NONDESTRUCTIVE DETECTION OF MILK FAT CONTENT BASED ON HYPERSPECTRAL TECHNOLOGY**

Q. Huang [1], Z. P. Xu [2], X. H. Jiang [1], J. P. Liu [1], H. R. Xue [1*]

[1] College of Computer and Information Engineering at Inner Mongolia Agricultural University, Hohhot, China; e-mail: xuehr@126.com
[2] Anhui Key Laboratory of Environmental Toxicology and Pollution Control Technology, Hefei Institutes of Physical Science, Chinese Academy of Sciences, Hefei, Anhui, China

Taking the fat content of six different brands of milk as the research object, hyperspectral reflectance data were obtained using hyperspectral imaging and image processing techniques. The raw data are preprocessed in seven different ways. Combine the three mature bionic algorithms – genetic algorithm (GA), ant colony optimization, and particle swarm optimization – with partial least squares (PLS) and support vector machine regression (SVR) models to filter characteristic bands, and to explore the linear and nonlinear relationship between milk spectral data and fat content. The correlation coefficient method, the uninformative elimination algorithm, the successive projection algorithm, the competitive adaptive reweighting sampling algorithm, four mature feature band selection methods, are compared with the bionic algorithm, and, according to the characteristics of each, are combined. The best combination of characteristic bands is selected to establish a regression model to detect milk fat content accurately. The band combination screened by GA and PLS achieved the best prediction results. A total of 72 bands were selected; the correlation coefficient of prediction was 0.9995, and the root mean square error of prediction was 0.0283. The experimental results show that higher accuracy can be obtained by establishing the PLS model using the characteristic bands screened by the linear relationship between spectral data and milk fat content. The SVR model was established based on the nonlinear relationship between spectral data and milk fat content. The accuracy of the SVR model was slightly lower than that of the PLS model. The selection of characteristic bands can improve the model's prediction accuracy, and the use of hyperspectral technology can realize the accurate detection of milk fat content.

Keywords: *milk, fat content detection, hyperspectral, genetic algorithm, feature band selection.*

НЕРАЗРУШАЮЩЕЕ ОПРЕДЕЛЕНИЕ ЖИРНОСТИ МОЛОКА НА ОСНОВЕ ГИПЕРСПЕКТРАЛЬНОЙ ТЕХНОЛОГИИ

Q. Huang [1], Z. P. Xu [2], X. H. Jiang [1], J. P. Liu [1], H. R. Xue [1*]

УДК 535.317.1:637.11

[1] Колледж компьютерной и информационной инженерии в сельском хозяйстве Университета Внутренней Монголии, Хух-Хото, Китай; e-mail: xuehr@126.com
[2] Хэфэйский институт физических наук Китайской академии наук, Хэфэй, Аньхой, Китай

(Поступила 10 июня 2022)

Данные гиперспектральной отражательной способности получены с использованием методов гиперспектральной визуализации и обработки изображений. Объект исследования — содержание жира в шести различных марках молока. Три зрелых бионических алгоритма — генетический алгоритм (GA), оптимизация муравьиной колонии и оптимизация роя частиц — объединены с моделями метода частичных наименьших квадратов (PLS) и вспомогательной векторной машинной регрессии

** Full text is published in JAS V. 90, No. 4 (http://springer.com/journal/10812) and in electronic version of ZhPS V. 90, No. 4 (http://www.elibrary.ru/title_about.asp?id=7318; sales@elibrary.ru).

(SVR) для фильтрации характеристических полос. Исследованы линейная и нелинейная зависимости между спектральными данными молока и содержанием жира. Метод коэффициента корреляции, алгоритм исключения неинформативности, алгоритм последовательной проекции, алгоритм конкурентной адаптивной выборки с повторным взвешиванием, четыре метода выбора зрелых полос признаков сравниваются с бионическим алгоритмом и в соответствии с характеристиками комбинируются. Наилучшая комбинация характеристических полос выбирается для создания регрессионной модели для точного определения содержания жира в молоке. Показано, что комбинация полос, проверенная с помощью GA и PLS, дает наилучшие результаты прогнозирования. Всего отобрано 72 полосы, коэффициент корреляции прогноза составил 0.9995, среднеквадратическая ошибка прогноза 0.0283. Экспериментальные результаты показывают, что более высокая точность может быть получена путем создания модели PLS с использованием характеристических полос, экранированных линейной зависимостью между спектральными данными и содержанием жира в молоке. Модель SVR построена на основе нелинейной зависимости между спектральными данными и содержанием жира в молоке. Точность модели SVR несколько ниже, чем у модели PLS. Выбор характеристических полос повышает точность модели, а использование гиперспектральной технологии обеспечивает точное определение жирности молока.

Ключевые слова: *молоко, определение жирности, гиперспектральный анализ, генетический алгоритм, отбор полос признаков.*

Introduction. Milk has a high nutritional value and is one of the essential nutrients for human health, containing many nutrients that maintain human health, such as water, protein, fat, and lactose. With the increasing standard of living, people's demand for quality milk is becoming higher and higher. Young people tend to choose whole milk for their physical growth needs, some adults choose low-fat milk for their fitness needs, and those in special groups with liver and gall bladder diseases need skimmed milk because of their physical limitations. In today's market, there are strict requirements for testing the fat content of milk and it is particularly important that milk is accurately tested for fat content [1].

Most of the traditional methods of milk fat content detection are based on chemical analysis [2, 3], which is a tedious process requiring professionals to perform the operation, long analysis time, and high cost. In recent years, with the development of spectroscopic instruments and computers, hyperspectral detection techniques have been gradually applied to food [4, 5], medical [6, 7], agricultural [8, 9], and environmental [10, 11] fields. Many scholars at home and abroad have used hyperspectral imaging to detect nutrients in milk. Lim et al. used near-infrared hyperspectral imaging technology combined with partial least squares (PLS) model regression coefficients to detect melamine particles in milk [12]. Zhao et al. applied image processing techniques to analyze hyperspectral data of milk based on hyperspectral imaging to finally establish PLS and N-PLS fat content prediction models [13]. Wang et al. used hyperspectral imaging to detect the content of the additive potassium sorbate in milk [14]. Based on hyperspectral image technology, Liu et al. used a competitive adaptive reweighted sampling (CARS) algorithm and a successive projection algorithm (SPA) to filter characteristic bands and finally established a PLS and support vector regression (SVR) model to predict milk's protein content [15]. Through literature research, it has been found that current research mainly focuses on how to build quantitative analysis models with higher accuracy, and most is based on linear models to predict the content of each component of milk, ignoring the nonlinear relationship between spectral bands and each component of milk, and there are fewer studies on the linear and nonlinear relationships between spectral bands and each component of milk. Studies on the screening of characteristic bands for milk fat properties are even more scarce. Many papers have demonstrated experimentally or theoretically that feature band screening is very important to obtain better prediction performance, and feature band selection is a very critical step to be taken before building calibration models [16–18]. In recent years, more and more band selection methods have been proposed, and in general, these methods can be divided into two main categories. One is based on mathematical statistics, such as correlation coefficient (CC) [19], uninformative variable elimination (UVE) [20], SPA [21], CARS [22, 23], and interval partial least squares method [24]. The other is based on bionic algorithms, such as the genetic algorithm (GA) [25–27], ant colony optimization (ACO) [28, 29], particle swarm optimization (PSO) [30], gray wolf optimization [31], the firefly algorithm [32], and the seagull optimization algorithm [33].

We use hyperspectral imaging to predict the fat content of milk, compare the excellence of different preprocessing methods for milk spectra, and combine GA, ACO, and PSO with PLS and SVR (choosing Gaussian kernel as a kernel function) respectively, to form six GA-PLS, GA-SVR, ACO-PLS, ACO-SVR, PSO-PLS,

and PSO-SVR waveband selection methods. A body check mechanism is also added to GA, and an elite retention strategy is used. These six band selection methods were used to explore the linear and nonlinear relationships between milk spectral reflectance data and milk fat content. The bionic algorithm is also compared with four classical mathematical and statistical methods, CC, UVE, SPA, and CARS, to explore the differences between the bionic algorithm and CC, UVE, SPA, and CARS. The methods are also combined according to their merits, and finally, GA-PLS is combined with SPA and CARS respectively (GA-PLS-PSO, GA-PLS-CARS). The advantages of each band selection method were effectively utilized, and finally, the best feature band selection method was selected to build a milk fat content prediction model from the screened data.

Materials and methods. Six different brands of milk, namely YiLi QQ Star, MengNiu High Calcium, YiLi Skimmed milk, TeLunSu, TheLand, and YiLiZhenNong, were selected as experimental samples with fat concentrations of 3.7, 3.7, 0, 4.4, 3.4, and 4.6 g/mL, respectively.

The hyperspectral image acquisition equipment is a hyperspectral imager (1003B-10141) from Headwall, USA, which consists of a hyperspectral imager, a quartz tungsten halogen illuminator, a mechanical scanning platform, a computer, and spectral acquisition software. The spectral range of measurement is 400–1000 nm with a resolution of 2.8 nm and 125 bands.

During the experimental process, in order to eliminate the uneven distribution of light intensity and the noise generated by dark currents in the camera sensor, black-and-white correction processing is required before each acquisition of experimental images [13]. The black background information (G_B) was first captured with the camera cover closed, then the white background information (G_W) was captured with a standard whiteboard, and the acquired hyperspectral image (G_R) was black and white corrected using Eq. (1) to obtain the final experimental image (G). Each sample was acquired three times and the clearest one was selected as the final experimental image:

$$G = \frac{G_R - G_B}{G_W - G_B}. \quad (1)$$

Although milk is liquid, it is not completely homogeneous and can be spectrally analyzed for regions of interest. Using ENVI software, 50 regions of interest with uniform light distribution were selected from the acquired hyperspectral images of milk samples, and finally, 300 samples were accumulated. The average of the spectral reflectance data of each region was exported as sample experimental data for spectral preprocessing operation.

In the image acquisition process, although the black-and-white correction operation is carried out, there is still some noise caused by other factors in the required spectral data, and it is necessary to carry out data pre-processing operations on the acquired spectral data. The vector normalization (VN), Savitzky–Golay (SG) smoothing, multiplicative scatter correction (MSC), standard normal variate transformation (SNVT), first derivative (1st Der), second derivative (2nd Der), and wavelet transform (WT) are selected to pre-process the original spectral data [34]. The method with the highest prediction accuracy was selected for subsequent experiments.

Spectral data have a large amount of information and a large number of bands, and there is a lot of redundant information and bands that are irrelevant to the attribute of interest. On the one hand, the prediction accuracy of the model is reduced, and on the other hand, the calculation time is increased [35, 36]. Therefore, it is necessary to select the characteristic band of the measured data before establishing the prediction model. The band combination itself is an NP-hard problem. As the number of bands increases, the number of possible band combinations increases exponentially. It is particularly important to choose a good band combination for spectral data modeling.

Three classical bionic algorithms, GA, ACO, and PSO, are selected to combine with PLS and SVR models for band screening respectively, to reduce the number of bands required for modeling, to explore the linear and nonlinear relationships between spectral data of milk and fat content in milk, and to improve model accuracy and stability. The fitness values are all calculated according to Eq. (2), where the larger and the smaller the RMSEP, the larger the fitness value and the higher the model accuracy. PLS is a typical linear regression model. After repeated iterations, GA-PLS, ACO-PLS, and PSO-PLS screened out the bands in which spectral data of milk correlated linearly with fat content. SVR is a typical nonlinear regression model. After repeated iterations, GA-SVR, ACO-SVR, and PSO-SVR screened out the bands of nonlinear correlation between milk spectral data and fat content. In GA, the final experimental results have an important relationship with the generation of the initial population, and the quality of the initial population determines the upper limit of the final experimental results. This paper proposes that the physical examination mechanism is

applied before the selection operator. By setting the threshold of physical examination, individuals with low fitness values are filtered, and individuals with high fitness values for subsequent operations are selected.

$$F = \frac{R_P^2}{\mathrm{RMSE}_P}. \quad (2)$$

In addition, four classical mathematical statistics methods, CC, UVE, SPA, and CARS, were selected to choose characteristic bands for full-band data. Because CC, UVE, SPA, and CARS reflect the linear relationship between the bands and attribute values, the PLS regression model is established on the band data screened by these methods. Finally, compared with GA-PLS, ACO-PLS, and PSO-PLS, the differences in various characteristic band selection methods in milk spectral data were explored.

After the experiment, it was found that the data after GA-PLS, ACO-PLS, and PSO-PLS feature band selection had a good prediction effect, but the number of selected bands was high, among which the GA-PLS algorithm had the best prediction accuracy. The data filtered by the SPA and CARS algorithms are less accurate than GA-PLS in prediction, although the number of bands is smaller. Based on this, the GA-PLS algorithm is combined with SPA and CARS (GA-PLS-SPA, GA-PLS-CARS) respectively, and the GA-PLS filtered data are used for secondary feature band selection with SPA and CARS respectively.

The nine feature band selection methods, GA-PLS, ACO-PLS, PSO-PLS, CC, UVE, SPA, CARS, GA-PLS-SPA, and GA-PLS-CARS, are based on the linear relationship between data and attribute values to screen the feature bands, and the screened band data are used to build PLS regression models to predict milk fat content. The three feature band selection methods, GA-SVR, ACO-SVR, and PSO-SVR, are based on the nonlinear relationship between data and attribute values to screen the feature bands, and the screened band data are used to build SVR regression models to predict milk fat content. The correlation coefficient of calibration $\left(R_C^2\right)$, the correlation coefficient of prediction $\left(R_P^2\right)$, the root mean square error of calibration (RMSEC), and the RMSEP were selected as the evaluation criteria for the regression models.

Results and discussion. *Spectral data analysis and data pre-processing.* Milk is a mixed liquid composed of various components. The main components are nutrients such as water, fat, protein, and lactose. The diversity of components leads to complex and changeable spectral curves. The original spectral data of six different brands of milk are shown in Fig. 1a. From the figure, it can be seen that the spectral curve changes are complex, and it is difficult to see the law of the spectral curve change with the naked eye. The collected 300 sample data are randomly divided into training and prediction sets in the ratio of 3:1 by using the Split function. The raw data were preprocessed with VN, SG, MSC, SVNT, 1stDer, 2ndDer, and WT respectively, and the preprocessed data were used to build PLS and SVR regression models. The modeling results of the preprocessed data and the original data are shown in Table 1. From Table 1, it can be seen that the MSC preprocessed data yielded the best results in terms of model accuracy compared with other data preprocessing methods, both the PLS regression model and the SVR regression model. Milk is a light-scattering liquid with an uneven distribution of emulsions, which makes physical sense using the MSC method. Compared with the original data, the PLS model increased from 0.9975 to 0.9988 and RMSEP decreased from 0.0698 to 0.0468. The SVR model increased from 0.9936 to 0.9938 and RMSEP decreased from 0.1116 to 0.1096. All preprocessed data achieved good results on the PLS model, with and above 0.99,

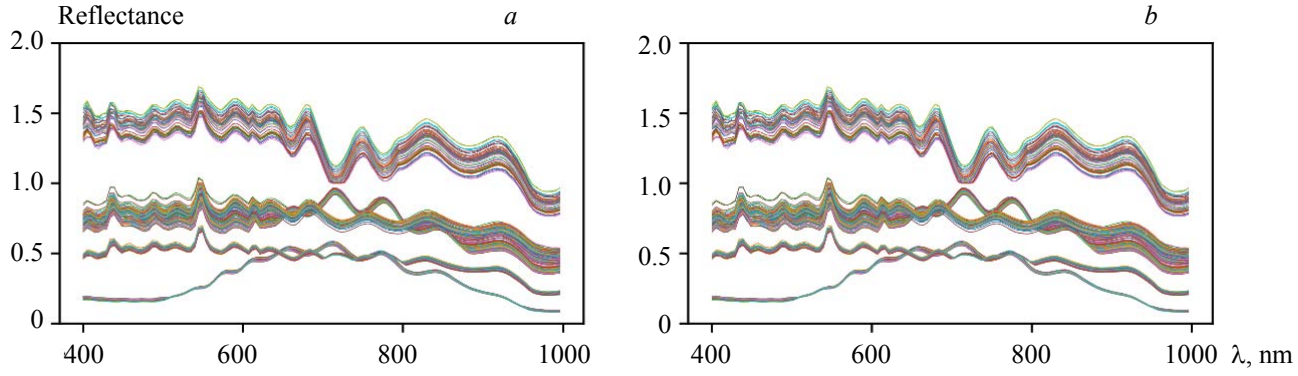

Fig. 1. (a) Spectral curves of original milk samples; (b) spectral profile of milk samples after MSC pre-treatment.

and RMSEC and RMSEP below 0.1 except for 2ndDer. Data preprocessed by methods other than MSC are not suitable for the SVR model, and the modeling accuracy is worse than with the original data. Finally, the MSC-preprocessed data were selected for feature variable screening. The spectral data of six different brands of milk pretreated with MSC are shown in Fig. 1b. Comparing Fig. 1a, and b, it is difficult to see any change in the preprocessed data, which should confirm the phenomenon that the modeling effect of the raw spectral data and the MSC pre-processed data is approximately the same.

TABLE 1. Results of PLS and SVR Prediction of Milk Fat Content by Different Pretreatment Methods

Means	R_C^2	R_P^2	RMSEC	RMSEP
Raw-PLS	0.9996	0.9975	0.0292	0.0698
VN-PLS	0.9997	0.9985	0.0265	0.0539
SG-PLS	0.9994	0.9974	0.0372	0.0712
MSC-PLS	0.9995	0.9988	0.0329	0.0468
SNVT-PLS	0.9994	0.9973	0.0359	0.0726
1stDer-PLS	0.9996	0.9958	0.0301	0.0899
2ndDer-PLS	0.9994	0.9933	0.0382	0.1144
WT-PLS	0.9996	0.9973	0.0304	0.0717
RAW-SVR	0.9934	0.9936	0.1274	0.1116
VN-SVR	0.9865	0.9822	0.1828	0.1866
SG-SVR	0.9924	0.9927	0.1370	0.1192
MSC-SVR	0.9952	0.9938	0.1086	0.1096
SNVT-SVR	0.7633	0.7655	0.7659	0.6777
1stDer-SVR	0.9864	0.9776	0.1835	0.2090
2ndDer-SVR	0.9624	0.9488	0.3052	0.3166
WT-SVR	0.9933	0.9935	0.1280	0.1124

Characteristic band selection results and analysis. The data after MSC preprocessing were screened for characteristic bands, and the GA-PLS, ACO-PLS, and PSO-PLS screened data were used to build a PLS regression model to predict milk fat content. The screened data from GA-SVR, ACO-SVR, and PSO-SVR were used to build SVR models to predict milk fat content; the specific results are shown in Table 2. From the linear relationship between milk spectral data and fat content to screen the characteristic bands, compared with the full-band data, the data after GA-PLS screening of bands was not significantly changed, but was reduced by 39.43%, and the number of required bands was only 57.6% of the original. The PLS regression modeling of the ACO-PLS-filtered data also yielded good results, with a 25.15% reduction in only 55.2% of the original number of bands required. PSO-PLS does not perform particularly well, with only a 5.84% reduction compared with full-band data PLS modeling, and the number of bands required is 65.6% of the original. From the nonlinear relationship between spectral data of milk and fat content to screen the characteristic bands, the RMSEP was reduced by 40.12% for the GA-SVR screened data to build the SVR regression model compared to the full-band data, and the number of bands required was only 24.8% of the original. The RMSEP of the ACO-SVR filtered data was reduced by 29.49% and the number of required bands was 28.8% of the original. The PSO-SVR algorithm does not perform particularly well, with only a 9.03% reduction in RMSEP and the number of required bands needing 64% of the original.

TABLE 2. Experimental Results of Band Selection for Different Bionic Algorithms

Featured selection methods	Number of bands	R_C^2	R_P^2	RMSEC	RMSEP
GA-PLS	72	0.9995	0.9995	0.0340	0.0283
ACO-PLS	69	0.9993	0.9993	0.0411	0.0350
PSO-PLS	82	0.9995	0.9990	0.0331	0.0435
GA-SVR	31	0.9977	0.9977	0.0742	0.0656
ACO-SVR	36	0.9973	0.9969	0.0816	0.0773
PSO-SVR	80	0.9949	0.9959	0.0999	0.0997

Among the three bionic algorithms, the GA performed the best in terms of both linear and nonlinear relationships between milk spectral data and milk fat content, with the best prediction model accuracy and a moderate number of bands required. The bands screened by GA-PLS and GA-SVR are shown in Fig. 2. GA-PLS selected 72 bands and GA-SVR selected 31 bands. Eighteen of the bands selected by both are the same, and 13 of the bands of GA-SVR are completely independent. It shows that most of the bands maintain a linear correlation with the fat content of the milk between the spectral data and the fat content, and some bands maintain a nonlinear relationship with the fat content. It can be seen from Fig. 2 that most of the bands screened by GA-PLS are uniformly distributed at a peak or trough of the spectral curve. The bands selected by GA-SVR are mainly around 405, 740, and 800 nm; among them, 740 nm corresponds to the quadruple frequency absorption band of the O-H bond, and 800 nm corresponds to the quadruple frequency absorption band of the N-H bond. Other selected bands are difficult to match to a specific chemical bond, but experimental results show that these bands play a key role in the modeling.

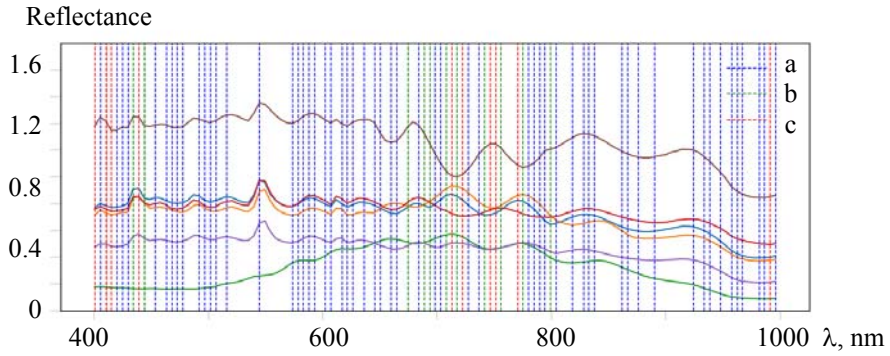

Fig. 2. Feature band selection diagram. Dotted line a: bands selected individually by GA-PLS; Dotted line b: bands selected individually by GA-SVR; dotted line c: GA-PLS and GA-SVR select the same band.

The MSC preprocessed data were screened for characteristic bands using four classical mathematical and statistical methods, CC, UVE, SPA, and CARS. Because these four methods use the nonlinear relationship between data and attribute values to screen the characteristic bands, the screened data were used to build a PLS regression model to predict milk fat content, and the experimental results are shown in Table 3. A total of 117 bands were selected for CC, and the model accuracy was improved to some extent, but the performance was less satisfactory than with other methods. The UVE algorithm selects 72 bands, requiring 57.6% of the original number of bands and reducing the RMSEP by 6.28%. The SPA algorithm selects a total of 37 bands, among all the methods, the SPA algorithm selects the least number of bands, but the model accuracy is the lowest, and the accuracy of the model is slightly reduced compared with the full-band model. The CARS algorithm performs the best among the four mathematical and statistical methods, with 43 bands selected, which is only 43.2% of the full band, whereas the RMSEP is reduced by 27.24%.

TABLE 3. Experimental Results of Four Kinds of Mathematical Statistics Methods for Feature Selection

Feature selection methods	Number of bands	R_C^2	R_P^2	RMSE_C	RMSE_P
CC	117	0.9994	0.9988	0.0352	0.0467
UVE	72	0.9997	0.9990	0.0230	0.0438
SPA	37	0.9992	0.9984	0.0430	0.0551
CARS	54	0.9997	0.9994	0.0263	0.0340

Observing Tables 2 and 3, it can be seen that among all the methods based on linear relationship screening bands, the biomimetic algorithm performs better in the prediction accuracy of milk fat content. In particular, GA-PLS has the best prediction results among all methods. The mathematical–statistical method runs faster and requires fewer bands than the bionic algorithm, mainly in the SPA and CARS algorithms, but the prediction accuracy is not as good as in the bionic algorithm. To investigate the optimal band for milk fat content prediction in a closer way, GA-PLS was combined with SPA and CARS algorithms, respectively.

The GA-PLS-filtered band data are screened again with the SPA and CARS algorithms for the characteristic bands respectively. The filtered band data were used to build a PLS fat content prediction model, and the experimental results are shown in Table 4. GA-PLS-SPA better integrates the advantages of GA-PLS and SPA, the number of bands is only 60.86% of GA-PLS, and the RMSEP is 32.54% lower than that of SPA. Compared with the full band, the number of bands is only 34.4% of the full band, and the $RMSE_P$ is reduced by 20.51%; GA-PLS-CARS also achieved good results with a band count of 89.85% of GA-PLS and a 10.84% reduction in RMSEP compared with CARS. GA-PLS-PSO and GA-PLS-CARS make a good combination of the number of bands and model prediction accuracy. In practical applications, the method of feature band screening can be chosen flexibly in terms of both prediction accuracy and prediction speed.

TABLE 4. Experimental Results of Screening Characteristic Bands by Combining Bionic Algorithm and Mathematical Statistics Method

Feature selection methods	Number of bands	R_C^2	R_P^2	$RMSE_C$	$RMSE_P$
GA-PLS-SPA	42	0.9995	0.9992	0.0346	0.0371
GA-PLS-CARS	62	0.9995	0.9995	0.0321	0.0303

Prediction results and analysis of milk fat content. The linear regression model PLS and nonlinear regression model SVR were selected to predict milk fat content, and the screened data from GA-PLS and GA-SVR were used to build a milk fat content prediction model. In order to compare the prediction results of the two models more intuitively, the deviation and variance between the predicted values of the two models and the actual value are calculated, and finally a histogram is drawn as shown in Fig. 3. C_Bias, T_Bias, C_Var, and T_Var represent the bias and variance of the calibration set and test set respectively. Bias reflects the accuracy of model predictions, whereas variance measures the stability of the model. On the three brands of Yili QQ Star, TheLand and Yili Skimmed milk, the accuracy and stability of the prediction results of the PLS model are better. The prediction effect of the SVR model on Telunsu brand milk is not very ideal, and the accuracy and stability are far lower than other brands, whereas the PLS model performs very well. On the two brands, Mengniu High Calcium and YiliZhenNong, the SVR model is more stable than the PLS model. Overall, the accuracy and stability of the PLS model are better.

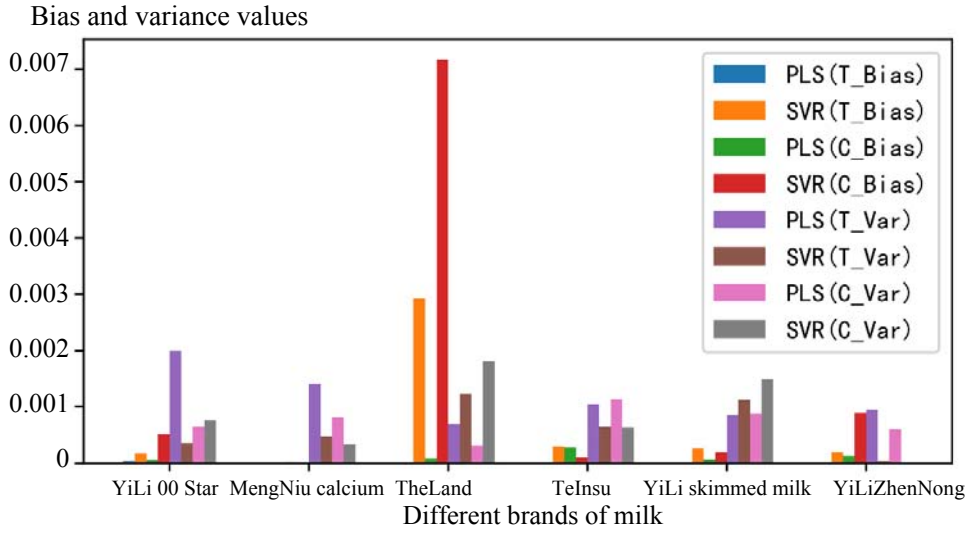

Fig. 3. The bias and variance of the predicted and true values of the PLS and SVR models.

Conclusions. A study of milk fat content detection methods using hyperspectral imaging technology was carried out to achieve nondestructive detection of milk fat content by combining spectral reflectance data with image information. The experimental results show that the model prediction accuracy can be improved by pre-processing the raw data, but not all pre-processing methods are suitable for milk spectral data, and some pre-processing methods may degrade the model. In milk production applications, different band

screening methods can be selected according to production needs. The GA-SVR can quickly and effectively identify the bands that have a nonlinear relationship with milk fat content in the spectral data, further improving the model accuracy and preventing overfitting. In terms of model pre-accuracy the PLS regression model is more accurate, but in terms of the number of bands, SVR requires fewer. The use of hyperspectral image detection technology can achieve accurate detection of milk fat content well, providing a new solution for the nondestructive detection of milk fat content.

Acknowledgements. The authors acknowledge financial support from the National Natural Science Foundation of China (No. 61461041), National Natural Science Foundation of China (No. 31960494), Inner Mongolia Autonomous Region Major Science and Technology Project (No. 2021ZD0003), Natural Science Foundation of Inner Mongolia Autonomous Region (No. 2022MS06026).

REFERENCES

1. S. L. Li, K. Yao, Z. J. Cao, C. Q. Liu Y. C. Wang, Z. H. Wang, J. X. Li, J. Q. Wang, H. P. Zhang, F. L. Kong, *Chin. J. Animal Sci.*, **58**, No. 3, 239–244 (2022).
2. D. M. Whitt, J. Pranata, B. G. Carter, D. M. Barbano, M. A. Drake, *J. Dairy Sci*., **105**, No. 7, 5700–5713 (2022).
3. L. Di Marzo, J. Pranata, D. M. Barbano, *J. Dairy Sci*., **104**, No. 7, 7448–7456 (2021).
4. Y. Y. Shao, Y. K. Shi, Y. D. Qin, G. T. Xuan, J. Li, Q. K. Li, F. J. Yang, Z. C. Hu, *Food Chem*., **386**, 132864 (2022).
5. Y. Q. Ren, D. W. Sun, *Food Chem*., **382**, 132346 (2022).
6. A. Panda, R. B. Pachori, N. Kakkar, M. Joseph John, N. D. Sinnappah-Kang, *Comp. Methods Programs Biomed*., **220**, 106836 (2022).
7. S. Karim, A. Qadir, U. Farooq, M. Shakir, A. A. Laghari, *Curr. Med. Imaging*, **19**, No. 5, 417–427 (2022), doi: 10.2174/1573405618666220519144435.
8. A. Kayad, F. A. Rodrigues, S. Naranjo, M. Sozzi, F. Pirotti, F. Marinello, U. Schulthess, P. Defourny, B. Gerard, M. Weiss, *Field Crops Res*., **282**, 108449 (2022).
9. L. Zheng, Q. Bao, S. Z. Weng, J. P. Tao, D. Y. Zhang, L. S. Huang, J. L. Zhao, *Spectrochim. Acta A: Mol. Biomol. Spectrosc*., **270**, 12084 (2022).
10. X. Q. Pan, J. B. Jiang, Y. M. Xiao, *Ecolog. Inform*., **68** (2022).
11. H. D. Zhang, J. Luo, S. M. Hou, Z. P. Xu, J. L. Evans, S. L. He, *Appl. Opt*., **61**, No. 12, 3400–3408 (2022).
12. J. Lim, G. Kim, C. Mo, M. S. Kim, K. Chao, J. Qin, X. Fu, I. Baek, B.-K. Cho, *Talanta*, **151**, 183–191 (2016).
13. Z. J. Zhao, Y. Wei, N. Q. Zhang, R. K. Chang, H. Y. Wu, H. Liu, H. Y. Shan, R. J. Yang, X. Y. Guo, *China Dairy Industry*, **46**, No. 2, 45–48 (2018).
14. Y. Wang, Y. Huang, W. Shen, F. Kong, M. Gao, H. Sun, *Spectr. Lett*., **54**, No. 4, 316–325 (2021).
15. M. C. Liu, H. R. Xue, J. P. Liu, R. R. Dai, P. W. Hu, Q, Huang, X. H. Jiang, *Spectrosc. Spectr. Anal.*, **42**, No. 5, 1601–1606 (2022).
16. M. Zhang, S. Zhang, J. Iqbal, *Chemometr. Intell. Lab. Systems*, **128**, 17–24 (2013).
17. P. Mishra, E. J. Woltering, *Talanta*, **224**, 121908 (2021).
18. R. M. Balabin, S. V. Smirnov, *Analyt. Chim. Acta*, **692**, No. 1–2, 63–72 (2011).
19. M. Radman, M. Moradi, A. Chaibakhsh, M. Kordestani, M. Saif, *IEEE Sensors J.*, **21**, No. 3, 3533–3543 (2021).
20. V. Centner, D. L. Massart, O. E. de Noord, S. de Jong, B. M. Vandeginste, C. Sterna, *Anal. Chem.*, **68**, No. 21, 3851–3858 (1996).
21. R. K. H. Galvao, M. C. Ugulino Araujo, W. D. Fragoso, E. C. Silva, G. E. Jose, S. F. Carreiro Soares, H. M. Paiva, *Chemometr. Intell. Lab. Systems*, **92**, No. 1, 83–91 (2008).
22. H.-Y. Zhen, R.-J. Ma, Y. Chen, X.-P. Sun, C.-L. Ma, *Spectrosc. Spectr. Anal.*, **40**, No. 5, 1601–1606 (2020).
23. H. Li, Y. Liang, Q. Xu, D. Cao, *Anal. Chim. Acta*, **648**, No. 1, 77–84 (2009).
24. F. Allegrini, A. C. Olivieri, *Anal. Chim. Acta*, **699**, No. 1, 18–25 (2011).
25. S. Cateni, V. Colla, M. Vannucci, *IEEE, 9th Int. Conf. Intell. Systems Design Appl*., 1278–1283 (2009).
26. R. G. Zhu, H. W. Duan, X. D. Yao, Y. Y. Qiu, B. X. Ma, C. J. Xu, *Spectrosc. Spectr. Anal.*, **36**, No. 9, 2925–2929 (2016).

27. Z. H. Tu, B. P. Ji, C. Y. Meng, D. Z. Zhu, B. L. Shi, Z. S. Qing, *Spectrosc. Spectr. Anal.*, **29**, No. 10, 2760–2764 (2009).
28. M. Shamsipur, V. Zare-Shahabadi, B. Hemmateenejad, M. Akhond, *J. Chemometr.*, **20**, No. 3–4, 146–157 (2006).
29. T. Liu, T. Xu, F. Yu, Q. Yuan, Z. Guo, B. Xu, *Comp. Electron. Agric*., **186** (2021).
30. F. Marini, B. Walczak, *Chemometr. Intell. Lab. Systems*, **149**, 153–165 (2015).
31. Y. Z. Hou, X. Gao, S. N. Li, X. Cai, P. Li, W. L. Li, Z. Li, *J. Pharmac. Innovation*, **17**, No. 4 (2022), https://doi.org/10.1007/s12247-022-09620-6.
32. H. Xu, S. Yu, J. Chen, X. Zuo, *Wireless Personal Commun.*, **102**, No. 4, 2823–2834 (2018).
33. L. Xu, Y. Mo, Y. Lu, J. Li, *Processes*, **9**, No. 6, 1037 (2021).
34. P.-Y. Diwu, X.-H. Bian, Z.-F. Wang, W. Liu, *Spectrosc. Spectr. Anal.*, **39**, No. 9, 2800–2806 (2019).
35. M. Li, Y. Feng, Y. Yu, T. Zhang, C. Yan, H. Tang, Q. Sheng, H. Li, *Spectrochim. Acta A: Mol. Biomol. Spectrosc*., **257** (2021).
36. M. M. Galera, D. P. Zamora, J. L. M. Vidal, A. G. Frenich, *Analyt. Lett*., **35**, No. 5, 921–941 (2002).