values. The coefficient of determination (R^2) was used as the primary indicator of predictive performance, while SD and CI were used to assess stability across validation runs.

As shown in Table 3, applying NIPPY spectral transformation slightly improved the prediction performance for MC and GR. The mean R^2 increased from 0.812 to 0.818 for MC and from 0.847 to 0.860 for GR. This indicates that spectral preprocessing contributes mainly to noise reduction, baseline stabilization, and correction of spectral distortion rather than producing large-scale predictive gains. This finding is consistent with previous studies reporting that appropriate spectral preprocessing can improve the stability of Vis/NIR-based prediction models, particularly when spectral data contain scattering effects, baseline variation, and overlapping absorption bands. In contrast, EC showed a slight decrease after NIPPY transformation, suggesting that spectral information related to membrane leakage may already be strongly preserved in the raw spectral measurements.

Table 3. Predictive performance of preprocessing scenarios reported as mean ± SD with 95% confidence interval

Preprocessing Scenario	MC (R^2 ± SD, 95% CI)	GR (R^2 ± SD, 95% CI)	EC (R^2 ± SD, 95% CI)
NONE	0.812 ± 0.004 (0.804-0.820)	0.847 ± 0.006 (0.835-0.859)	0.862 ± 0.007 (0.848-0.876)
PCA	0.801 ± 0.005 (0.791-0.811)	0.840 ± 0.006 (0.828-0.852)	0.861 ± 0.006 (0.849-0.873)
GA	0.815 ± 0.004 (0.807-0.823)	0.863 ± 0.005 (0.853-0.873)	0.869 ± 0.004 (0.861-0.877)
PSO	0.812 ± 0.004 (0.804-0.820)	0.853 ± 0.005 (0.843-0.863)	0.864 ± 0.004 (0.856-0.872)
NIPPY	0.818 ± 0.004 (0.810-0.826)	0.860 ± 0.005 (0.850-0.870)	0.859 ± 0.006 (0.847-0.871)
NIPPY+PCA	0.818 ± 0.004 (0.810-0.826)	0.860 ± 0.005 (0.850-0.870)	0.859 ± 0.005 (0.849-0.869)
NIPPY+GA	0.825 ± 0.004 (0.817-0.833)	0.864 ± 0.004 (0.856-0.872)	0.862 ± 0.004 (0.854-0.870)
NIPPY+PSO	0.818 ± 0.004 (0.810-0.826)	0.856 ± 0.005 (0.846-0.866)	0.861 ± 0.004 (0.853-0.869)

The optimized NIPPY transformation results are illustrated in Figure 8. The figure shows that the preprocessing procedure modifies the spectral profile by smoothing high-frequency fluctuations and enhancing relevant spectral patterns. For MC, the optimal preprocessing pipeline includes baseline correction, normalization, first-derivative transformation, and Savitzky-Golay filtering with a window size of 7 and polynomial order of 3. For GR and EC, a larger Savitzky-Golay window size of 11 provides better generalization, indicating that these parameters are more strongly influenced by broader spectral trends than by narrow local fluctuations.

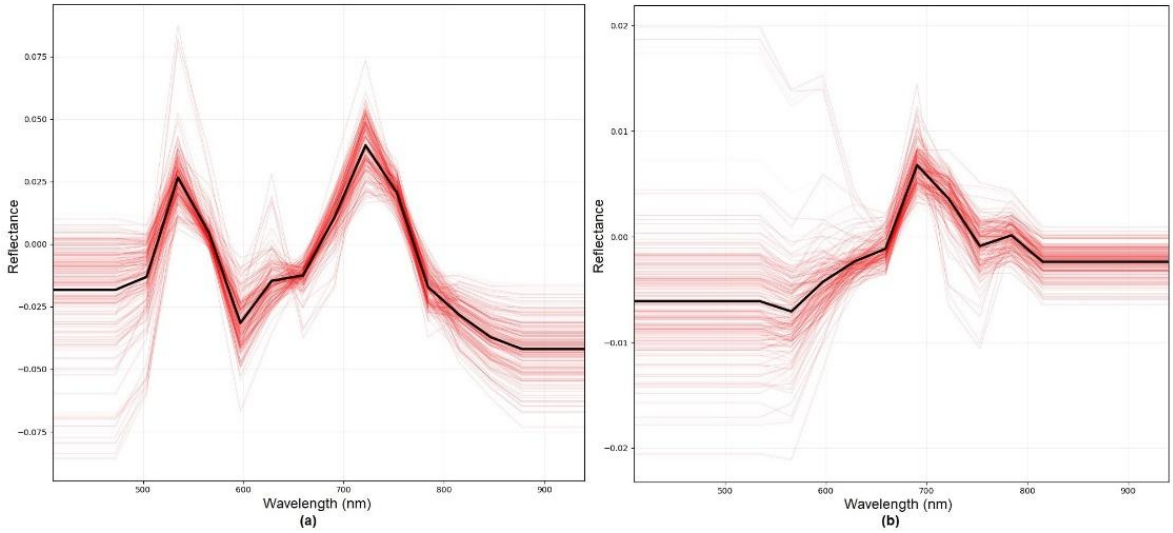

Figure 8. Optimized NIPPY spectral transformation results for (a) MC, (b) GR, and EC

Feature selection analysis is presented in Figure 9. PCA tends to reduce R^2 across the evaluated parameters, particularly for MC and GR. This may occur because PCA transforms the original wavelength variables into latent components, which can reduce direct physical interpretability and remove subtle wavelength-specific information relevant to seed physiological responses [31]. In contrast, GA consistently improves or maintains predictive performance by selecting informative wavelength subsets in the original spectral domain. This is important for Vis/NIR applications because wavelength-level selection allows model behavior to remain interpretable in relation to biochemical absorption mechanisms [51]. PSO provides moderate improvement, particularly for GR and EC, although its performance is generally slightly lower than GA. This difference may be related to the broader exploration capability of GA through selection, crossover, and mutation mechanisms, while PSO may converge faster but can be less exploratory in high-dimensional wavelength selection problems [52].

The overall comparison of preprocessing scenarios is shown in Figure 10. The NIPPY+GA scenario achieves the highest mean R^2 for MC (0.825 ± 0.004) and GR (0.864 ± 0.004), suggesting that spectral denoising followed by evolutionary wavelength selection provides the most stable pipeline for these parameters. However, for EC, GA alone yields the highest mean R^2 value (0.869 ± 0.004), indicating that localized wavelength selection may be more important than global spectral transformation for parameters related to membrane leakage and electrolyte release [14], [53]. The overlapping confidence intervals among several scenarios indicate that preprocessing effects should be interpreted in terms of stability and robustness rather than large absolute gains in predictive accuracy. Overall, these results confirm that preprocessing effectiveness is strongly dependent on the biochemical nature of the target variable and should not be assumed to be universally optimal across all seed quality indicators.

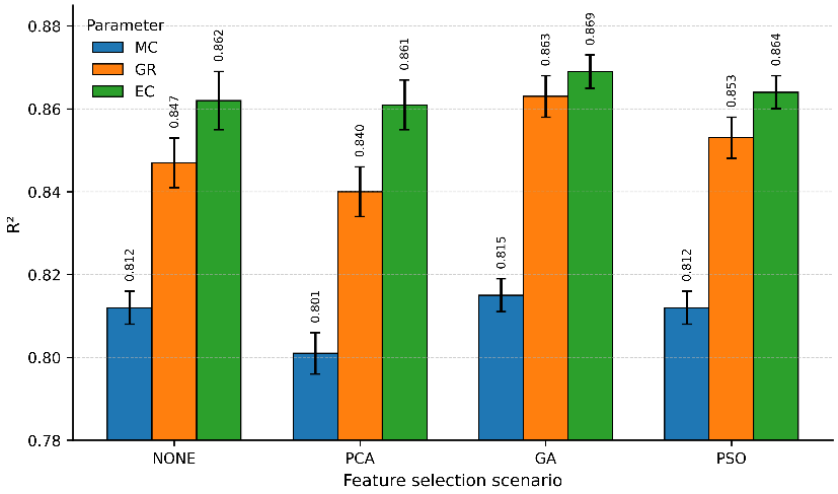

Figure 9. Comparison of feature selection effects on prediction performance with standard deviation error bars

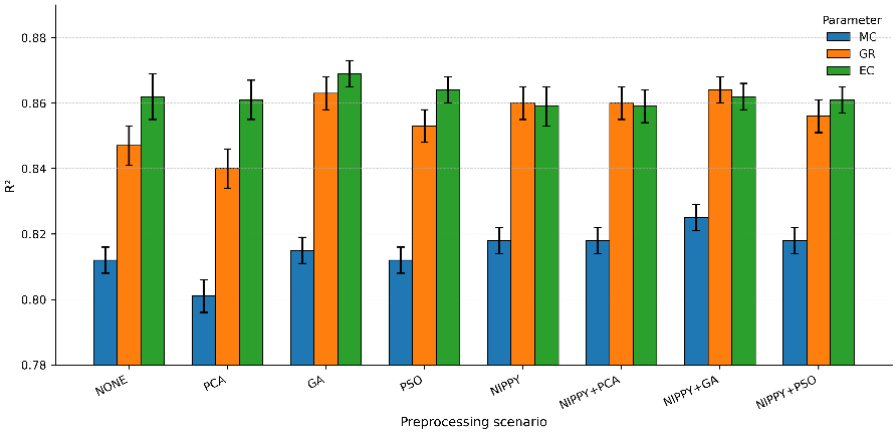

Figure 10. Comparison of preprocessing scenarios for MC, GR, and EC prediction with standard deviation error bars

3.3. Prediction Results Based on Algorithms

Table 4 presents the predictive performance of baseline models, conventional ensemble learning, the 1D-CNN deep learning baseline, and the proposed fuzzy ensemble approaches under the optimal preprocessing conditions identified in Section 3.2. To strengthen statistical reporting, the results are reported as mean ± SD with 95% CI obtained from repeated cross-validation. This enables model evaluation to consider both average predictive performance and variability across validation folds. As shown in Table 4, the strongest predictive performance is obtained by Random Forest and the two fuzzy ensemble architectures. The Fuzzy Residual-Corrected Ensemble achieves R^2 values of 0.8249 ± 0.004 for MC, 0.8636 ± 0.004 for GR, and 0.8689 ± 0.004 for EC. Random Forest achieves highly comparable results, with R^2 values of 0.8251 ± 0.004, 0.8626 ± 0.004, and 0.8685 ± 0.004, respectively. The RF+XGB Fuzzy Ensemble also produces similar performance. The overlap in 95% CI among these top-performing models indicates that their differences in mean predictive accuracy are very small. For example, the improvement of the Fuzzy Residual-Corrected Ensemble over Random Forest for EC is only 0.0004 in R^2. Therefore, the result should not be interpreted as evidence of substantial practical superiority in mean accuracy.

To further examine whether the observed numerical differences are statistically meaningful, additional non-parametric statistical tests were conducted using repeated cross-validation outputs. The Friedman test was used to evaluate overall differences among algorithms, the Nemenyi post-hoc

test was applied to assess pairwise rank differences, and the Wilcoxon signed-rank test was used for paired comparisons between Random Forest and the proposed fuzzy ensemble models. The results are summarized in Table 5.

Table 4. Comparison of predictive performance across algorithms reported as mean ± SD with 95% confidence interval

Algorithm	MC (R^2 ± SD, 95% CI)	GR (R^2 ± SD, 95% CI)	EC (R^2 ± SD, 95% CI)
Fuzzy Residual-Corrected Ensemble	0.8249 ± 0.004 (0.817-0.833)	0.8636 ± 0.004 (0.856-0.872)	0.8689 ± 0.004 (0.861-0.877)
Random Forest	0.8251 ± 0.004 (0.817-0.833)	0.8626 ± 0.004 (0.855-0.871)	0.8685 ± 0.004 (0.861-0.877)
RF+XGB Fuzzy Ensemble	0.8251 ± 0.004 (0.817-0.833)	0.8626 ± 0.004 (0.855-0.871)	0.8685 ± 0.004 (0.860-0.876)
1D-CNN	0.8042 ± 0.004 (0.796-0.812)	0.8513 ± 0.005 (0.841-0.861)	0.8557 ± 0.004 (0.848-0.864)
XGBoost	0.6326 ± 0.004 (0.625-0.641)	0.7127 ± 0.005 (0.703-0.723)	0.7104 ± 0.004 (0.702-0.718)
Stacking Ensemble	0.3962 ± 0.005 (0.386-0.406)	0.4496 ± 0.006 (0.438-0.462)	0.5185 ± 0.006 (0.507-0.531)
Ridge Regression	0.1920 ± 0.005 (0.182-0.202)	0.2819 ± 0.006 (0.270-0.294)	0.2767 ± 0.007 (0.263-0.291)
PLS Regression	0.1885 ± 0.005 (0.179-0.199)	0.2725 ± 0.005 (0.263-0.283)	0.2637 ± 0.006 (0.254-0.274)
SVR	0.2129 ± 0.004 (0.205-0.221)	0.1972 ± 0.006 (0.185-0.209)	0.1313 ± 0.005 (0.122-0.142)

Table 5. Summary of non-parametric statistical validation between Random Forest and the Fuzzy Residual-Corrected Ensemble

Scope	Friedman χ^2, p-value	Kendall's W	Best Average Rank	RF Average Rank	Wilcoxon RF vs Fuzzy, p-value; effect size	Nemenyi RF vs Fuzzy, p-value	Interpretation
MC	337.5267, $p < 0.001$	0.9644	Fuzzy = 1.3400	RF = 1.8800	$p < 0.001$; r = 0.6251	p = 0.9565	Large overall model difference; RF-Fuzzy difference not significant by Nemenyi
GR	323.5667, $p < 0.001$	0.9245	Fuzzy = 1.2400	RF = 2.4400	$p < 0.001$; r = 0.9733	p = 0.2176	Large overall model difference; RF-Fuzzy difference not significant by Nemenyi
EC	332.2200, $p < 0.001$	0.9492	Fuzzy = 1.2600	RF = 2.2000	$p < 0.001$; r = 0.9467	p = 0.5378	Large overall model difference; RF-Fuzzy difference not significant by Nemenyi
Overall	962.6422, $p < 0.001$	0.9168	Fuzzy = 1.2800	RF = 2.1733	$p < 0.001$; r = 0.8610	p = 0.0340	Fuzzy shows the best overall rank, but practical gains in mean R^2 remain marginal

The statistical comparison in Table 5 shows that the Friedman test confirmed significant overall differences among algorithms for MC, GR, EC, and overall evaluation, with large Kendall's W effect sizes. This indicates that model choice has a substantial influence on predictive behavior. The Fuzzy Residual-Corrected Ensemble obtained the best average rank across all targets and overall evaluation, suggesting improved ranking stability across validation folds. However, the Nemenyi post-hoc comparisons at the individual target level did not show consistently significant differences between Random Forest and the Fuzzy Residual-Corrected Ensemble. Therefore, the advantage of the fuzzy ensemble should be interpreted cautiously.

The Wilcoxon signed-rank test indicates that the Fuzzy Residual-Corrected Ensemble tends to provide more consistent fold-level performance than Random Forest. Nevertheless, because the mean R^2 difference is extremely small and the confidence intervals overlap, the proposed fuzzy ensemble framework is not positioned as a model that substantially improves mean predictive accuracy over Random Forest. Instead, its contribution lies in residual correction, uncertainty-aware modeling, and improved stability under spectral variability.

The 1D-CNN baseline achieves competitive performance, with R^2 values ranging from 0.8042 to 0.8557, but it does not outperform the tree-based and fuzzy ensemble models. This suggests that deep learning models may require larger and more diverse spectral datasets to fully exploit their representation learning capacity. In small-sample Vis/NIR spectroscopy, structured ensemble models remain advantageous because they can combine nonlinear modeling capability with stable behavior under limited data conditions.

In contrast, classical linear models such as Ridge Regression, PLS Regression, and SVR show substantially lower predictive performance, with R^2 values below 0.30 for most target parameters. This confirms that the relationships between Vis/NIR spectra and maize seed quality parameters are nonlinear and cannot be adequately captured by linear assumptions [54], XGBoost improves performance compared with linear baselines but remains below Random Forest and fuzzy ensemble models. The stacking ensemble also performs lower than expected, which may be attributed to error propagation during meta-learning and sensitivity to multicollinearity among spectral variables under limited sample conditions [41]. Overall, the results indicate that the fuzzy ensemble framework should be interpreted primarily as an uncertainty-aware and statistically stable modeling approach rather than as a method that produces large absolute gains in predictive accuracy.

3.4. Fuzzy-Ensemble Approach and SHAP-Based Explainability

The proposed fuzzy ensemble framework provides a mechanism for handling spectral uncertainty and nonlinear relationships in Vis/NIR data through adaptive weighting and residual-based correction. However, based on the statistical validation in Section 3.3, the contribution of the fuzzy ensemble models should not be interpreted as substantial superiority in mean predictive accuracy. Instead, the main scientific value lies in uncertainty-aware prediction, residual correction, ranking stability, and interpretable spectral modeling. The Fuzzy Residual-Corrected Ensemble achieved performance comparable to Random Forest while obtaining the best average rank across MC, GR, EC, and overall evaluation. This indicates that modeling residual errors as structured uncertainty can improve stability across validation folds, even when the mean R^2 improvement is marginal.

To strengthen interpretability and assess whether the predictive model captures physiologically meaningful spectral regions, SHAP analysis was conducted using the Random Forest primary learner within the proposed fuzzy ensemble framework [46]. The SHAP results are presented in Figure 11 and summarized by wavelength-level importance for MC, GR, and EC. Unlike a general feature importance plot, SHAP provides an additive attribution framework that estimates how each wavelength contributes to model prediction, allowing the spectral behavior of the model to be interpreted in relation to maize seed physiology.

For MC, the most influential wavelengths were 535 nm, 585 nm, 560 nm, 645 nm, and 485 nm. The importance of 485 nm is particularly meaningful because it falls within the 410-500 nm visible region, which is associated with pigment-related and surface optical characteristics of biological materials. In addition, the contribution of 535-645 nm suggests that visible and early near-infrared responses are relevant for moisture-related prediction in maize seeds. For GR, the dominant wavelengths were 585 nm, 535 nm, 560 nm, 610 nm, and 760 nm, indicating that germination-related prediction is influenced by visible to early near-infrared biochemical responses. For EC, the most important wavelengths were 585 nm, 535 nm, 760 nm, 560 nm, and 645 nm, suggesting that the model captures spectral responses associated with physiological deterioration and membrane-related seed quality changes.

The SHAP profile also directly addresses the physiological relevance of the 410-500 nm and 850-940 nm regions. The 410-500 nm region contributes to model interpretation through pigment-related and surface optical responses, particularly for MC where 485 nm appears among the top-ranked wavelengths. Meanwhile, the 850-940 nm region, although not always ranked as the dominant contributor, remains physiologically meaningful because wavelengths in this range are associated with moisture-related O-H overtone absorption in grain materials. This indicates that the model captures both dominant biochemical regions and targeted physiological regions relevant to maize seed quality.

Overall, the SHAP results indicate that the proposed framework does not operate merely as a black-box predictor. Instead, it provides a transparent link between wavelength-level spectral information, seed physiological characteristics, and predictive behavior. This strengthens the scientific contribution of the study by combining uncertainty-aware fuzzy ensemble prediction with interpretable spectral attribution in a portable Vis/NIR-based maize seed quality assessment framework.

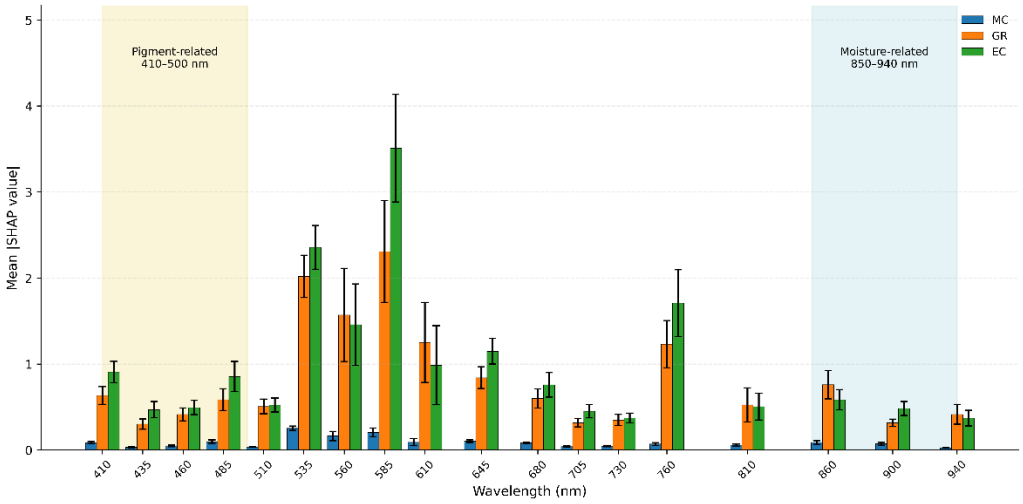

Figure 11. SHAP-based wavelength importance for MC, GR, and EC prediction using the Random Forest

3.5. Practical Implications, Limitations, and Future Work

The integration of portable Vis/NIR spectroscopy with machine learning and fuzzy ensemble modeling provides a practical and non-destructive framework for maize seed quality assessment, enabling rapid estimation of MC, GR, and EC in controlled laboratory and post-harvest environments. This approach demonstrates potential applicability for early-stage decision support in seed quality screening and classification systems. However, based on the statistical validation results, the proposed framework should be understood primarily as an uncertainty-aware and interpretable spectral modeling approach rather than as a system that provides large accuracy gains over the strongest baseline model.

From a methodological perspective, the proposed framework contributes to uncertainty-aware spectral modeling by integrating spectral preprocessing, metaheuristic wavelength selection, fuzzy residual correction, statistical validation, and SHAP-based explainability. All experiments were conducted under controlled laboratory conditions using a dataset of 800 samples across four maize varieties, which may limit the generalization capability of the model under broader agro-ecological conditions. In addition, the absence of external multi-location validation restricts the assessment of domain robustness and cross-environment adaptability.

Although repeated cross-validation and leakage-controlled data partitioning were applied to ensure unbiased evaluation, the relatively small sample size combined with high spectral correlation may still introduce optimistic bias in performance estimation, particularly for deep learning and ensemble-based models. Another limitation relates to computational complexity. The use of metaheuristic feature selection (GA and PSO), combined with ensemble learning and fuzzy inference mechanisms, increases computational overhead, which may limit real-time deployment on low-resource or edge computing devices. Despite these limitations, the proposed framework demonstrates strong potential for improving spectral learning stability through uncertainty-aware fuzzy modeling, residual correction, statistical validation, and explainable AI analysis.

Future research should focus on extending validation to multi-location and multi-season datasets to improve external generalization capability. Domain adaptation techniques and transfer learning strategies should also be explored to address cross-environment variability in spectral data. Optimization of computational efficiency is necessary to support real-time implementation in portable or embedded Vis/NIR systems. Furthermore, hybrid approaches combining deep learning-based spectral feature extraction with fuzzy ensemble mechanisms may further enhance predictive robustness, scalability, and interpretability in precision agriculture applications.

4. CONCLUSION

This study proposed a statistically validated, uncertainty-aware, and interpretable framework for maize seed quality prediction using portable Vis/NIR spectroscopy. The framework integrates NIPPY-based spectral preprocessing, metaheuristic wavelength selection, and fuzzy ensemble learning to address spectral noise, wavelength redundancy, nonlinear relationships, and prediction uncertainty in small-sample spectral data. The results show that preprocessing combined with GA-based wavelength selection improved model stability, particularly for MC and GR, while GA alone provided the strongest performance for EC, indicating that the effectiveness of preprocessing is target-dependent. Among the evaluated models, Random Forest and the proposed fuzzy ensemble architectures achieved the strongest predictive performance, with R^2 values ranging from 0.8249 to 0.8689 across MC, GR, and EC. The Fuzzy Residual-Corrected Ensemble obtained the best average rank in the statistical validation results. However, the improvement in mean accuracy over Random Forest was marginal, and the confidence intervals among the top-performing models overlapped. Therefore, the main contribution of the fuzzy ensemble framework should not be interpreted as a large gain in predictive accuracy, but as improved ranking stability, residual correction, and uncertainty-aware prediction under spectral variability. The results also confirm that classical baseline models, including Ridge Regression, PLS Regression, and SVR, were less effective in capturing the nonlinear relationship between Vis/NIR spectra and maize seed quality parameters. The 1D-CNN baseline showed competitive performance but did not outperform the tree-based and fuzzy ensemble models, suggesting that deep learning may require larger and more diverse spectral datasets to achieve stronger generalization. SHAP-based explainability further demonstrated that the model identified physiologically meaningful wavelength regions, including pigment-related visible bands and moisture-related near-infrared bands, thereby improving the transparency of the proposed framework. Overall, the proposed framework provides a reliable and interpretable approach for portable Vis/NIR-based maize seed quality prediction. Nevertheless, this study remains limited to laboratory-scale data consisting of 800 maize seed samples from four varieties. Future work should focus on external validation using multi-location and multi-season datasets, improving computational efficiency for embedded implementation, and exploring hybrid deep learning–fuzzy ensemble architectures to enhance robustness, scalability, and field applicability.

ACKNOWLEDGEMENTS

This research was funded through the 2025 Research Grant Program from the Directorate of Research and Community Service, Directorate General of Research and Development, Ministry of Education, Culture, Research, and Technology, Indonesia, under Contract No. 270/UN.61.4/PFR/2025 (Hibah Fundamental). We sincerely appreciate the financial support that has facilitated the completion of this study.

DECLARATIONS

AI USAGE STATEMENT

The authors declare that AI-assisted tools, including OpenAI and QuillBot, were used only for language refinement, grammar checking, and improving sentence clarity. These tools were not used for the conceptual framework, research design, data analysis, interpretation of results, or conclusions. The authors reviewed and approved the final manuscript and take full responsibility for its content.

AUTHOR CONTRIBUTION

Ridwan Raafi'udin developed the research idea, designed the experimental framework, performed the Vis/NIR spectroscopy data processing, conducted the model development and numerical experiments, analyzed the results, and drafted the manuscript. Ali Khumaidi contributed to the methodological design, metaheuristic optimization strategy, fuzzy ensemble modelling, validation of the

computational workflow, result interpretation, and manuscript refinement. Indra Permana Solihin supported the literature review, data preparation, model implementation, visualization, and discussion of the machine learning results. Erik Mulyana contributed to the agronomic interpretation, maize seed quality assessment, validation of seed-related parameters, and discussion of the practical relevance of the findings. All authors reviewed, revised, and approved the final version of the manuscript.

CONFLICTING INTERESTS

The authors declare no conflicts of interest.

REFERENCES

[1] H. Sidahmed, A. Vad, and J. Nagy, "Advances in Sweet Corn (Zea mays L. saccharata) Research from 2010 to 2025: Genetics, Agronomy, and Sustainable Production," *Agronomy*, vol. 15, no. 5, p. 1260, May 2025, doi: 10.3390/agronomy15051260.

[2] R. M. Kamel, D. V. Assaha, M. El-kholy, and H. I. A. Elsawy, "Influence of maize (Zea mays L.) moisture content stored in various compacted multi-layer hermetic bags on seed quality and mycotoxins contamination," *J. Stored Prod. Res.*, vol. 107, p. 102319, Jun. 2024, doi: 10.1016/j.jspr.2024.102319.

[3] M. B. Abadía, L. A. Castillo, Y. N. Alonso, M. G. Monterubbianesi, G. Maciel, and R. E. Bartosik, "Germination and Vigor of Maize Seeds: Pilot-Scale Comparison of Low-Oxygen and Traditional Storage Methods," *Agriculture*, vol. 14, no. 8, p. 1268, Aug. 2024, doi: 10.3390/agriculture14081268.

[4] S. Wang *et al.*, "A Rapid and Quantitative Method for Determining Seed Viability Using 2,3,5-Triphenyl Tetrazolium Chloride (TTC): With the Example of Wheat Seed," *Molecules*, vol. 28, no. 19, p. 6828, Sep. 2023, doi: 10.3390/molecules28196828.

[5] R. B. Hegde, V. Kudva, S. Nayak, N. Sampathila, and A. Thalengala, "Advanced techniques for seed quality assessment and germination monitoring," *Discov. Appl. Sci.*, vol. 7, no. 7, p. 690, Jul. 2025, doi: 10.1007/s42452-025-07284-8.

[6] Z. Ji, J. Zhu, J. Deng, F. Meng, H. Jiang, and Q. Chen, "High-precision identification of zearalenone contamination in wheat based on olfactory sensor combined with near-infrared spectroscopy," *J. Food Compos. Anal.*, vol. 145, p. 107805, Sep. 2025, doi: 10.1016/j.jfca.2025.107805.

[7] M. L. F. Simeone *et al.*, "Portable near-infrared (NIR) spectroscopy and multivariate calibration for reliable quality control of maize and sorghum grain chemical composition," *J. Food Compos. Anal.*, vol. 134, p. 106502, Oct. 2024, doi: 10.1016/j.jfca.2024.106502.

[8] I. J. S. Ferreira, D. dos Santos Costa, L. A. Rolim, S. T. de Freitas, N. A. C. de Souza, and B. Teruel, "Monitoring pesticides with portable NIR spectroscopy in different intact fruits," *J. Food Compos. Anal.*, vol. 148, p. 108024, Dec. 2025, doi: 10.1016/j.jfca.2025.108024.

[9] D. A. S. Saputri, M. Fahri Reza Pahlawan, B. M. A. Murti, and R. E. Masithoh, "Vis/NIR spectroscopy for non-destructive method in detecting soybean seeds viability," *IOP Conf. Ser. Earth Environ. Sci.*, vol. 1038, no. 1, p. 012043, Jun. 2022, doi: 10.1088/1755-1315/1038/1/012043.

[10] D. G. D. Bernardo Oliveira, G. Gorla, E. P. de Medeiros, S. da S. Simões, J. M. Amigo, and M. F. Pimentel, "Cottonseed germination prediction with non-destructive NIR strategies," *Talanta*, vol. 298, p. 129017, Feb. 2026, doi: 10.1016/j.talanta.2025.129017.

[11] K. B. Beć, J. Grabska, and C. W. Huck, "Interpretability in near-infrared (NIR) spectroscopy: Current pathways to the long-standing challenge," *TrAC Trends Anal. Chem.*, vol. 189, p. 118254, Aug. 2025, doi: 10.1016/j.trac.2025.118254.

[12] X. Wang *et al.*, "Simultaneous estimation of multiple soil properties from vis-NIR spectra using a multi-gate mixture-of-experts with data augmentation," *Geoderma*, vol. 453, p. 117127, Jan. 2025, doi: 10.1016/j.geoderma.2024.117127.

[13] Y. Wang *et al.*, "NIRS-based detection advances in agriculture: Data enhancement, characteristic wavelength selection and modelling techniques," *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.*, vol. 343, p. 126611, Dec. 2025, doi: 10.1016/j.saa.2025.126611.

[14] A. Khumaidi, R. Raafi'udin, and N. S. Triastuti, "Enhancing Ship Coating Quality Detection via Machine Learning-Optimized Visible Near-Infrared Spectroscopy," *Instrum. Mes. Métrologie*, vol. 23, no. 6, pp. 441–450, Dec. 2024, doi: 10.18280/i2m.230604.

[15] T. Zhu and J. Xing, "Near-infrared spectroscopy and ensemble learning modeling for moisture detection in forest floor leaf litter," *Vib. Spectrosc.*, vol. 140, p. 103841, Sep. 2025, doi: 10.1016/j.vibspec.2025.103841.

[16] Z. Zou *et al.*, "Research on nondestructive detection of sweet-waxy corn seed varieties and mildew based on stacked ensemble learning and hyperspectral feature fusion technology," *Spectrochim. Acta Part A Mol.*

Biomol. Spectrosc., vol. 322, p. 124816, Dec. 2024, doi: 10.1016/j.saa.2024.124816.
[17] R. Saatchi, "Fuzzy Logic Concepts, Developments and Implementation," *Information*, vol. 15, no. 10, p. 656, Oct. 2024, doi: 10.3390/info15100656.
[18] A. Khumaidi, Y. A. Purwanto, H. Sukoco, and S. H. Wijaya, "Using Fuzzy Logic to Increase Accuracy in Mango Maturity Index Classification: Approach for Developing a Portable Near-Infrared Spectroscopy Device," *Sensors*, vol. 22, no. 24, p. 9704, Dec. 2022, doi: 10.3390/s22249704.
[19] H. Ouifak and A. Idri, "A comprehensive review of fuzzy logic based interpretability and explainability of machine learning techniques across domains," *Neurocomputing*, vol. 647, p. 130602, Sep. 2025, doi: 10.1016/j.neucom.2025.130602.
[20] Z. Shi, L. Chen, J. Duan, G. Chen, and K. Zhao, "Robust and fuzzy ensemble framework via spectral learning for random projection-based fuzzy-c-means clustering," *Eng. Appl. Artif. Intell.*, vol. 117, p. 105541, Jan. 2023, doi: 10.1016/j.engappai.2022.105541.
[21] J. Li *et al.*, "Explainable AI opens the black-box model of spectral technique to trace agro-products origin: The case of kiwifruit," *Artif. Intell. Agric.*, vol. 235, p. 110354, Jun. 2026, doi: 10.1016/j.aiia.2026.05.010.
[22] M. T. Ahmed, M. W. Ahmed, and M. Kamruzzaman, "A systematic review of explainable artificial intelligence for spectroscopic agricultural quality assessment," *Comput. Electron. Agric.*, vol. 235, p. 110354, Aug. 2025, doi: 10.1016/j.compag.2025.110354.
[23] A. Savari, Y. Li, K. Alnefaie, and N. S. S. Singh, "State-of-the-art Machine Learning Advances in Reliability-based Design, Integrity Assessment, Inspection and Maintenance of Pipelines: A Systematic Review," *J. Pipeline Sci. Eng.*, p. 100528, May 2026, doi: 10.1016/j.jpse.2026.100528.
[24] P. Muthudoss *et al.*, "Machine Learning-Enabled NIR Spectroscopy in Assessing Powder Blend Uniformity: Clear-Up Disparities and Biases Induced by Physical Artefacts," *AAPS PharmSciTech*, vol. 23, no. 7, p. 277, Oct. 2022, doi: 10.1208/s12249-022-02403-9.
[25] J. Workman, Jr. and L. Weyer, *Practical Guide to Interpretive Near-Infrared Spectroscopy*. CRC Press, 2007. doi: 10.1201/9781420018318.
[26] A. Khumaidi, R. Raafiudin, S. Muis, D. Susanti, and R. Budiarto, "Embedded System Design and Control of a Portable Vis/NIR Spectroscopy with Hybrid XGBoost-ANFIS for Soybean Seed Quality Prediction," *Int. J. Robot. Control Syst.*, vol. 6, no. 1, 2026, doi: https://doi.org/10.31763/ijrcs.v6i1.2392.
[27] M. Eslamifar, H. Tavakoli, E. Thiessen, R. Kock, J. Correa, and E. Hartung, "Effective spectral pre-processing methods enhance accuracy of soil property prediction by NIR spectroscopy," *Discov. Appl. Sci.*, vol. 7, no. 8, p. 896, Aug. 2025, doi: 10.1007/s42452-025-07580-3.
[28] C. Yan, "A review on spectral data preprocessing techniques for machine learning and quantitative analysis," *iScience*, vol. 28, no. 7, p. 112759, Jul. 2025, doi: 10.1016/j.isci.2025.112759.
[29] J. Torniainen, I. O. Afara, M. Prakash, J. K. Sarin, L. Stenroth, and J. Töyräs, "Open-source python module for automated preprocessing of near infrared spectroscopic data," *Anal. Chim. Acta*, vol. 1108, pp. 1–9, Apr. 2020, doi: 10.1016/j.aca.2020.02.030.
[30] E. Pamukçu, "Efficient Wavelength Selection for Limited Near-Infrared Spectral Data via Genetic Algorithm and Hybrid Regression," *J. Chemom.*, vol. 39, no. 3, Mar. 2025, doi: 10.1002/cem.70015.
[31] X.-H. Ma, Z.-G. Chen, S. Liu, J.-M. Liu, and X. Tian, "Wavelength selection method for near-infrared spectroscopy based on the combination of mutual information and genetic algorithm," *Talanta*, vol. 286, p. 127573, May 2025, doi: 10.1016/j.talanta.2025.127573.
[32] H. Xue, X. Xu, and Y. Yang, "Variety classification and identification of corn seeds based on hyperspectral imaging technology," *IET Conf. Proc.*, vol. 2022, no. 18, pp. 174–177, Dec. 2022, doi: 10.1049/icp.2022.3219.
[33] J. H. Holland, *Adaptation in Natural and Artificial Systems*. The MIT Press, 1992. doi: 10.7551/mitpress/1090.001.0001.
[34] L. Breiman, "Random Forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, Oct. 2001, doi: 10.1023/A:1010933404324.
[35] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of ICNN'95 - International Conference on Neural Networks*, Boston, MA: IEEE, 2011, pp. 1942–1948. doi: 10.1109/ICNN.1995.488968.
[36] D. H. Wolpert, "Stacked generalization," *Neural Networks*, vol. 5, no. 2, pp. 241–259, Jan. 1992, doi: 10.1016/S0893-6080(05)80023-1.
[37] L. A. Zadeh, "Fuzzy sets," *Inf. Control*, vol. 8, no. 3, pp. 338–353, Jun. 1965, doi: 10.1016/S0019-9958(65)90241-X.
[38] J. H. Friedman, "Greedy function approximation: A gradient boosting machine.," *Ann. Stat.*, vol. 29, no. 5, Oct. 2001, doi: 10.1214/aos/1013203451.
[39] J. Ssali Nantongo, E. Serunkuma, G. Burgos, M. Nakitto, F. Davrieux, and R. Ssali, "Machine learning methods in near infrared spectroscopy for predicting sensory traits in sweetpotatoes," *Spectrochim. Acta*

Part A Mol. Biomol. Spectrosc., vol. 318, p. 124406, Oct. 2024, doi: 10.1016/j.saa.2024.124406.

[40] M. J. Yacomelo Hernández *et al.*, "Application of Vis–NIR Spectroscopy and Machine Learning for Assessing Soil Organic Carbon in the Sierra Nevada de Santa Marta, Colombia," *Sustainability*, vol. 18, no. 1, p. 513, Jan. 2026, doi: 10.3390/su18010513.

[41] X. Hao, Z. Chen, S. Yi, and J. Liu, "Application of improved Stacking ensemble learning in NIR spectral modeling of corn seed germination rate," *Chemom. Intell. Lab. Syst.*, vol. 243, p. 105020, Dec. 2023, doi: 10.1016/j.chemolab.2023.105020.

[42] O. Khater, A. Khater, A. S. Al-Nasr, S. Abozyd, B. Mortada, and Y. M. Sabry, "Advancing near-infrared spectroscopy: A synergistic approach through Bayesian optimization and model stacking," *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.*, vol. 318, p. 124492, Oct. 2024, doi: 10.1016/j.saa.2024.124492.

[43] N. Veček, M. Črepinšek, and M. Mernik, "On the influence of the number of algorithms, problems, and independent runs in the comparison of evolutionary algorithms," *Appl. Soft Comput.*, vol. 54, pp. 23–45, May 2017, doi: 10.1016/j.asoc.2017.01.011.

[44] D. Rey and M. Neuhäuser, "Wilcoxon-Signed-Rank Test," in *International Encyclopedia of Statistical Science*, Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 1658–1659. doi: 10.1007/978-3-642-04898-2_616.

[45] G. E. Gignac and E. T. Szodorai, "Effect size guidelines for individual differences researchers," *Pers. Individ. Dif.*, vol. 102, pp. 74–78, Nov. 2016, doi: 10.1016/j.paid.2016.06.069.

[46] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in *Proceedings of the 31st International Conference on Neural Information Processing Systems*, Red Hook, New York: USA: Curran Associates Inc, 2017, pp. 4768–4777.

[47] S. Rezaei, J. Buitink, and F. R. Hay, "Seed responses to change in ambient humidity and the consequences for storability," *J. Stored Prod. Res.*, vol. 109, p. 102477, Dec. 2024, doi: 10.1016/j.jspr.2024.102477.

[48] O. Matsuda and Y. Ohara, "Last-percent improvement in eligibility rates of crop seeds based on quality evaluation using near-infrared imaging spectrometry," *PLoS One*, vol. 18, no. 9, p. e0291105, Sep. 2023, doi: 10.1371/journal.pone.0291105.

[49] L. Feng, T. Hou, B. Wang, and B. Zhang, "Assessment of rice seed vigour using selected frequencies of electrical impedance spectroscopy," *Biosyst. Eng.*, vol. 209, pp. 53–63, Sep. 2021, doi: 10.1016/j.biosystemseng.2021.06.011.

[50] D. Yang, Y. Zhou, Y. Jie, Q. Li, and T. Shi, "Non-destructive detection of defective maize kernels using hyperspectral imaging and convolutional neural network with attention module," *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.*, vol. 313, p. 124166, May 2024, doi: 10.1016/j.saa.2024.124166.

[51] M. Masoudi and R. Khodabakhshian, "Genetic algorithm-optimized PLS for detecting adulteration in cinnamon powder via FT-IR spectroscopy," *Expert Syst. Appl.*, vol. 290, p. 128522, Sep. 2025, doi: 10.1016/j.eswa.2025.128522.

[52] A. Azedou, A. Amine, I. Kisekka, and S. Lahssini, "Genetic algorithm optimization of ensemble learning approach for improved land cover and land use mapping: Application to Talassemtane National Park," *Ecol. Indic.*, vol. 177, p. 113776, Aug. 2025, doi: 10.1016/j.ecolind.2025.113776.

[53] N. Singh *et al.*, "Integrating NIR spectroscopy with machine learning and heuristic algorithm-assisted wavelength selection algorithms for protein content prediction in rice bean (Vigna umbellata L.)," *Food Humanit.*, vol. 3, p. 100399, Dec. 2024, doi: 10.1016/j.foohum.2024.100399.

[54] Y. Wang, L. Xing, H.-J. He, J. Zhang, K. W. Chew, and X. Ou, "NIR sensors combined with chemometric algorithms in intelligent quality evaluation of sweetpotato roots from 'Farm' to 'Table': Progresses, challenges, trends, and prospects," *Food Chem. X*, vol. 22, p. 101449, Jun. 2024, doi: 10.1016/j.fochx.2024.101449.