Agricultural Science; Vol. 7, No. 1; 2025
ISSN 2690-5396 E-ISSN 2690-4799
https://doi.org/10.30560/as.v7n1p84

Study On the Detection of Dry Matter in Silage Corn Feed Based on Near Infrared Spectroscopy

Changfeng Shao[1,2] & Wenjing Zhu[1,2]

[1] School of Agricultural Engineering, Jiangsu University, Zhenjiang, China

[2] High-tech Key Laboratory of Agricultural Equipment and Intelligence, Jiangsu Province, China

Correspondence: Wenjing Zhu, School of Agricultural Engineering, Jiangsu University, Zhenjiang, Jiangsu, China. Tel: 86-153-5388-4867. E-mail: zwj0410@foxmail.com

Received: March 7, 2025 Accepted: April 10, 2025 Online Published: April 12, 2025

The research is financed by Key Laboratory of Modern Agricultural Equipment and Technology Project (MAET202322) and National Natural Science Foundation of China Youth Fund Project (61901194).

Abstract

This study explored the application of a portable near-infrared (NIR) spectrometer for analyzing silage corn feed quality, specifically focusing on developing a quantitative detection model for dry matter content. Spectral data were collected within the 855-1890 nm range using a portable NIR spectrometer, and the dataset was partitioned into calibration and prediction sets using the SPXY algorithm. An Extreme Learning Machine (ELM) model optimized by Particle Swarm Optimization (PSO) was employed for modeling. Five preprocessing methods were evaluated: Moving Average Filter (MAF), Savitzky-Golay Filter (SGF), Multiplicative Scatter Correction (MSC), Standard Normal Variate (SNV) transformation, and First Derivative (FD). To enhance model performance, feature wavelengths were selected using three methods: Bootstrap Soft Shrinkage (BOSS), Competitive Adaptive Reweighted Sampling (CARS), and Iterative Retained Information Variable (IRIV). The optimal model combining SNV preprocessing with BOSS feature selection achieved a prediction correlation coefficient (R_p) of 0.8708 and Root Mean Square Error of Prediction (RMSEP) of 0.6802. These results demonstrate the potential of portable NIR spectroscopy for rapid dry matter content determination in silage corn feed.

Keywords: near-infrared spectroscopy, silage corn, dry matter content, machine learning optimization

1. Introduction

As one of the most cultivated and productive crops in the world, maize is not only an important link in the human food supply, but also a core raw material for livestock rearing(García-Chávez et al., 2020). In aquaculture, maize silage plays a decisive role as a key feed resource to improve productivity and economic efficiency(Ribas et al., 2021). Silage technology maintains and enhances the nutritional value of the feed by creating anaerobic fermentation conditions (Fuglestad et al., 2021), which maximizes the use of plant nutrients and lays the foundation for the sustainable progress of the farming industry (Yin et al., 2024). Straw dry matter content is a basic and critical indicator in the silage process. Ideal storage results depend on the rapid detection and adjustment of the dry matter ratio. However, conventional chemical analysis is time-consuming and labor-intensive, and a breakthrough is needed(Büning-Pfaue, 2003) . Therefore, a rapid and non-destructive way to monitor the dry matter indexes in silage maize stover not only improves the speed of inspection and reduces the running cost, but also provides an important basis for establishing an efficient analytical system and optimizing feed utilization.

Several studies have been conducted on the application of NIRS technology in agriculture.(Wu et al., 2005) developed a NIRS calibration model to accurately predict the cellulose content of corn stover by partial least squares regression (PLS).(Aureli et al., 2017) developed a NIRS calibration method for predicting the concentration of total phosphorus, phytate-P, and protein in feeds based on 14 plant samples from around the world. calibration method.(Ma et al., 2019) developed a high-precision model for predicting moisture, crude protein and soluble sugar content of mulberry leaves using a handheld NIRS. (Zhang et al., 2019)constructed a high-precision prediction model for moisture content by spectral analysis of 208 corn stover silage samples. These studies provide an important theoretical foundation and technical reference for this research. However, although the feasibility of NIR detection in silage nutrient quality testing has been verified, there are fewer studies on the

index of dry matter of silage corn feed, especially in the pre-processing methods, wavelength selection methods and modeling algorithms have not yet been studied in depth. In addition, silage corn feed generally requires on-site field testing, which has certain requirements for portability.

The aim of this study is to use a portable near-infrared spectrometer to collect the spectral data of silage corn feed, and to establish a high-precision dry matter prediction model for silage corn feed by combining multiple preprocessing methods and wavelength selection algorithms. The feature wavelengths were screened by Competitive Adaptive Resampling Algorithm (CARS), Bootstrap Soft Shrinkage Algorithm (BOSS), and Iterative Retention of Information Variables Algorithm (IRIV), and the optimization algorithm was used to preferentially select the modeling parameters, and the optimal preprocessing method, wavelength selection strategy, and modeling parameters were ultimately determined in order to establish a dry matter content prediction model with high prediction accuracy. This not only helps to improve the efficiency and accuracy of feed quality testing, but also provides technical support for the sustainable development of the farming industry.

2. Method

2.1 Test Materials

The silage corn feed samples used in this experiment were collected in December 2024 from the silage corn fermentation pond of a dairy farm located in Tai'an City, Shandong Province. A total of 270 samples were gathered. Immediately after collection, the samples were placed into vacuum bags and vacuum-sealed. They were then transported and stored under refrigerated conditions to ensure their preservation.

2.2 Near Infrared Spectroscopy Data Acquisition

The NIR spectral acquisition system comprises an NIR spectrometer, light source, acquisition lens, and computer. The system utilizes the NIR-NT compact fiber-optic near-infrared spectrometer, manufactured by INSION in Germany, with a nominal wavelength range of 900-1700 nm. However, due to the manufacturing process, the actual wavelength range is 855-1890 nm. Prior to testing, the system requires a 30-minute preheating process to eliminate baseline drift. Subsequently, the spectrometer's focal length must be adjusted to ensure optical accuracy. The system's exposure time was set to 20 ms for the experimental parameters. During data acquisition, samples should be evenly distributed in the container to ensure they are within the aperture's coverage area. Using near-infrared spectral data acquisition software, spectral data within the 855-1890 nm wavelength range were collected in this study.

2.3 Measurement of Dry Matter Content

The dry matter content was measured following the standard method outlined in GB/T 6435-2014 "Determination of Moisture in Feed" using the oven-drying technique. The procedure is described as follows:

A beaker was first dried at 103 °C for 30 minutes in a drying oven and cooled to room temperature in a desiccator. Its mass (m_1) was then recorded. Subsequently, 5 g of silage corn feed was weighed into the pre-dried beaker and dried at 103 °C for 4 hours. During the drying process, the sample was periodically monitored until a constant weight was achieved, defined as a difference of less than 0.1% between consecutive measurements. After drying, the beaker and dried sample were allowed to cool to room temperature in a desiccator, and the total mass (m_2) was recorded. The dry matter content was calculated using Equation (1).

$$X = \frac{m_3 - m_1}{m_2} \times 100\% \quad (1)$$

In the Equation, m_1 represents the initial weight of the beaker (g), m_2 is the weight of the sample (g), and m_3 refers to the weight of the beaker and the sample after drying (g).

The statistical analysis results for the dry matter indices of the test samples are presented in Table 2.

Table 1. Statistical table of dry matter content (%) of silage maize forage

Quality Indicators	Minimum	Maximum	Mean	Standard Deviation
DM	28.63	35.17	31.47	2.30

2.4 Spectral Data Preprocessing

The acquisition of near-infrared spectral data is inevitably affected by various factors, including stray light, signal-to-noise interference, diffuse light reflection, and spectral baseline shifts. To eliminate these interferences and establish a reliable prediction model, preprocessing of the spectral data is essential. In this study, five preprocessing methods were applied to the spectral data: moving average filtering (MAF), Savitzky-Golay filtering (SGF), multiplicative scatter correction (MSC), standard normal variate (SNV) transformation, and first-order derivative (FD) processing.

2.5 Characteristic Wavelength Selection

The selection of wavelengths for near-infrared (NIR) spectroscopy is crucial for reducing model complexity, improving prediction accuracy, and enhancing scanning speed. By identifying wavelengths closely related to the target analyte, data dimensionality can be reduced, and the model structure can be optimized. This, in turn, minimizes the model size and enhances prediction performance. Additionally, selecting characteristic wavelengths helps eliminate noise and interference, thereby increasing the model's robustness. In this study, Competitive Adaptive Reweighted Sampling (CARS), Bootstrap Soft Shrinkage (BOSS), and Iterative Retention of Informative Variables (IRIV) methods were employed to select the key wavelengths.

2.6 Spectral Modeling Methods

Extreme Learning Machine (ELM) is a machine learning approach derived from the Single Hidden Layer Feedforward Neural Network (SLFN). It is effective in handling complex nonlinear fitting problems and is particularly suitable for large-scale training datasets (Wang et al., 2018) . Based on the principle of structural risk minimization, ELM enhances data fitting capabilities by randomly mapping the linear output function to achieve efficient nonlinear transformations in high-dimensional feature spaces. In this study, the particle swarm optimization (PSO) algorithm was employed to optimize the key parameters of the ELM model.

Particle Swarm Optimization (PSO) is an evolutionary algorithm inspired by the collective intelligence of groups. It efficiently solves complex optimization problems by simulating the collaborative search mechanisms among individuals in a population(Huang et al., 2021) . The PSO algorithm utilizes particle position and velocity as core parameters, conducting searches through dynamic updates of the individual best and global best solutions, ultimately converging on the global optimal solution. Compared to traditional optimization algorithms, PSO is characterized by its simplicity, high efficiency, and ease of parallel implementation. It has been widely applied in function optimization and parameter tuning in machine learning. In this study, PSO was used to optimize the hyperparameters of the ELM model, including the number of neurons in the hidden layer and the weights, thereby significantly improving the model's generalization performance and prediction accuracy.

2.7 Model Evaluation Metrics

In model evaluation, the key metrics include the calibration set correlation coefficient (R_c), the calibration set root mean square error (*RMSEC*), the prediction set correlation coefficient (R_p), and the prediction set root mean square error (*RMSEP*). The closer the values of R_c and R_p are to 1, the stronger the model's explanatory and predictive abilities. Conversely, lower values of *RMSEC* and *RMSEP* indicate higher model precision.

2.8 Data Processing Methods

Spectral preprocessing and dataset partitioning were performed using NIRSA 5.3 software. Wavelength selection and model development were conducted in MATLAB R2022a. The IRIV algorithm for feature wavelength extraction was implemented using the libPLS 1.98 toolbox (Li et al., 2018).

3. Results

3.1 Dataset Division

The dataset was divided using the SPXY algorithm. A total of 189 samples (70% of the dataset) were used as the calibration set, while 81 samples (30% of the dataset) were assigned to the test set.

3.2 Spectral Preprocessing Analysis

To analyze the effects of different near-infrared (NIR) spectral preprocessing methods on the prediction model for the dry matter content of silage corn feed, five preprocessing techniques—Savitzky-Golay filtering (SGF), moving average filtering (MAF), multiplicative scatter correction (MSC), standard normal variate (SNV), and first-order derivative (FD)—were applied to the raw spectral data. The original spectra and the preprocessed spectral data are shown in Figure 1.

As illustrated in Figure 1, after SGF and MAF preprocessing, the spectral curves became smoother, noise fluctuations were significantly reduced, and the curves were largely consistent with the original spectral profile. Following MSC and SNV preprocessing, the spectral curves demonstrated higher concentration, scattering effects were effectively reduced, and spectral features became more prominent. After FD preprocessing, small feature variations in the spectral data were revealed, characteristic peaks were significantly enhanced, and the signal-to-noise ratio was improved to a certain extent.

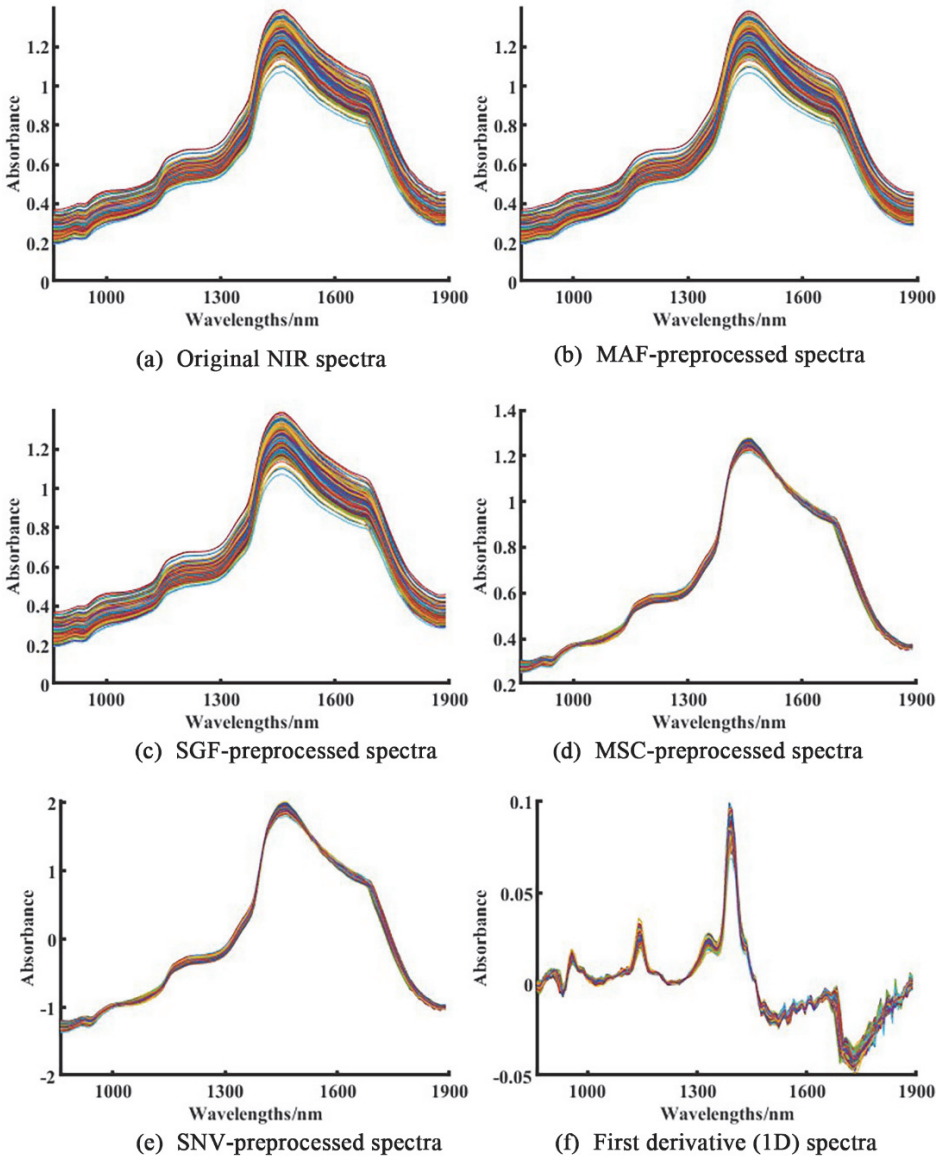

Figure 1. Original Spectra and Preprocessed Spectra

Table 2 shows the performance of PSO-ELM prediction models developed using different preprocessing methods across the full wavelength spectrum. Models developed using raw spectral data yielded a prediction set correlation coefficient (R_p)and *RMSEP* values of 0.8228 and 0.7774, respectively, indicating the feasibility of direct modeling despite the potential interference from noise and scattering effects.

Models based on MAF preprocessing exhibited slightly diminished prediction set performance (R_p= 0.8179, *RMSEP* = 0.7871). Although MAF effectively smooths spectral data and reduces high-frequency noise, it may simultaneously remove critical spectral information, resulting in reduced prediction accuracy. In contrast, SGF preprocessing improved prediction set performance (R_p= 0.8375, *RMSEP* = 0.7472), successfully mitigating overfitting tendencies in the original data and enhancing the model's generalization ability.

The MSC and SNV preprocessing techniques produced distinct outcomes, with SNV delivering superior prediction set performance (R_p= 0.8413, *RMSEP* = 0.7680), outperforming both MSC-processed and raw spectral data. The superior performance of SNV likely results from its standardized normalization process, which effectively reduces scattering-induced variability, minimizes bias, and improves the reliability of the model.

First-order derivative (FD) preprocessing significantly improved calibration set performance but paradoxically reduced prediction set performance R_p= 0.8132, *RMSEP* = 0.7687). This inverse relationship suggests that while FD enhances the resolution of spectral features, it also amplifies noise components and increases the risk of overfitting, leading to diminished predictive performance.

Table 2. Modeling Performance Based on Full-Wavelength Spectral Data

Preprocessing	Modeling Methods	Calibration Set		Test Set	
		R_C	*RMSEC*	R_P	*RMSEP*
RAW	PSO-ELM	0.8817	0.7309	0.8228	0.7774
MAF		0.8658	0.7752	0.8179	0.7871
SGF		0.872	0.7585	0.8375	0.7472
MSC		0.8754	0.7444	0.8369	0.778
SNV		**0.8723**	**0.7529**	**0.8413**	**0.768**
1D		0.8873	0.7529	0.8132	0.7687

3.3 Characteristic Wavelength Extraction

3.3.1 Competitive Adaptive Reweighted Sampling

The Competitive Adaptive Reweighted Sampling (CARS) algorithm is a wavelength selection method based on Darwinian evolutionary theory(Li et al., 2009) , designed to optimize wavelength selection through iterative statistical analysis. This algorithm integrates adaptive reweighted sampling (ARS) with partial least squares (PLS) to identify wavelengths with larger absolute regression coefficients during each iteration. Through multiple iterations, the CARS algorithm searches for the subset of wavelengths that minimizes the root mean square error of cross-validation (*RMSEC*V), thereby identifying the most representative set of optimal wavelengths.

In this study, the optimal principal component of the PLS model is determined by Monte Carlo cross-validation (MCCV) and input into the competitive adaptive reweighted sampling (CARS) algorithm to realize the feature wavelength screening. The specific process(Liu et al., 2023) is as follows: first, the spectral data are centered and preprocessed to eliminate baseline bias; based on the MCCV method, the model stability is assessed and the optimal principal component fraction is determined to be 15 by 50 independent random samples (80% of samples taken each time) combined with a five-fold cross-validation; subsequently, the principal component fraction is inputted into the CARS algorithm as a parameter to carry out the Wavelength screening: in 500 iterations of optimization, a subset of candidate variables is generated by Monte Carlo sampling, low regression coefficient variables are gradually eliminated using exponential decreasing function (EDF), and high weighted wavelengths are competitively retained in combination with adaptive reweighting sampling (ARS). Finally, the optimal wavelength subset with the smallest *RMSEC*V is screened by 50 independent modeling validations using the 50-fold cross-validation root mean square error (*RMSEC*V) as a criterion.

As shown in Figure 2, Figure 2-a depicts the relationship between the number of samples and the number of sampled variables, which overall exhibits a smooth decreasing trend. This indicates that a large number of irrelevant variables are gradually eliminated as the number of samples increases. Figure 2-b illustrates the changing relationship between the number of samples and the root mean square error of cross-validation (*RMSEC*V). At the early stages, the *RMSEC*V decreases rapidly, consistent with the trend observed in Figure 2-a. The *RMSEC*V reaches its minimum value at 21 samples, after which it begins to increase.

Figure 2-c shows the relationship between the regression coefficients of the wavelength variables and the number of samples. The blue dashed line marks the location of the minimum *RMSEC*V value, which corresponds to Figure 2-b, indicating the optimal subset of wavelengths. Following screening by the CARS algorithm, a total of 24 characteristic wavelengths were selected, accounting for 18.8% of the overall wavelengths. The distribution of the characteristic wavelength variables obtained through the CARS algorithm is shown in Figure 3.

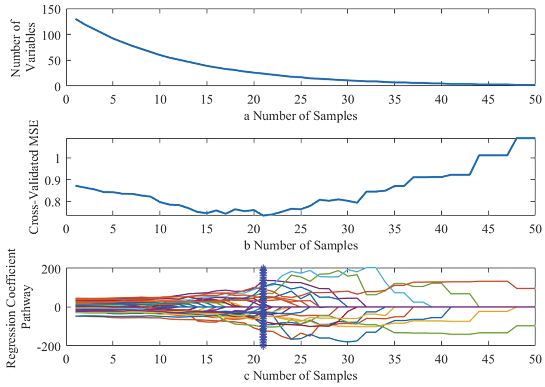

Figure 2. CARS screening results

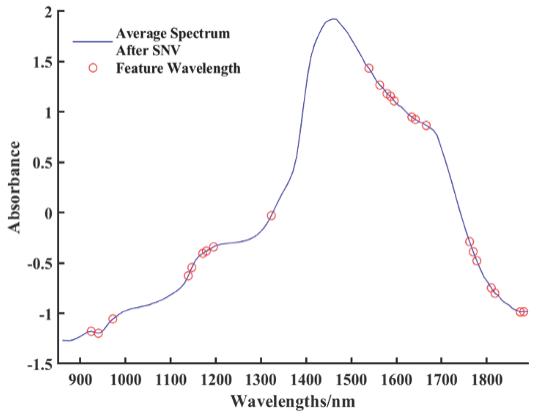

Figure 3. CARS-Selected Wavelength Distribution

3.3.2 Bootstrap Optimized Subset Selection

The BOSS algorithm (Bootstrap Optimized Subset Selection) is an iterative optimization method for feature wavelength screening. Its core concept is to combine bootstrap sampling and dynamic assignment strategies to identify the optimal sub-model by generating random feature subsets and employing Partial Least Squares (PLS) modeling with Root Mean Square Error of Cross-Validation (*RMSECV*) as the criterion (Deng et al., 2016). During the iterative process, regression coefficients are utilized to construct weight vectors, dynamically adjusting the probability of variable selection.

This approach prioritizes the retention of highly contributory variables through weighted bootstrap selection, ultimately selecting the variable combination with the lowest *RMSECV* as the optimal solution while balancing data diversity and model stability.

As a feature selection algorithm based on model cluster analysis, this section highlights the results of its practical application. Figure 4illustrates the relationship between the *RMSECV* and the number of iteration rounds during cross-validation using the BOSS algorithm. The results indicate that the *RMSECV* exhibits a dynamic trend of first decreasing and then increasing as the number of iterations progresses. At the 3rd iteration round, the *RMSECV* reaches its minimum value, marking the precise localization of the optimal variable combination. Further increasing the number of iteration rounds results in a rebound of the *RMSECV*. This rebound occurs because key information variables are unintentionally eliminated. The distribution of feature wavelengths selected by the BOSS algorithm across the entire wavelength range is shown in the accompanying figure, with the selected wavelengths accounting for 26.6% of the total wavelengths. The distribution of the characteristic wavelength variables obtained through the CARS algorithm is shown in Figure 5.

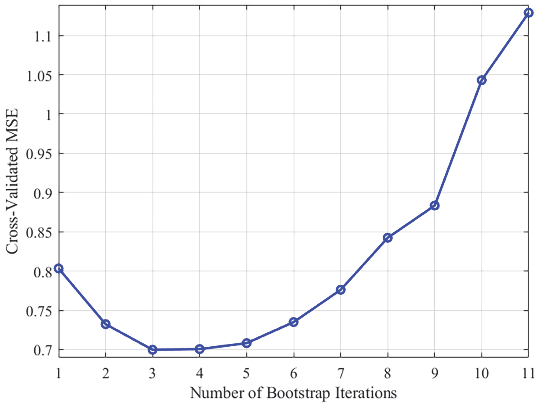

Figure 4. BOSS screening results

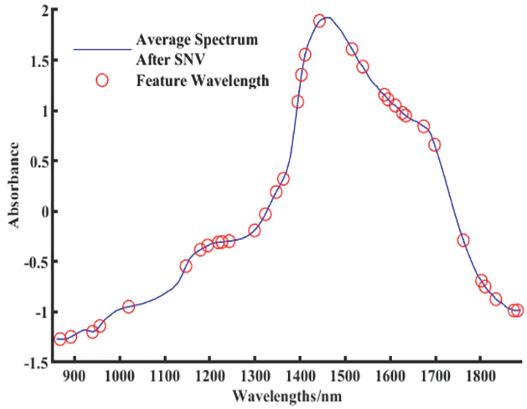

Figure 5. BOSS-Selected Wavelength Distribution

3.3.3 Iterative Retained Information Variable

The Iterative Retained Information Variable (IRIV) method is a feature selection algorithm based on binary random matrix rearrangement filtering(Yun et al., 2014). This method classifies variables into four categories: strong information, weak information, no information, and interfering variables. Through multiple rounds of iterations, the algorithm gradually filters out the optimal subset of features, thereby reducing the impact of redundant and interfering variables on model performance.

As shown in Figure 6, the IRIV algorithm exhibits a gradual reduction in the number of variables during the feature extraction process for the spectral data of silage corn feed. The process begins with an initial screening of global variables, followed by a reverse elimination mechanism to evaluate the contribution of each variable to the model's predictive performance. Through this stepwise elimination of low-contributing variables, 23 key feature wavelengths—accounting for approximately 18% of the original feature space—were ultimately retained. During the backward elimination process, the root mean square error of cross-validation (*RMSECV*) is used as the evaluation metric, and the model is re-modeled at each step of variable removal to ensure that overall model performance is not significantly degraded. The distribution of the characteristic wavelength variables obtained through the CARS algorithm is shown in Figure 7.

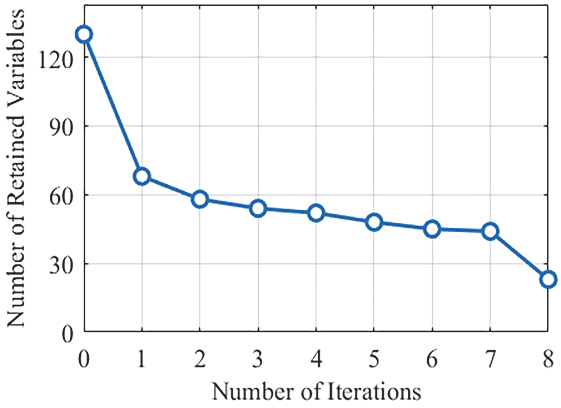

Figure 6. IRIV screening results

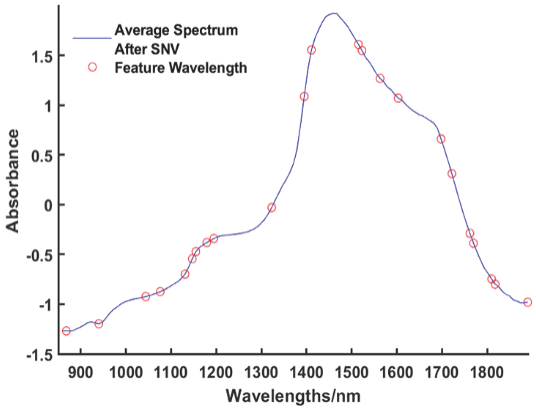

Figure 7. IRIV-Selected Wavelength Distribution

3.4 Modeling Based on Characteristic Wavelengths

Table 3 summarizes the performance of prediction models built using different characteristic wavelengths. During the construction of the ELM model, the modeling parameters were optimized using the PSO algorithm. From the table, it can be observed that the CARS, BOSS, and IRIV algorithms effectively reduce the number of wavelengths required for modeling. Consequently, the input parameters for the prediction model are reduced, which enhances the real-time performance of online detection systems.

The correlation coefficients (R_p) of the model prediction sets based on the characteristic wavelengths all exceed 0.85, demonstrating superior performance compared to models built using the original data. Among the tested models, the SNV-BOSS-PSO-ELM model demonstrated the best performance, achieving an R_p of 0.8708 and an *RMSEP* of 0.6802.

Table 3. Modeling Performance Based on Characteristic Wavelengths

Modeling Methods	Wavelength Selection Methods	Number of Features	Calibration Set		Test Set	
			R_c	*RMSEC*	R_P	*RMSEP*
PSO-ELM	CARS	24	0.8995	0.687	0.8664	0.6895
	BOSS	34	**0.9142**	**0.6306**	**0.8708**	**0.6802**
	IRIV	23	0.9027	0.666	0.867	0.6927

3.5 Experimental Summary

In this experiment, a quantitative detection model for the dry matter content of silage corn feed using portable near-infrared spectroscopy was developed, with silage corn feed as the research object. The results indicate that spectral data preprocessed using the Standard Normal Variate (SNV) method outperformed the original data and data processed with other preprocessing methods. During the PSO-ELM modeling process, the model constructed using the characteristic wavelengths selected by the BOSS algorithm demonstrated the best prediction performance.

By analyzing the fitted scatter plots (Figures 8–10) of the prediction results from the three models based on different wavelength selection algorithms for the dry matter content of silage corn feed, it is evident that the SNV-BOSS-PSO-ELM model exhibited a better fitting effect compared to the other models. Its prediction results are more concentrated, which aligns with the data presented in the table. These results confirm that the SNV-BOSS-PSO-ELM model provides better predictive accuracy for the dry matter content of silage corn feed than the other models.

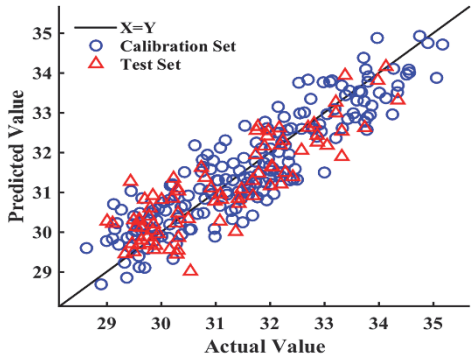

Figure 8. Scatter Plot of SNV-CARS-PSO-ELM Model Performance

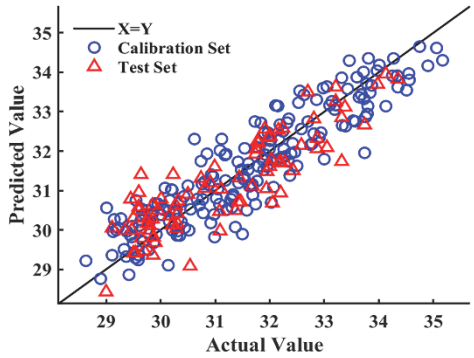

Figure 9. Scatter Plot of SNV-BOSS-PSO-ELM Model Performance

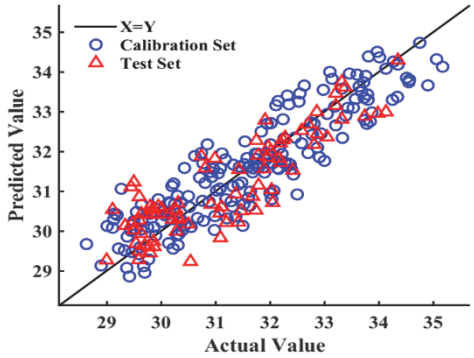

Figure 10. Scatter Plot of SNV-IRIV-PSO-ELM Model Performance

4. Discussion

In this study, 270 silage maize feed samples were analyzed, focusing on the dry matter physicochemical values. The raw spectral data were preprocessed using five methods: SGF, MAF, MSC, SNV, and FD. Based on an analysis of model performance, the most suitable preprocessing method was selected for subsequent modeling. Characteristic wavelengths were selected using the CARS, BOSS, and IRIV algorithms, while prediction models were constructed based on full wavelengths and characteristic wavelengths using the ELM algorithm. The PSO algorithm was employed to optimize the modeling parameters during this process. The main conclusions are summarized as follows:

(1) By comparing the spectral preprocessing results of SGF, MAF, MSC, SNV, and FD and combining them with the quantitative evaluation of the PSO-ELM model, it was found that SNV preprocessing yielded the best performance in the quantitative dry matter model. For the calibration set, the correlation coefficient (R) was 0.8723, with a root mean square error (RMSE) of 0.7529. For the prediction set, the R was 0.8413, and the RMSE was 0.768. These results indicate that SNV preprocessing effectively reduces the effects of spectral scattering, baseline drift, and background interference, thereby enhancing the consistency and reliability of the spectral data and significantly improving the accuracy and stability of the model.

(2) The characteristic wavelength screening focused on the range of 1100–1800 nm, which encompasses the core absorption regions of the primary chemical components in dry matter, including the vibrational absorption regions of the O-H bond in water and the C-H bond in cellulose and hemicellulose. Among the algorithms, BOSS demonstrated a broader and more balanced wavelength distribution, comprehensively covering key intervals and enhancing the model's generalization ability. CARS concentrated on key absorption regions, producing results with strong relevance to the target, though slightly limited in the breadth of information covered. IRIV, on the other hand, covered a wider wavelength range and captured the features across multiple intervals but included a higher number of redundant variables, which negatively impacted model efficiency.

(3) In the evaluation of the models built using characteristic wavelengths, the BOSS-based PSO-ELM model demonstrated the best overall performance. For the quantitative dry matter model, the BOSS-based model achieved an R of 0.9142 and an RMSE of 0.6306 in the calibration set, while in the prediction set, the R was 0.8708 and the RMSE was 0.6802. These results outperformed other algorithms, indicating that BOSS effectively balances model prediction performance and improves the correlation coefficient while maintaining high accuracy.

Acknowledgments

The research is financed by Key Laboratory of Modern Agricultural Equipment and Technology Project (MAET202322) and National Natural Science Foundation of China Youth Fund Project (61901194).

References

Aureli, R., Ueberschlag, Q., Klein, F., Noel, C., & Guggenbuhl, P. (2017). Use of near infrared reflectance spectroscopy to predict phytate phosphorus, total phosphorus, and crude protein of common poultry feed ingredients. *Poultry Science, 96*(1), 160–168. https://doi.org/10.3382/ps/pew214

Büning-Pfaue, H. (2003). Analysis of water in food by near infrared spectroscopy. *Food Chemistry, 82*(1), 107–115. https://doi.org/10.1016/S0308-8146(02)00583-6

Deng, B. C., Yun, Y. H., Cao, D. S., Yin, Y. L., Wang, W. T., Lu, H. M., Luo, Q. Y., & Liang, Y. Z. (2016). A bootstrapping soft shrinkage approach for variable selection in chemical modeling. *Analytica Chimica Acta, 908*, 63–74. https://doi.org/10.1016/j.aca.2016.01.001

Fuglestad, E. M., Vik, J., Finstad, T., & Soraa, R. A. (2021). Compressed growth—The transforming power of the round bale technology. *Journal of Rural Studies, 84*, 174–179. https://doi.org/10.1016/j.jrurstud.2021.04.005

García-Chávez, I., Meraz-Romero, E., Castelán-Ortega, O., Zaragoza Esparza, J., Osorio Avalos, J., Robles Jimenez, L. E., & González-Ronquillo, M. (2020). Corn silage, meta-analysis of the quality and yield of different regions in the world. *Preprints*. https://doi.org/10.20944/preprints202010.0094.v1

Huang, W. C., Liu, H. Y., Zhang, Y., Mi, R. W., Tong, C. G., Xiao, W., & Shuai, B. (2021). Railway dangerous goods transportation system risk identification: Comparisons among SVM, PSO-SVM, GA-SVM and GS-SVM. *Applied Soft Computing, 109*, Article 107541. https://doi.org/10.1016/j.asoc.2021.107541

Li, H. D., Liang, Y. Z., Xu, Q. S., & Cao, D. S. (2009). Key wavelengths screening using competitive adaptive reweighted sampling method for multivariate calibration. *Analytica Chimica Acta, 648*(1), 77–84. https://doi.org/10.1016/j.aca.2009.06.046

Li, H. D., Xe, Q. S., & Liang, Y. Z. (2018). libPLS: An integrated library for partial least squares regression and linear discriminant analysis. *Chemometrics and Intelligent Laboratory Systems, 176*, 34–43. https://doi.org/10.1016/j.chemolab.2018.03.003

Liu, L. H., Zareef, M., Wang, Z., Li, H. Q., Chen, Q. S., & Ouyang, Q. (2023). Monitoring chlorophyll changes during Tencha processing using portable near-infrared spectroscopy. *Food Chemistry, 412*, Article 135505. https://doi.org/10.1016/j.foodchem.2023.135505

Ma, Y., Zhang, G. Z., & Rita-Cindy, S. A. A. (2019). Quantification of water, protein and soluble sugar in mulberry leaves using a handheld near-infrared spectrometer and multivariate analysis. *Molecules, 24*(24), Article 4439. https://doi.org/10.3390/molecules24244439

Ribas, T. M. B., Neumann, M., Horst, E. H., Cristo, F. B., Moresco, E. M., & de Almeida, E. R. (2021). Effect of inoculants in corn silage on dry matter digestibility, ingestive behavior, performance and carcass of heifers. *Semina: Ciências Agrárias, 42*(3), 1271–1285. https://doi.org/10.5433/1679-0359.2021v42n3p1271

Wang, M. C., Zhang, X. Y., Niu, X. F., Wang, F. Y., Xu, Z. S., Zhang, X. Q., Chen, S. B., Wang, M. C., & IEEE. (2018, August 19–20). Inversion of heavy metal content in a copper mining area based on extreme learning machine optimized by particle swarm algorithm. In *10th IAPR Workshop on Pattern Recognition in Remote Sensing (PRRS)*, Beijing, China.

Wu, J., Bai, Q. L., Su, S. B., Chen, S. J., Meng, Q. X., & Yan, Y. L. (2005). Near-infrared reflectance spectroscopy analysis of cellulose content in corn stalk. *Chinese Journal of Analytical Chemistry, 33*(10), 1421–1423.

Yin, X. T., Wang, P. F., Yan, Z. Q., Yang, Q. L., Huang, X. Y., & Gun, S. (2024). Effects of whole-plant corn silage on growth performance, serum biochemical indices, and fecal microorganisms in Hezuo pigs. *Animals, 14*(5), Article 662. https://doi.org/10.3390/ani14050662

Yun, Y. H., Wang, W. T., Tan, M. L., Liang, Y. Z., Li, H. D., Cao, D. S., Lu, H. M., & Xu, Q. S. (2014). A strategy that iteratively retains informative variables for selecting optimal variable subset in multivariate calibration. *Analytica Chimica Acta, 807*, 36–43. https://doi.org/10.1016/j.aca.2013.11.032

Zhang, M. Q., Zhao, C., Shao, Q. J., Yang, Z. D., Zhang, X. F., Xu, X. F., & Hassan, M. (2019). Determination of water content in corn stover silage using near-infrared spectroscopy. *International Journal of Agricultural and Biological Engineering, 12*(6), 143–148. https://doi.org/10.25165/j.ijabe.20191206.4914

Copyrights