DETERMINATION OF RICE SEED VIGOR BY LOW-FIELD NUCLEAR MAGNETIC RESONANCE COUPLED WITH MACHINE LEARNING
/
低场核磁共振技术结合机器学习判别水稻种子活力方法研究

Ensi CHENG [1], Ping SONG [2], Boxiao WANG [1], Tiangang HOU [1], Liyan WU [*1], Benhua ZHANG [*3]
[1]College of Engineering, Shenyang Agricultural University, Shenyang/ China;
[2]College of Information and Electrical Engineering; Shenyang Agricultural University, Shenyang/ China;
[3]School of Mechanical and Electrical Engineering, Suqian College, Suqian/ China
Tel: +86-139-4209-8712; E-mail:wly78528@syau.edu.cn
Corresponding author: Wu Liyan; Zhang Benhua
DOI: https://doi.org/10.35633/inmateh-67-52

Keywords: *LF-NMR, Machine learning, Rice seeds, Seed vigor*

ABSTRACT
Physiological index data and low-field nuclear magnetic resonance (LF-NMR) spectral data of rice seed samples from three varieties harvested in different years were collected through a combination of the standard germination test and an LF-NMR test. Three parameters of seed vigor: germination energy, germination percentage, and germination index, were calculated based on the physiological index data of the rice seed samples to determine their vigor over the years after harvest. LF-NMR Carr-Purcell-Meiboom-Gill (CPMG) sequence echo-peak data were used as the input, and rice seed vigor was used as the output to establish discriminative models using principal component analysis, support vector machine, logistic regression, K-nearest neighbor, artificial neural network, and Fisher's linear discriminant. The results showed that models constructed using any algorithm, except for principal components analysis-algorithm distinguished between seeds with high and low vigor, while models constructed using Fisher's linear discriminant algorithm gave the best results. This study provided a rapid, accurate, and non-destructive method to test rice seed vigor, offering theoretical support and a reference for rice seed-sorting and storage research.

摘要
采用标准发芽试验和低频核磁共振(LFNMR)相结合的方法，采集了 3 个不同年份水稻种子的生理指标数据和低场核磁共振波普数据。根据水稻种子生理指标数据，计算发芽势、发芽率和发芽指数 3 个参数，对不同收获年份的水稻种子进行活力高低的区分。将 LF-NMR 硬脉冲序列回波峰点数据作为输入，水稻种子活力水平作为输出，建立结合主成分分析、支持向量机、逻辑回归、k 近邻、人工神经网络和 Fisher 线性判别法建立判别模型。结果表明，除主成分分析算法外，任何一种算法所构造的模型都能区分种子的活力高低，而 Fisher 线性判别算法所构造的模型效果最好。本研究为水稻种子活力的快速、准确、无损检测提供了一种方法，为水稻分选贮藏研究提供了理论支持和参考。

INTRODUCTION
Rice (*Oryza sativa* L.) is an important food crop for humans, with more than 50% of the world population relying on rice as a staple food. Seed vigor is the sum of a rapid and neat germination, growth, and productive potentials of a seed lot under a wide range of field conditions. Seed vigor and vitality levels are critical in determining seed quality, and can directly affect seed germination, seedling growth rate, and plant vigor *(Reed, 2022)*. High seed vigor plays a key and decisive role in achieving high and stable rice yields. Rice seed vigor is highest at harvest and decreases gradually as the seeds age during subsequent storage. Sowing of low-vigor rice seed causes great economic losses to agricultural production *(Yan, 2018)*. Hence, a rapid, accurate, and non-destructive method for testing rice seed vigor is paramount.

Methods for testing seed vigor can be divided into two categories: traditional testing methods and non-destructive testing (NDT) methods *(Milosevic, 2010)*. Among the traditional methods, the standard germination test is widely used and allows visual determination of seed vigor. However, it requires manual measurement of root and shoot lengths of seedings during germination, which has the disadvantages of requiring complex operation procedures, high technical expertise, high destructive effect, long test cycle, and shows low testing efficiency *(Romano and Stevanato, 2020)*. The International Seed Testing Association (ISTA) recommends the conductivity and the tetrazolium tests to assess seed vigor by analyzing the relationship between seed physiological and biochemical indices and germination rate and potential *(Ma, 2020)*. Other methods include the accelerated aging test *(Alahakoon, 2021)*, the cool germination test *(Singh, 2021)*, the complex stress vigor test *(Zhang, 2021)*, and the cold test *(Sharma, 2021)*. However, rapid, accurate, and non-destructive testing methods are gaining increasing acceptance and are the future trend for measurement of seed vigor *(Zhang, 2022)*. Thus, for example, *Wang et al.* (*2020*) used near-infrared (NIR) spectroscopy coupled with partial least squares-discriminant analysis (PLS-DA) to identify the vigor of single maize seeds under different treatments *(Wang, 2020)*. Similarly, *Zhang et al.* (*2017*) used an electronic nose to analyze the odor of maize seeds with different vigor *(Zhang, 2017)*; furthermore, they used principal component analysis (PCA), support vector machine (SVM), and Fisher linear discriminant (FLD) to establish a qualitative analysis model of maize seed vigor. In turn, He et al. (2019) used near-infrared hyperspectral imaging (NIR-HSI) to distinguish rice seeds from different years and identified the non-viable ones using PLS-DA *(He, 2019)*. Yasmin et al. (2020) used the machine vision technique to identify healthy and non-healthy tomato seeds and discriminated seed quality in real time by image analysis *(Yasmin, 2020)*. Similarly, Feng et al. (2021) built a discriminative model based on electrical impedance spectroscopy parameters for rapid NDT of rice seed viability *(Feng, 2021)*. Song et al. (2021a) used the low-field nuclear magnetic resonance (LF-NMR) technique for NDT of moisture content and moisture distribution in rice seeds with different vigor *(Song, 2021)*. However, to date, few studies have reported using the LF-NMR technique coupled with machine learning algorithms to determine rice seed vigor.

In this study, three varieties of newly harvested rice seeds and naturally aged rice seeds were used as the research objects. Firstly, the standard germination test was conducted to discriminate the vigor of rice seeds of each variety in different years, and then the LF-NMR technique was used to collect the LF-NMR spectral data of all test samples to obtain information regarding their internal hydrogen ions. Subsequently, six machine learning algorithms, including PCA, SVM, logistic regression (LR), K-nearest neighbor (KNN), artificial neural network (ANN), and FLD, were used to discriminate the vigor level of the same rice seeds, and discriminative models of rice seed vigor were constructed. Finally, the validity of these models was verified. The study aimed to find an efficient and accurate method to determine the vigor of rice seeds and provide a reference for rice seed sorting and storage research.

MATERIALS AND METHODS

Experimental materials

This study used seeds of three different rice varieties. Two of them are conventional japonica rice varieties, namely, Shen Nong 9816 (SN) and Bei Geng 3 (BG), bred by the Rice Research Institute at Shenyang Agricultural University, and suitable for cultivation in the mid- to late-maturing rice areas in northeastern China and south of Shenyang. The third variety is the conventional indica rice Tai You 871 (TY), bred by the Agricultural College of Jiangxi Agricultural University and suitable for cultivation in areas less prone to rice blast disease, in Jiangxi Province, China. All test samples were provided by the Rice Research Institute at Shenyang Agricultural University. Rice seeds were selected and those newly harvested (October 2021) were numbered 2021SN, 2021BG, and 2021TY, according to the names above, respectively. Meanwhile, naturally aged rice seeds (harvested in October 2020) were numbered 2020SN, 2020BG, and 2020TY. Rice seeds of the same varieties were sourced from the same seed lots. All seeds had initial moisture contents lower than 14% and were stored at room temperature away from light. Among the test samples of different varieties in different years, 380 rice seeds were selected as mature, full, and free of mold, germination, and insects, of which 300 seeds were used for the standard germination test (three groups of 100 seeds each, were used as replicates), and 80 seeds were used for the LF-NMR test (16 groups of five seeds each, were used as replicates).

Germination test

To evaluate the differences in rice seed vigor due to natural aging, a 7-d standard germination test was conducted in compliance with International Rules for Seed Testing (2022). Rice seeds were first soaked in a 0.2% potassium permanganate solution for 15 min and then rinsed five times with distilled water. The germination box was lined with three layers of germination paper, moistened with distilled water, and then 100 seeds were evenly placed in it. To prevent water evaporation, the germination box was covered with a lid, and placed in an RTOP-268D intelligent artificial climate chamber (Zhejiang Tuopu Instruments Co., Ltd., China). The germination temperature was set at 27°C, with a 12/12 h light/dark regime in an alternating mode. To keep the germination paper moist, distilled water was sprayed quantitatively every day. The total germination time was 7 d, and the number of germinated seeds was recorded daily. Germination energy (GE) was calculated 4 d after germination, and germination percentage (GP) and germination index (GI) were calculated 7 d after germination using Equations 1, 2 and 3. The results were expressed using the mean of the three replicates used for the standard germination test.

$$GE=\frac{N_4}{N}\times 100 \quad (1)$$

where: N_4 is the number of germinated seeds on the fourth day, and N is the total number of tested seeds.

$$GP=\frac{N_7}{N}\times 100 \quad (2)$$

where: N_7 is the number of final germination seeds, and N is the total number of tested seeds.

$$GI=\sum_{t=1}^{7}\frac{N_t}{t} \quad (3)$$

where: t is the germination day, and N_t is the number of germinated seeds per day corresponding to t.

LF-NMR relaxation measurements

The transverse relaxation time T_2 is an important LF-NMR parameter to characterize the hydrogen-ion relaxation inside rice seeds. Different components of the same material have different Carr-Purcell-Meiboom-Gill (CPMG)-echo-sequence peak curves, which can be visualized using LF-NMR. The superimposed relaxation signals of all hydrogen ions in rice seeds were measured at a resonance frequency of 21 MHz using an NMI20-015V-I Nuclear Magnetic Resonance analyzer (Shanghai Niumag Science and Technology Co., Ltd., China). The magnet temperature was set to 32°C to ensure a constant rate of hydrogen-ion motion in the detection system.

The CPMG pulse sequence was used to scan all the test samples five times with 3,000 echoes; the repeated-sampling waiting time was 0.150 ms, pulse width for 90° radiofrequency (RF) was 17 μs, pulse width for 180° RF was 36 μs, and repeated sampling interval was 1,000 ms. After measurement, raw PEA files were obtained, and the CPMG-echo-sequence peak data of the samples were extracted.

Data processing and analysis

This study adopted the LF-NMR technique to collect CPMG-echo-sequence peak data of rice seed samples for multivariate analysis, and the algorithms used for the analysis included PCA, SVM, LR, KNN, ANN, and FLD *(Ge, 2017)*. Cross-validation of data was performed using the hold-out method to ensure the discriminative accuracy of the final model trained on all data.

Principal component analysis is an analytical method to reduce the dimensionality of high-dimensional data. LF-NMR CPMG-echo-sequence peak data contain a large amount of redundant information and random noise values that are not conducive to building an accurate model for discriminating rice seed vigor. Therefore, the extraction of feature variables using PCA techniques is necessary. To retain the main features of the original data variation, the multivariate data were linearly transformed and dimensionally reduced, and the reduced feature vectors were extracted. First, the CPMG-echo-sequence peak data of each tested sample was set as a multivariate in one row, and each row vector represented one sample. The CPMG-echo-sequence peaks of the rice seed samples of the same variety and different years were grouped together to obtain the original matrix M using Equation 4.

$$M=n\times a \quad (4)$$

where: n denotes the number of seed samples, and a refers to the CPMG-echo-peak data of the samples.
Then, the data matrix M was analyzed according to the maximum variance theory by PCA.

$$M=M_1\times M_2 \quad (5)$$

where: M_1 is an $n \times m$ dimensional matrix, n is the number of rice seed samples, m is the number of principal components, and M_2 is an $m \times m$ dimensional matrix. Based on the principle of PCA, matrix M_1 is the principal component score matrix, i.e., the first column is principal component 1 (PC1), the second column is principal component 2 (PC2), and the other principal components are derived in turn.

Support vector machine (SVM) is a supervised learning approach to solve the problem of data binary classification in machine learning. Its goal is to find a hyperplane using two classes of data as far away from the hyperplane as possible, thus classifying the new data more accurately *(Hu, 2020)*.

Logistic regression is a typical probabilistic statistical classification model; a linear algorithm widely used to discriminate binary classification problems with k-dimensional independent variables. Using a sigmoid function allows LR to output only two values, so that it is possible to classify both naturally aged and newly harvested rice seeds, thereby distinguishing between their vigor.

As for KNN, this is a classification model often applied in pattern recognition, machine learning, and data mining. The principle of this algorithm is that a naturally aged rice seed sample is the nearest neighbor to *k* samples in the dataset in the feature space. If most of these *k* samples belong to the naturally aged rice seed category, then the sample also belongs to this category; otherwise, the sample belongs to the newly-harvested seed category.

In turn, ANN is a classification algorithm that mimics the behavioral characteristics of animal neural networks and performs distributed parallel information processing. It consists of an input layer, multiple hidden layers, and an output layer. ANN can acquire relevant knowledge through learning and store it in each connection weight. It forms memory in the process of continuous training to complete more complex classification discriminations with strong fault tolerance and nonlinear mapping capability.

Finally, FLD is a popular binary classification algorithm in machine learning and data mining. The LF-NMR CPMG-echo-sequence peak data are projected in low dimensions. After projection, the projection points of similar samples are as close as possible, and those of dissimilar samples are as far as possible, i.e., the projection minimizes intra-class variance and maximizes inter-class variance, thus discriminating between seed samples contrasting for vigor.

For each rice variety, 160 test seed samples were used for the LF-NMR test; five seeds were used as a group, and each test was repeated five times. Thus, a total of 160 groups of CPMG-echo-sequence peak data were obtained for each variety. Among them, 100 groups were randomly selected as the training set, and the remaining 60 groups were used for model validation. The indices of accuracy, sensitivity, and specificity were used to evaluate the binary classification performance of the discriminative model, as shown in Equations 6, 7, and 8.

$$Accuracy\ (\%)=\frac{FP+NA}{TT}\times 100 \tag{6}$$

$$Sensitivity\ (\%)=\frac{FP}{FP+FN}\times 100 \tag{7}$$

$$Specificity\ (\%)=\frac{NA}{NA+NF}\times 100 \tag{8}$$

where: FP are freshly harvested seeds predicted to be freshly picked seeds; NA are naturally aged seeds predicted to be naturally aged seeds; TT is the total number of seeds of the same variety from different years; NF are naturally aged seeds predicted to be freshly harvested seeds, and FN are freshly harvested seeds predicted to be naturally aged seeds.

The statistical analysis software SPSS (version 25.0, IBM, USA) was used to analyze and process all the data collected in the LF-NMR spectral test and the standard germination test. The function plotting software Origin (version 2021; OriginLab Company, USA) was used for graph plotting. In terms of machine learning algorithms, MATLAB (R2021b; MathWorks Company, USA) was used to build and validate the models.

RESULTS

Analysis of the results of the standard germination test

Figure 1 shows the germination of a group of rice seed samples of variety BG on day 7 after initiation of vigor testing using the standard germination test. Fig. 1a shows the newly harvested seeds, while Fig. 1b shows the naturally aged seeds.

As Fig. 1 clearly shows, the growth rate of both roots and shoots of the naturally-aged seeds were lower than that of the newly harvested seeds.

(a) Newly picked rice seeds (b) Naturally aged rice seeds

Fig. 1 - Germination of naturally aged and newly picked seeds (BG, 7d)

The number of germinated seeds per day during the seed germination test of the three rice varieties in different years was recorded; additionally, seedling height was measured and GP, GE, and GI of the seeds, which indicate the level of vigor, were calculated.

The results are shown in Table 1.

Table 1

Effects of natural aging on GP, GE and GI
(*Newly harvested seed. **naturally-aged seed)

Variety	GP (%)	EG (%)	GI
2021SN*	96	93	28.7
2020SN**	45	34	13.1
2021BG*	98	96	29.2
2020BG**	35	28	10
2021TY*	95	94	27.9
2020TY**	21	15	6

As can be seen, the germination rate of SN decreased from 96% to 45% with aging, that of BG decreased from 98% to 35%, and that of TY decreased from 95% to 21%. The highest difference (74%) in GP between newly harvested and naturally aged seed was observed for variety TY, and the lowest difference (51%) was observed for variety SN. The germination rate of rice seeds of the same variety in different years showed a large difference, so that the germination rate of newly harvested rice seeds was high, whereas that of naturally aged rice seeds was significantly lower.

Furthermore, the germination potential of newly harvested rice seeds of SN, BG, and TY was 93%, 96%, and 94%, respectively, whereas that of naturally aged rice seeds was 34%, 28%, and 15%, respectively. Indeed, germination potential followed the same trend as germination rate. According to the International Rules for Seed Testing (2022), GI is an important index of high rice seed vigor. In this study, GI of 2021BG was the largest, indicating that its seed vigor was the highest among the tested samples. Conversely, GI of 2020TY was the lowest, indicating that its seed vigor was the lowest among the tested samples. Vigor index of newly harvested seeds was higher than that of naturally-aged seeds regardless of variety.

Therefore, rice seeds from 2021 were defined as high-vigor seeds, whereas seeds from 2020 were defined as low-vigor rice seeds.

Analysis of LF-NMR spectral test results

LF-NMR spectral data of seed samples of three rice varieties were examined under the same conditions. Fig. 2 shown the CPMG-echo-sequence peaks of the tested seeds of the same variety in different years.

Figure 2 indicates that the CPMG-echo-sequence peak curves of the seeds of the three rice varieties showed similar trends. The decay rate of the CPMG-echo-sequence peaks for naturally-aged seed was lower than that for the newly harvested seeds, from 20 to 100 ms. After 100 ms, the decay rate of the rice seeds in different years gradually tended to weaken until the end of the decay at 450 ms.

The decay of the CPMG-echo-peak curve is closely related to the content and the free state of hydrogen ions in the seed samples, which in turn is related to seed vigor. Hence, the curve data can be used to evaluate the vigor of the seed samples.

Small differences were observed between the CPMG-echo-sequence peak curves of seed samples from different years, and the difference between curvature and signal intensity was also small. Thus, it is impossible to distinguish and discriminate between newly-harvested and naturally-aged rice seeds using this curve alone. Addressing this problem requires parsing the curve data using machine learning methods. The extracted CPMG-echo-sequence peaks of rice seeds from different years were used to discriminate between newly-harvested and naturally-aged rice seeds.

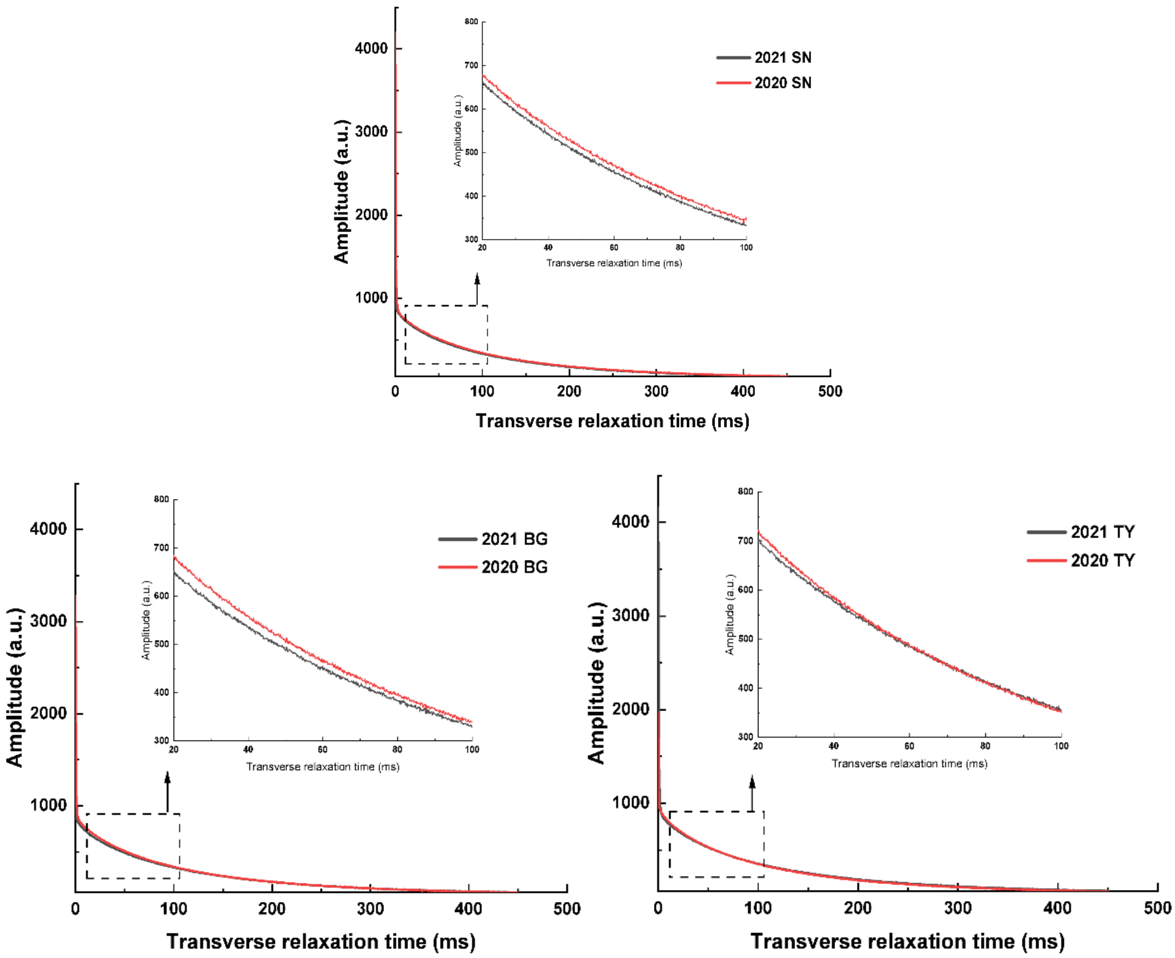

Fig. 2 - Carr-Purcell-Meiboom-Gill (CPMG)-echo-sequence peaks of seeds of three rice varieties

Analysis of PCA algorithm results

The PCA algorithm was used to linearly transform and downscale the LF-NMR CPMG-echo-sequence peak data. Then, a multivariate analytical model was established based on the different varieties. The intra-class vigor difference of different seed samples was characterized by the interval distance on the principal component score plot. The farther the interval between two classes of samples, the greater the difference in vigor, and vice versa. The contribution rates of PC1 and PC2 obtained in the PCA linear transformation were included in PC1 and PC2. The larger the contribution rate, the better the primary component reflected index information.

The PCA algorithm was adopted to extract the principal feature components of the LF-NMR CPMG-echo-sequence peak data to distinguish between rice seeds of the same variety in different years with similar feature information. The results of PCA analysis showed that the PC1 and PC2 explained 44.8%, 46.9%, and 51.3% of the information original variances in SN, BG, and TY, respectively, as illustrated in Fig. 3.

Figure 3 revealed that the separation effect for TY was the best, whereas that for BG was the poorest. However, regardless of the variety, it was difficult to distinguish naturally-aged from newly-harvested seed using the PCA method. Therefore, other machine learning classification algorithms were required to parse these LF-NMR CPMG-echo-sequence peak data. Rice seed vigor was discriminated by extracting information that reflected different vigor characteristics.

The PCA method usually selects a cumulative contribution of ≥ 80% to determine the first *k* components to be defined as the principal components of the research question. The feature variables of seed samples of the same rice variety in different years were extracted based on this method. The selected results were analyzed, and the top 10 principal components were found to have a contribution rate of ≥ 80%. Hence, the top 10 principal components of CPMG-echo-peak data were selected as the feature variables for the analysis model, and five other machine learning classification algorithms were used to further construct the models to determine the vigor of rice seed.

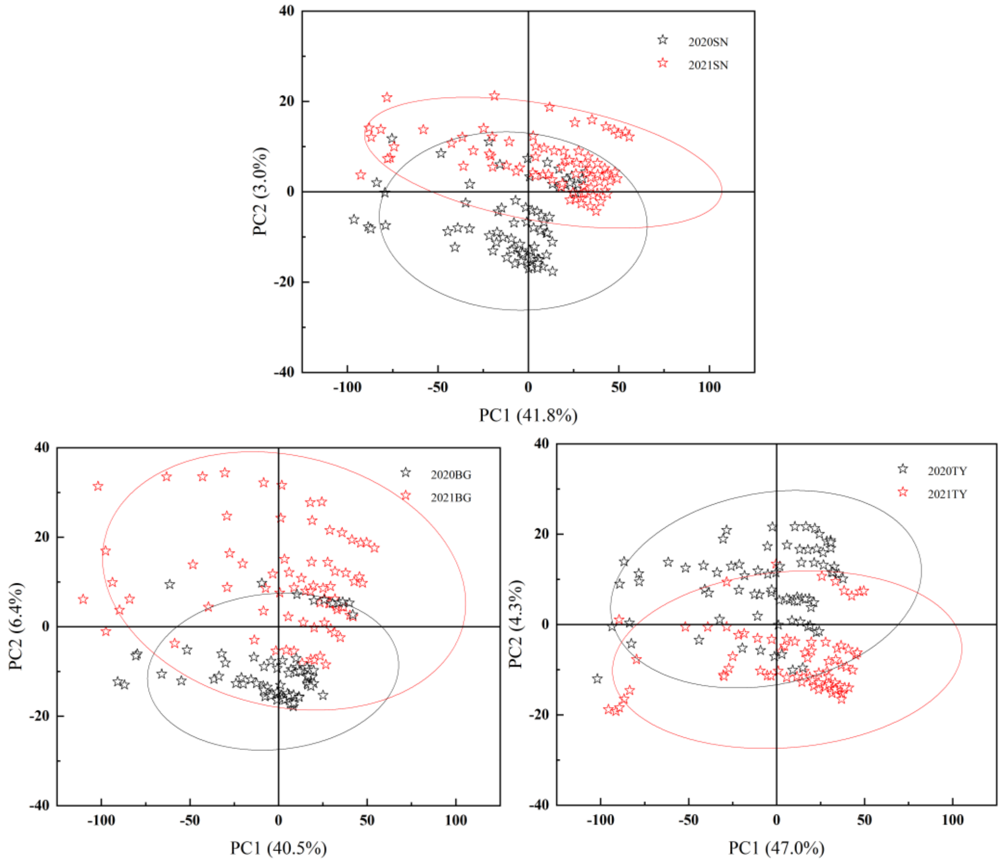

Fig. 3 - **Principal component analysis (PCA) plots of the CPMG-echo-sequence peak data for the seed of three rice varieties from two different years**

Analysis of the results of five machine learning classification algorithms

Model training was performed using the SVM, LR, KNN, ANN, and FLD algorithms on 100 groups of randomly obtained training set data for each variety, and the model validation was conducted with 60 groups of validation set data. The results of the evaluation of the indices of accuracy, sensitivity, and specificity of the five models of multivariate analysis for the determination of seed vigor are shown in Table 2.

Table 2

Results of the evaluation of seed vigor indices based on support vector machine (SVM), logistic regression (LR), K-nearest neighbor (KNN), artificial neural network (ANN), and Fisher's linear discriminant (FLD)

Model	Index	SN	BG	TY
SVM	**Accuracy (%)**	96.7	95	93.4
	Sensitivity (%)	96.7	96.7	90
	Specificity (%)	96.7	93.4	96.7
LR	**Accuracy (%)**	93.4	95	93.4
	Sensitivity (%)	90	93.4	86.7
	Specificity (%)	96.7	96.7	100
KNN	**Accuracy (%)**	95	95	83.4
	Sensitivity (%)	93.3	96.7	70
	Specificity (%)	96.7	93.4	96.7
ANNs	**Accuracy (%)**	83.3	85	81.6
	Sensitivity (%)	86.6	73.4	80
	Specificity (%)	80	96.7	83.3
FLD	**Accuracy (%)**	100	98.3	100
	Sensitivity (%)	100	96.7	100
	Specificity (%)	100	100	100

By comparing and analyzing each index of the discriminative models of newly-harvested and naturally-aged seeds of three rice varieties, constructed using SVM, LR, KNN, ANN and FLD algorithms, it was found that each model discriminated seed vigor. Additionally, among these classification algorithms, the discriminative model constructed by the ANN algorithm performed poorest in discriminating seed vigor, whereas the model constructed by the FLD algorithm performed best. The rice seed-vigor discriminative model constructed by the FLD algorithm was visualized as shown in Fig. 4. This model distinguished well between the newly-harvested and the naturally-aged seeds of the three varieties tested herein.

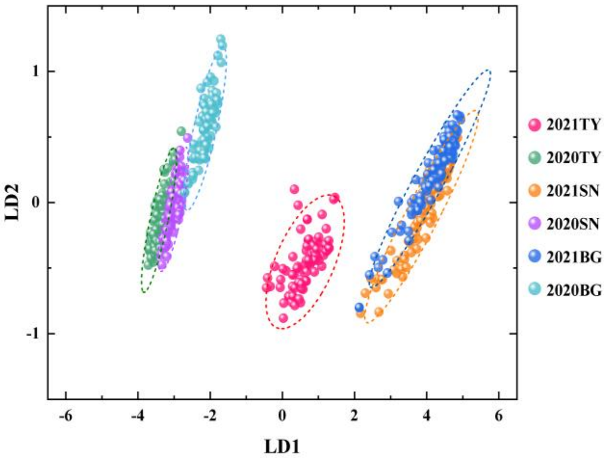

Fig. 4 - FLD plots of the CPMG-echo-sequence peak data of the same variety of rice seeds in different years Model validation

Pre-processed CPMG-echo-sequence peak data of 50 groups of seed of the BG rice variety from two different years were randomly selected to validate the effectiveness of the discriminative model constructed using the FLD algorithm for determining rice seed vigor. The discriminative model constructed using the FLD algorithm, which performed best among the five classification algorithms, was selected, and the classification results are shown in the form of a visualized confusion matrix in Fig. 5.

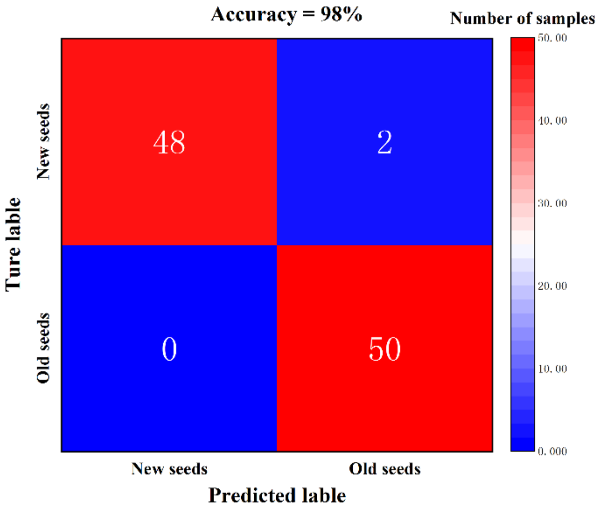

Fig. 5 - Confusion matrix of classification results as per the FLD model

In Fig. 5, new seeds are the newly-harvested seeds, and old seeds are the naturally-aged seeds. The true labels are the number of actual seed samples in different years, and the values in the upper left and lower right cells are the number of correctly discriminated seed samples in different years. The higher values in the upper left and in the lower right corners indicate that the model constructed using the FLD algorithm is more capable of discriminating newly-harvested and naturally-aged seeds. The color of the matrix grid is related to the size of the values in the grid, and the color bar on the right side gives the specific numerical correspondence. The accuracy in BG recognition was 98%, similar to the results in Table 2. Thus, 48 samples of newly-harvested seeds and 50 of naturally-aged seeds were correctly classified, and the intra-class accuracies were 96% and 100%, respectively. Furthermore, the recognition accuracy of BG seeds of different years by the FLD algorithm-based model was greater than 95%, thereby showing strong generalization performance, indicating that the model constructed using the FLD algorithm effectively discriminated the level of vigor between rice seeds from two different years.

CONCLUSIONS

1. The standard germination test was used to obtain values of three indices to characterize rice seed vigor, namely, GP, GE, and GI. The values of all three indices of the newly-harvested seeds were higher than those of the naturally-aged seeds. Seed samples were classified for vigor according to the year of harvest.

2. The LF-NMR test was conducted to obtain the spectral data of seed samples of three rice varieties from two different years. The PCA algorithm was adopted to discriminate the vigor of the seed samples by the CPMG-echo-sequence peak data. The results led to the conclusion that distinguishing a naturally-aged rice seed from a newly-harvested one was difficult.

3. The PCA algorithm was used to select the first 10 principal components of the CPMG-echo-sequence peak data as the feature variables of the analysis model. Subsequently, the SVM, LR, KNN, ANN, and FLD algorithms were used to build and validate the corresponding seed-vigor discrimination models for three rice varieties. All models discriminated for seed vigor but best performance was observed for the model constructed using the FLD algorithm.

ACKNOWLEDGEMENT

This research was supported by the National Natural Science Foundation of China (31701318, 31811540396) and by the Basic Research Project of Liaoning Province (LSNJC201916, LSNJC201908).

REFERENCES

[1] Alahakoon, A.A.C.B., Abeysiriwardena, D.S.d Z., Damunupola, J.W., Hay, F.R & Gama-Arachchige, N. S. (2021). Accelerated aging test of seed vigor for predicting field emergence of wet direct-seeded rice. *Crop & Pasture Science,* 72**,** 773-781.

[2] Feng, L., Hou, T., Wang, B. & Zhang, B. (2021). Assessment of rice seed vigor using selected frequencies of electrical impedance spectroscopy. *Biosystems Engineering,* 209**,** 53-63.

[3] Ge, Z., Song, Z., Deng, S.X. & Huang, B. (2017). Data Mining and Analytics in the Process Industry: The Role of Machine Learning. *IEEE Access,* 5**,** 20590-20616.

[4] He, X., Feng, X., Sun, D., Liu, F., Bao, Y. & He, Y. (2019). Rapid and Nondestructive Measurement of Rice Seed Vitality of Different Years Using Near-Infrared Hyperspectral Imaging. *Molecules,* 24.

[5] Hu, R., Zhu, X., Zhu, Y. & Gan, J. (2020). Robust SVM with adaptive graph learning. *World Wide Web-Internet and Web Information Systems,* 23**,** 1945-1968.

[6] Ma, W.G., Zhang, Z.H., Zheng, Y.Y., Pan, W., Qiu, T., Guan, Y.J. & Hu, J. (2020). Determination of tobacco (Nicotiana tabacum) seed vigor using controlled deterioration followed by a conductivity test. *Seed Science and Technology,* 48**,** 1-10.

[7] Milosevic, M., Vujakovic, M. & Karagic, D. (2010). Vigor tests as indicators of seed viability. *Genetika-Belgrade,* 42**,** 103-118.

[8] Reed, R.C., Bradford, K.J. & Khanday, I. (2022). Seed germination and vigor: ensuring crop sustainability in a changing climate. *Heredity,* 128**,** 450-459.

[9] Romano, A. & Stevanato, P. (2020). Germination Data Analysis by Time-to-Event Approaches. *Plants-Basel,* 9.

[10] Sharma, N., Reinke, R. & Sacks, E.J. (2021). Comparison of Methods to Evaluate Rice (Oryza sativa) Germplasm for Tolerance to Low Temperature at the Seedling Stage. *Agronomy-Basel,* 11.

[11] Singh, J., Guzman, I., Begna, S., Trostle, C. & Angadi, S. (2021). Germination and early growth response of guar cultivars to low temperatures. *Industrial Crops and Products,* 159.

[12] Song, P., Kim, G., Song, P., Yang, T., Yue, X. & Gu, Y. (2021). Rapid and non-destructive detection method for water status and water distribution of rice seeds with different vigor. *International Journal of Agricultural and Biological Engineering,* 14**,** 231-238.

[13] Wang, Y., Peng, Y., Zhuang, Q. & Zhao, X. (2020). Feasibility analysis of NIR for detecting sweet corn seeds vigor. *Journal of Cereal Science,* 93.

[14] Yan, S.J., Huang, W.J., Gao, J.D., Fu, H. & Liu, J. (2018). Comparative metabolomic analysis of seed metabolites associated with seed storability in rice (*Oryza sativa L*.) during natural aging. *Plant Physiology and Biochemistry,* 127**,** 590-598.
[15] Yasmin, J., Lohumi, S., Ahmed, M.R., Kandpal, L.M., Faqeerzada, M.A., Kim, M.S. & Cho, B.K. (2020). Improvement in Purity of Healthy Tomato Seeds Using an Image-Based One-Class Classification Method. *SENSORS,* 20.
[16] Zhang, R., Hussain, S., Wang, Y., Liu, Y., Li, Q., Chen, Y., Wei, H., Gao, P. & Dai, Q. (2021). Comprehensive Evaluation of Salt Tolerance in Rice (*Oryza sativa L.*) Germplasm at the Germination Stage. *Agronomy-Basel,* 11.
[17] Zhang, T., Sun, Q., Yang, L., Yang, L. & Wang, J. (2017). Vigor detection of sweet corn seeds by optimal sensor array based on electronic nose. *Nongye Gongcheng Xuebao/Transactions of the Chinese Society of Agricultural Engineering,* 33**,** 275-281.
[18] Zhang, X., Hou, Z. & Xuan, C. (2022). Design and Experiment of Recognition System for Coated Red Clover Seeds Based on Machine Vision. *INMATEH Agricultural Engineering***,** vol.66 (1), 62-72.