Journal of Electronic Research and Application, 2026, Volume 10, Issue 2
http://ojs.bbwpublisher.com/index.php/JERA
ISSN Online: 2208-3510
ISSN Print: 2208-3502

Study on pH Value Detection of Silage Corn Feed Based on Micro Near Infrared Spectrometer

Xiao Li*, Wenjing Zhu, Changfeng Shao

Jiangsu University, Zhenjiang, Jiangsu, China

**Author to whom correspondence should be addressed.*

Abstract: Silage corn feed is an important source of roughage for cattle and sheep breeding. Its pH value can quickly reflect the quality of feed and is one of the main indicators for evaluating the fermentation quality of feed. In this study, the pH value of silage corn feed was non-destructively and rapidly detected based on near-infrared spectroscopy analysis technology. The silage corn feed of dairy farms in Tai'an City, Shandong Province was taken as the research object. The spectra of 855~1890 nm of samples were collected based on a miniature near-infrared spectroscopy acquisition system, and the pH value prediction model of silage corn feed was established by combining chemometrics methods. Firstly, a miniature near-infrared spectroscopy acquisition system was introduced. The system was used to collect the near-infrared spectral data of 260 silage corn feed samples. The near-infrared spectral data were preprocessed by standard normal transformation, moving mean filtering, and multiple scattering correction. The wavelength was optimized by algorithms such as elimination of uninformative variables, competitive adaptive reweighted sampling algorithm, and variable iterative space shrinkage. The support vector regression (SVR) model based on snake optimization (SO) was used. The model was established by combining the pH value of the leaching solution of silage corn feed with the pH meter. The results show that the full-band modeling based on multiple scattering correction (MSC) pretreatment is the best, the modeling effect of the characteristic wavelength selected based on variable iterative space shrinkage approach (VISSA) is the best, and the modeling effect of MSC-VISSA-SO-SVR model is the best, and the correction correlation coefficient is 0.990531. The corrected root mean square error is 0.083545, the predicted correlation coefficient is 0.980487, and the predicted root mean square error is 0.127718. The results of this study show that the combination of MSC-VISSA-SO-SVR model based on micro near-infrared spectroscopy acquisition system can provide a reference for non-destructive detection of pH value of silage corn feed by near-infrared spectroscopy.

Keywords: Near infrared spectroscopy; Silage corn feed; pH value; Characteristic wavelength

Online publication: April 3, 2026

1. Introduction

Silage corn feed is made of the whole corn, including corn ears, which is harvested and chopped at the appropriate growth stage, and fermented by anaerobic lactic acid bacteria in a closed environment [1]. Silage corn feed is an important source of roughage for ruminants such as cattle and sheep. It has the advantages of wide source, low cost and simple production. The main indicators for evaluating fermentation quality are pH, organic acid and ammonia nitrogen [2]. The pH value can quickly reflect the quality of silage corn feed, and the pH value range of good corn silage is 3.4 to 4.7. The pH value of corn silage with poor quality is greater than 4.8. Excessive pH will lead to aerobic corruption and reduce the utilization rate of crude protein. It is necessary to accurately and quickly determine the pH value of corn silage.

The traditional method for determining the pH value of silage is to take the feed leaching solution and determine it via a pH meter. Although the method is more accurate, it is time-consuming, laborious and destroys the sample. Near-infrared spectroscopy analysis technology has the advantages of high efficiency and speed, so this technology has been widely used in the field of feed-rapid analysis [3–7]. Bell *et al.* (2018) proved that near infrared spectrum (NIRS) can be used to monitor and evaluate the seasonal changes of nutrient concentrations in different forages [8]. Tian *et al.* (2023) used infrared instruments to collect the spectral data of oat forage silage, combined with partial least squares regression (PLSR) to establish a quantitative analysis model of nutrients in fresh samples and dry samples [9]. Wang *et al.* (2021) used NIRS combined with modified partial least squares (MPLS) to construct near-infrared prediction models for six nutrients of dry matter, crude protein, acid detergent fiber and crude ash in whole-plant corn silage products and verified the advantages and disadvantages of their prediction models [10]. Zhang *et al.* (2019) used multiple scattering correction (MSC) combined with first-order derivative (FD) continuous pretreatment to improve partial least squares to establish a wide range of moisture quantitative model for corn straw silage [11]. The correlation coefficients of the calibration set and the validation set were 0.974 and 0.949, respectively. Liu *et al.* (2007) collected near-infrared diffuse reflectance spectra of fresh samples and straw silage samples after drying and crushing, established models of pH values and fermentation products of samples, collected transmission spectra of extract samples and established models [12]. The correlation determination coefficients of calibration set and validation set were both greater than 0.8. In the aspect of hyperspectral detection of silage corn feed, Zhang *et al.* (2023) used hyperspectral imaging system combined with PLSR model to establish a quantitative detection model of pH value of silage corn feed with different quality [13]. The determination coefficient of prediction set of the optimal algorithm combination was 0.9170, but when collecting hyperspectral images, feed samples need to be placed on an electronically controlled mobile platform, and the collection speed is slow and the time is long. Although some scholars have used near infrared spectroscopy to study the field of silage, there are few reports on the non-destructive detection of pH value of silage corn feed by micro near infrared spectrometer.

In this study, the spectral data of silage corn feed in the range of 855~1890 nm were collected by a miniature near-infrared spectroscopy acquisition system, and different methods were used for spectral preprocessing and different wavelength selection algorithms were used to screen the characteristic wavelengths. At the same time, the pH value of the sample was measured by a pH meter, and a quantitative detection model of pH value based on snake-optimized support vector regression was established.

2. Materials and methods

2.1. Sample preparation

In December 2024, 260 silage corn feed samples were collected from the silage corn fermentation tank of the dairy farm in Tai'an City, Shandong Province. Samples were randomly selected. All samples were packaged in vacuum bags, transported by refrigerated trucks, and stored in refrigerators (**Figure 1**).

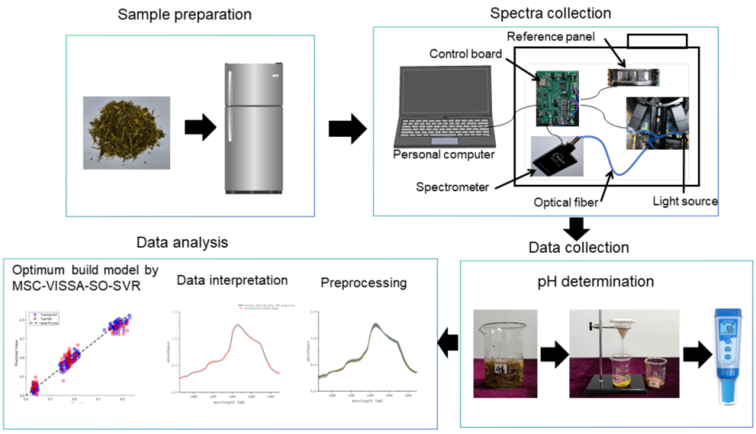

Figure 1. Schematic diagram of pH value modeling process based on miniature near-infrared spectrometer.

2.2. Acquisition system construction

The acquisition system includes a spectral detection module, a detection accessory design and a detection software. The acquisition process is shown in **Figure 1**. The spectral detection module is a miniature near-infrared spectrometer (NIR-NT near-infrared spectrometer produced by Insion, Germany). The actual wavelength range is 855–1890 nm. The resolution is less than 16 nm). The halogen tungsten lamp light source, optical fiber, built-in reference plate and control communication module are composed. The control module controls the light source and built-in reference plate. The communication module sends the converted spectral signal of the spectrometer to the detection software, as shown in **Figure 2**. To ensure the stability of the acquisition system in complex agricultural environments such as farms, the detection accessories are designed, including a protective shell and a sample cup. The base and the protective shell cooperate to form a closed dark box to reduce external light interference. The detection software is based on the Python language, combined with the PyQt5 framework and the Matplotlib library. The near-infrared spectrum detection software is modularly designed. The established silage detection model is imported into the detection software, and the spectral acquisition device is controlled by the USB communication function to complete the collection of the sample spectrum. The collected original spectral data is analyzed by loading the detection model to obtain the predicted value.

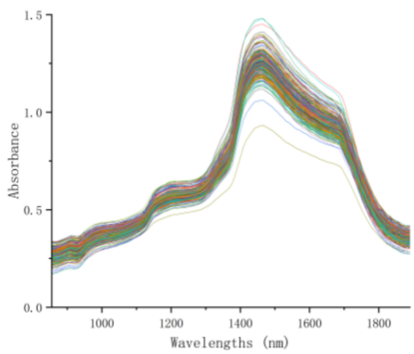

Figure 2. Original near infrared spectra.

2.3. pH determination

The pH value of silage corn feed samples was determined according to the standard method DB15T1458-2018. The specific operation steps are as follows: 10 g silage samples collected by near-infrared spectroscopy were placed in a 200 mL glass beaker, 90 mL distilled water (the mass ratio of sample to distilled water was 1:9) was added, sealed with a plastic wrap and numbered. The samples were fully soaked in distilled water for 30 min and filtered through 4 layers of gauze to obtain the extract (2024) [14]. The PH5F functional pen-type plane pH meter (Shanghai Sanxin, measurement range: -2.00 to 16.00 pH, resolution: 0.0 level) was used to measure the feed extract. Three parallel samples were set for each sample, and the final result was taken as the arithmetic average.

2.4. Collection of spectral data

The acquisition system based on the micro-miniature near-infrared spectrometer was initialized, and the corresponding detection software was launched. The instrument was allowed to preheat for 30 minutes to minimize baseline drift. The operational parameters were configured with a wavelength range of 855–1890 nm, an integration time of 20 ms, and an averaging number of three scans. Prior to sample measurement, dark current spectra and reference spectra from a calibration plate were collected for system calibration. Subsequently, the sample was placed into a sample cup, positioned on the measurement platform, and its spectral data were acquired.

2.5. Stoichiometric analysis methods

In the process of near-infrared spectrum acquisition, due to the unevenness of the sample surface, noise, baseline drift, scattering distortion and feature ambiguity are generated. Three methods, standard normal variate (SNV), moving average filter (MAF) and MSC, are used to preprocess the collected raw data.

In order to improve the efficiency of subsequent data processing, uninformative variables elimination (UVE), competitive adaptive reweighted sampling (CARS) and variable iterative space shrinkage approach (VISSA) are used for wavelength selection. A support vector regression (SVR) model based on snake optimization (SO) optimization was established to establish a quantitative detection model for pH value of silage corn feed. UVE is a variable selection method based on regression coefficients. The core idea is to artificially add a random noise matrix and use the noise value as a threshold to eliminate variables that do not provide information (2018) [15]. As a wavelength selection method based on Darwin's evolutionary theory, CARS uses iterative statistical analysis to optimize the wavelength selection process (2025) [16]. The algorithm uses adaptive reweighted sampling technology, combined with partial least squares regression analysis. In each iteration, the algorithm filters out the wavelengths with larger absolute value of regression coefficient and eliminates the wavelengths with smaller absolute value of regression coefficient. The wavelength subset with the smallest root mean square error of cross-validation was found, which is considered as the optimal wavelength set. Based on the model cluster analysis strategy, VISSA statistically evaluates the performance of the variable space at each optimization step. The weighted binary matrix sampling is used to build sub-models across variable subspaces, and the variable space is reduced in each step and the new variable space takes precedence over the previous space (2022) [17].

The performance of the model is based on the correlation coefficient of calibration (Rc) and root mean square error of calibration (RMSEC). Correlation coefficient of prediction (Rp) and root mean squared error of prediction (RMSEP) of the prediction set are used as evaluation indicators (2023) [18]. In this paper, the snake optimization algorithm is used to optimize the penalty coefficient C and gamma parameters of SVR. SO optimization algorithm is an optimization algorithm proposed by Hashim and Hussien to simulate snake mating behavior (2022) [19]. It has the characteristics of strong global optimization ability and fast convergence speed, which further improves the performance of the model in practical applications.

3. Result and discussion

3.1. Statistics and division of sample set

In this study, KS algorithm (Kennard-Stone, KS) was used to divide 260 silage corn feed pH sample sets, which were divided into calibration set and prediction set according to the ratio of 7:3. The calibration set contained 182 samples and the prediction set contained 78 samples. The measured value of pH value of silage corn feed is shown in **Table 1**.

Table 1. Statistical table of pH value of silage corn feed

Quality indicators	Minimum value	Maximum value	Average value	Standard deviation
pH	3.55	4.36	3.87	0.24

3.2. Preprocessing of spectral data

The original spectral data were preprocessed by MAF, MSC and SNV. The original spectra are shown in **Figure 2**. The optimization effect of MSC treatment on the model performance is the most obvious. The spectra after MSC pretreatment are shown in **Figure 3**. MSC treatment is mainly used to eliminate or compensate the multiple scattering effect caused by the physical properties of the sample, so as to improve the accuracy and reliability of the spectral data (2019) [20]. The overall volatility of the data is reduced, and the reliability and accuracy of the modeling are improved. The modeling effect after preprocessing is shown in **Table 2**.

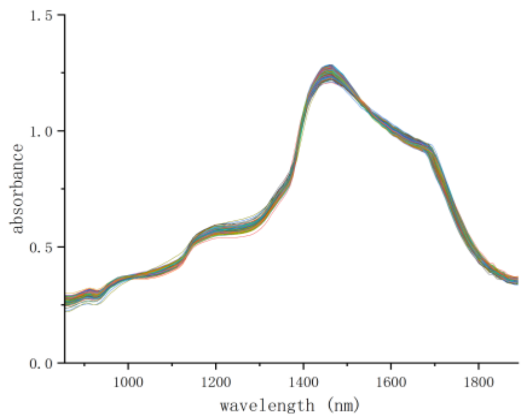

Figure 3. The spectra after MSC pretreatment.

Table 2. Modeling effect based on full-wavelength spectral data

Pretreatment	Modeling method	Correction set		Prediction set	
		Rc	RMSEC	Rp	RMSEP
RAW	SO-SVR	0.980289	0.121698	0.953924	0.189744
MAF		0.971500	0.150541	0.919277	0.230001
MSC		0.986350	0.099984	0.969147	0.151687
SNV		0.988410	0.092244	0.967660	0.151089

3.3. Characteristic wavelength optimization

In the process of establishing the model, the characteristic wavelengths are optimized, and the uninformative bands and bands with low correlation in the model establishment are screened and removed, so as to reduce the amount of calculation, simplify the model and improve the effect of establishing the model.

In this study, UVE, VCPA and VISSA algorithms were used to select the characteristic wavelengths. The UVE algorithm is used to filter the wavelength, and the noise random variable is set. The 1~128 on the abscissa is the actual spectral band variable, and the random noise variable is added after 128. The black dotted line is the upper and lower thresholds. The stability elimination threshold is 99%, and the wavelength variable outside the threshold line is the selected variable. **Figure 4** shows the algorithm screening process. A total of 48 characteristic wavelengths were screened, accounting for 37.50 % of the total wavelength, and its distribution is shown in **Figure 5**.

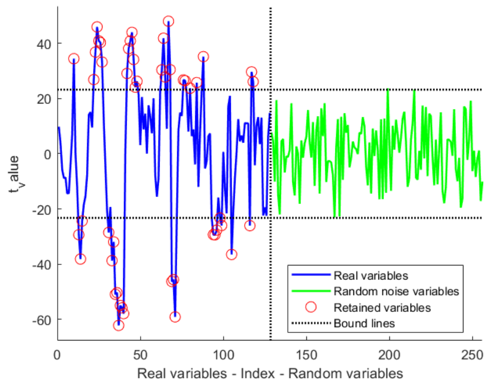

Figure 4. UVE algorithm preferred wavelength.

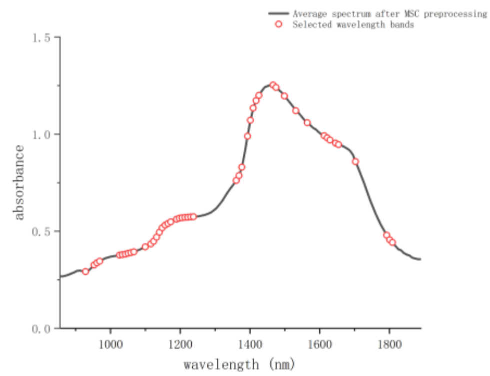

Figure 5. The preferred wavelength distribution of UVE.

In the use of CARS algorithm to filter the wavelength, the first centralized processing, the application of interactive verification method for iterative optimization, set the number of 1000 iterations, using 50 sampling, **Figure 6(a)** shows that with the increase of sampling times, a large number of invalid variables are gradually eliminated. **Figure 6(b)** shows that RMSECV reaches the minimum value at 14 sampling times. **Figure 6(c)** describes the relationship between the regression coefficient of the wavelength variable and the number of sampling times. The blue dotted line is the location of the minimum RMSECV, which corresponds to **Figure 6(b)**. This is the optimal subset. After screening by the CARS algorithm, a total of 22 characteristic wavelengths were screened, accounting for 17.18 % of the total wavelengths. The distribution is shown in **Figure 7**.

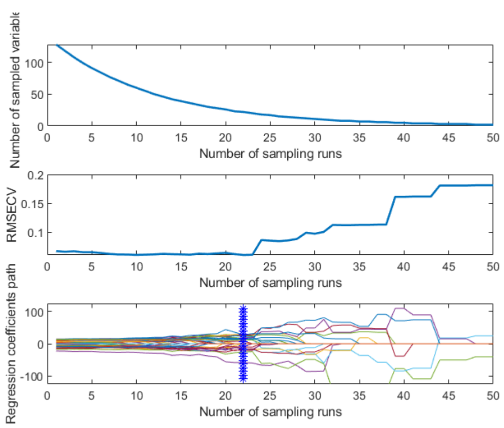

Figure 6. CARS algorithm preferred wavelength.

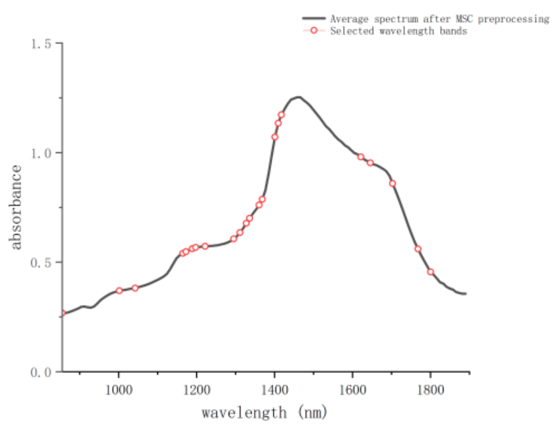

Figure 7. The preferred wavelength distribution of CARS.

The VISSA algorithm performs feature extraction on the spectral data after MSC processing. **Figure 8** shows the trend of RMSECV in the iterative process of the algorithm. When the number of variables is between 0 and 46, RMSECV shows a downward trend, indicating that effective information highly related to the detection of PH value is extracted during the iteration process. From 46 to 128, RMSECV began to show an upward trend, indicating that the algorithm added interference information in this process. Finally, 46 key characteristic wavelengths were selected, accounting for about 35.93 % of the total number of original features, and its distribution was shown in **Figure 9**.

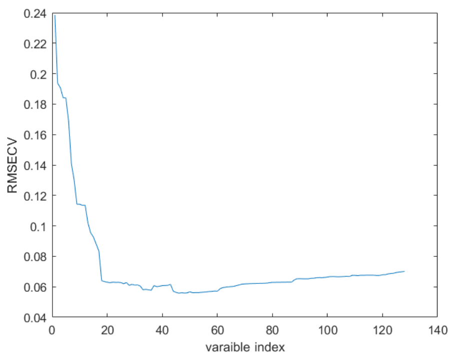

Figure 8. VISSA algorithm preferred wavelength.

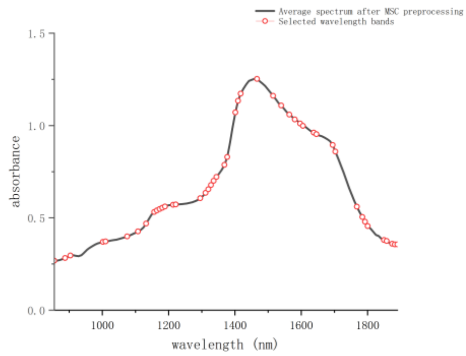

Figure 9. The preferred wavelength distribution of VISSA.

3.4. Establishment of quantitative detection model

The results of the quantitative detection model of SO-SVR silage corn feed pH based on different wavelengths are shown in **Table 3**. In the process of establishing the SO-SVR model, the effect of MSC pretreatment is better than others, and the model established by the characteristic wavelength screened by the VISSA algorithm has the best prediction effect. In this study, the MSC-VISSA-SO-SVR model was selected. **Figure 10** is the distribution map of model training set and prediction set.

Table 3. Comparison of modeling effects based on different characteristic wavelength selection

Wavelength selection algorithm	Modeling method	Wavelength number	Correction set		Prediction set	
			Rc	RMSEC	Rp	RMSEP
UVE	SO-SVR	48	0.985138	0.104582	0.973221	0.143228
CARS		22	0.986566	0.102071	0.971339	0.142622
VISSA		46	0.990531	0.083545	0.980487	0.127718

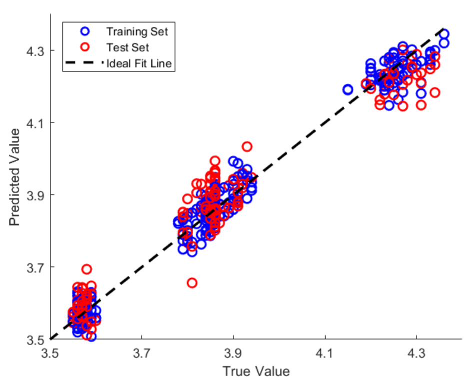

Figure 10. Scatter plot of silage corn feed pH value model prediction.

4. Conclusion

In this study, the spectral data of 260 silage corn feeds were collected based on the miniature near-infrared spectroscopy system, the composition of the acquisition system was introduced, and the pH value of silage corn feed was predicted by combining near-infrared spectroscopy and chemometrics. In the process of model establishment, the MSC processing in preprocessing and the VISSA algorithm in characteristic wavelength selection have the best prediction effect. The modeling effect Rp of MSC-VISSA-SO-SVR model is 0.980487, and RMSEP is 0.127718. The MSC-VISSA-SO-SVR model based on a miniaturized near-infrared spectroscopy system can be used as a valuable method to predict the pH value of silage corn feed. In addition, the main indicators for evaluating the fermentation quality of silage corn feed also include organic acids and ammonia nitrogen. In this study, only the detection and prediction model of pH was established, and the prediction models of other indicators were established for detection, so as to evaluate the quality of silage corn feed in many aspects.

Disclosure statement

The author declares no conflict of interest.

References

[1] Zhang W, Fan P, Wu K, 2025, Application of Silage Corn Feed in Beef Cattle Breeding. Animal Industry and Environment, 2025(2): 18–19.

[2] Liu N, Wang J, Tu Y, et al., 2019, Quality Evaluation of Corn Silage and Establishment of Rapid Detection Method for its Nutritional Components. Journal of Domestic Animal Ecology, 40(10): 1–7.

[3] Ning H, Wang J, Jiang H, et al., 2022, Quantitative Detection of Zearalenone in Wheat Grains based on Near-Infrared Spectroscopy. Spectrochimica acta, Part A, Molecular and Biomolecular Spectroscopy, 2022(280): 121545–121545.

[4] Wang J, Zareef M, He P, et al., 2019, Evaluation of Matcha Tea Quality Index using Portable NIR Spectroscopy Coupled with Chemometric Algorithms. Journal of the Science of Food and Agriculture, 99(11): 5019–5027.

[5] Yanna R, Muhammad Z, Lihua L, et al., 2023, Application of Portable Vis-NIR Spectroscopy for Rapid Detection of Myoglobin in Frozen Pork. Meat Science, 2023(201): 109170–109170.

[6] Na R, Yin H, Hu B, 2022, Application of Near-Infrared Spectroscopy in the Field of Rapid Detection of Feed. Feed Review, 2022(2): 44–46+50.

[7] Dai C, Xu X, Huang W, et al., 2022, Monitoring of Critical Parameters in Thermophilic Solid-State Fermentation Process of Soybean Meal using NIR Spectroscopy and Chemometrics. Journal of Food Measurement and Characterization, 17(1): 576–585.

[8] Luca M, James D, 2018, The Use of Mobile Near-Infrared Spectroscopy for Real-Time Pasture Management. Frontiers in Sustainable Food Systems, 2018, 2.

[9] Tian L, Chen F, Lin Y, et al., 2023, Study on Nutritional Quality of Oat Forage Silage by Near Infrared Spectroscopy. Acta Agrestia Sinica, 31(12): 3841–3850.

[10] Wang X, Guo T, Pan F, et al., 2021, Near Infrared Spectroscopy was used to Rapidly Analyze the Nutritional Components of Whole Corn Silage. Journal of Domestic Animal Ecology, 42(1): 52–55.

[11] Zhang M, Zhao C, Shao Q, et al., 2019, Determination of Water Content in Corn Stover Silage using Near-Infrared Spectroscopy. International Journal of Agricultural and Biological Engineering, 12(6): 143–148.

[12] Liu X, Han L, Yang Z, et al., 2007, Rapid Analysis of Silage pH Value and Fermentation Products by Near Infrared

Spectroscopy. Chinese Journal of Analytical Chemistry, 2007(9): 1285–1289.

[13] Zhang M, Hao M, Tian H, et al., 2023, Nondestructive Detection of pH Value of Silage Corn Feed based on Hyperspectral Imaging Technology. Transactions of the Chinese Society of Agricultural Engineering, 39(4): 239–247.

[14] Yu Y, Tian H, Zhao K, et al., 2024, Rapid pH Value Detection in Secondary Fermentation of Maize Silage Using Hyperspectral Imaging. Agronomy, 14(6): 1204.

[15] Wang F, Li Y, Peng Y, et al., 2018, Study on Nondestructive Detection Method of Lycopene Content based on Visible/Near Infrared Transmission Spectroscopy. Chinese Journal of Analytical Chemistry, 46(9): 1424–1431.

[16] Zhang F, Wang M, Yan B, et al., 2025, Identification of Egg Varieties by Hyperspectral Imaging Combined with ELM. Spectroscopy and Spectral Analysis, 45(3): 836–841.

[17] Wang F, Wang C, Shi S, 2022, Rapid and Low-Cost Detection of Millet Quality by Miniature Near-Infrared Spectroscopy and Iteratively Retaining Informative Variables. Foods, 11(13): 1841–1841.

[18] Ma Y, Li L, Shi Y, et al, 2023, Rapid Detection of Cotton Seed Protein and Oil Content based on Near Infrared Spectroscopy. Cotton Science, 35(3): 211–219.

[19] Hashim F, Hussien G, 2022, Snake Optimizer: A Novel Meta-Heuristic Optimization Algorithm. Knowledge-Based Systems, 2022(242).

[20] Sun J, Zhou X, Hu Y, et al., 2019, Visualizing Distribution of Moisture Content in Tea Leaves Using Optimization Algorithms and NIR Hyperspectral Imaging. Computers and Electronics in Agriculture, 2019(160): 153–1.