

Article

Not peer-reviewed version

Research on the Detection Method of Organic Matter in Tea Garden Soil based on Image Information and Hyperspectral Data Fusion

[Haowen Zhang](#) , Chongshan Yang , Min Lu , Zhongyuan Liu , [Xiaojia Zhang](#) ^{*} , [Chunwang Dong](#) ^{*}

Posted Date: 8 November 2023

doi: 10.20944/preprints202311.0447.v1

Keywords: Hyperspectral; machine visualization properties; data fusion; tea plantation soils; organic matter



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

Research on the Detection Method of Organic Matter in Tea Garden Soil based on Image Information and Hyperspectral Data Fusion

Haowen Zhang ¹, Chongshan Yang ¹, Min Lu ¹, Zhongyuan Liu ¹, Xiaojia Zhang ^{1,*} and Chunwang Dong ^{1,2,*}

¹ Tea Research Institute, Shandong Academy of Agricultural Sciences, Jinan 250100, China; 18853875771@163.com (H.Z.); yang1029345485@163.com (C.Y.); lumin_1994@163.com (M.L.); lzy13112350601@163.com (Z.L.)

² College of Mechanical and Electrical Engineering, Shihezi University, Shihezi 832000, China

* Correspondence: zhangxiaojia@saas.ac.cn (X.Z.); dongchunwang@163.com (C.D.)

Abstract: Soil organic matter is an important component that reflects soil fertility and promotes plant growth. The soil of typical Chinese tea plantations was used as the research object in this work, and by combining soil hyperspectral data and image texture characteristics, a quantitative prediction model of soil organic matter based on machine vision and hyperspectral imaging technology was built. Three methods, standard normalized variate (SNV), multisource scattering correction (MSC) and smoothing, were first used to preprocess the spectra. After that, random forest (RF), variable combination population analysis (VCPA) and variable combination population analysis and iterative retained information variable algorithm (VCPA-IRIV) algorithms were used to extract the characteristic bands. Finally, the quantitative prediction model of nonlinear support vector regression (SVR) and linear partial least squares regression (PLSR) for soil organic matter was established by combining nine color features and five texture features of hyperspectral images. The outcomes demonstrate that, in comparison to single spectral data, fusion data may greatly increase the performance of the prediction model, with MSC+VCPA-IRIV+SVR ($R^2_c=0.995$, $R^2_p=0.986$, $RPD=8.155$) being the optimal approach combination. This work offers excellent justification for more investigation into nondestructive methods for determining the amount of organic matter in soil.

Keywords: hyperspectral; machine visualization properties; data fusion; tea plantation soils; organic matter

1. Introduction

Worldwide, tea plants are produced as a significant revenue crop [1,2]. Because of its distinct flavor and scent, tea is the most consumed nonalcoholic beverage in the world [3]. The organic matter in the soil has a large influence on how the tea tree grows. The whole soil ecology depends heavily on soil organic matter, which also helps with nitrogen fixation, carbon sequestration, and plant nutrient retention [4]. As a result, accurate information on the amount of organic matter in the soil is crucial for enhancing tea plants' ability to thrive. However, conventional techniques for determining soil organic matter are labor intensive, costly, time consuming, and leave behind dangerous chemical residues. The quick detection of soil organic materials during the growth of tea trees cannot be met by this [5].

Machine vision and hyperspectral imaging are two nondestructive testing techniques that are widely utilized in many industries, including forestry [6], agriculture [7], animal husbandry [8], and fisheries [9]. To explore the impact of moisture variations on soil urease activity, Chenbo Yang et al. [10] developed a prediction model for soil urease activity based on hyperspectral data, of which the PLSR model was the most successful (R^2_v of 0.8564). Rahimi-Ajdadi et al. [11] employed soil

photographs to estimate the soil water content using feature-finding in several color spaces, and their findings had an average absolute error of less than 1.1%. However, data from a single source may not be perfect in defining some qualities, and individuals have begun to employ fused data in the field of NDT to further enhance target identification. Ting An et al. [12] collected visual (color pictures) and olfactory (sensor array spectra) data on tea leaves during the fermentation of black tea and combined them with hyperspectral data to create a discriminative model for the level of fermentation that had a 95% prediction set accuracy. Yingqian Yin et al. [13] combined black tea image features with hyperspectral features to characterize different grades of black tea. They then created discriminant models using partial least squares discriminant analysis (PLS-DA), support vector machines (SVM), and probabilistic neural networks (PNN), with the SVM model having the highest validation set accuracy (98.33%). On the other hand, few studies have been performed on the identification of soil organic matter utilizing fused data. The efficacy of the hyperspectral and machine vision data-fused quantitative soil organic matter prediction model is still being assessed.

Based on the aforementioned research state, the soil of typical Chinese tea plantations was chosen as the research object in this work, and the soil's spectrum information and color and texture attributes were retrieved using the obtained hyperspectral pictures. To combine the spectral and color texture features of the soil, middle-level data fusion is used. From there, a quantitative prediction model of soil organic matter is established to find the best algorithmic strategy, which establishes a theoretical foundation for future advancements in the nondestructive detection technology of soil organic matter.

2. Materials and Methods

2.1. Collection of Soil Samples and Estimation of the Organic Matter Content

The soil samples utilized in this study were gathered in the Shandong Province locations of Jufeng Town, Rizhao City; Lanshan District, Qingdao City, and Junan County, Linyi City. In all of the municipal tea gardens, 15 separate soil collection locations with a combined sampling area of 25 m² and a depth of 0~40 cm were chosen. Soil samples were then taken ten times from each area for a final total of 150 soil samples.

The gathered samples were taken back to the lab at the Shandong Academy of Agricultural Sciences in Jinan, Shandong Province, China, for the analysis of the organic matter content. To ensure that each soil sample weighed the same weight, contaminants such as stones and animal and plant remains were cleaned before drying. Following that, the criteria for assessing soil organic matter content (NY/T 1121.6-2006) were used to evaluate the organic matter content of all soil samples. Weigh each soil sample that passes through a screen with an aperture of 0.25 mm at 0.05~0.5 g, oxidize the soil's organic carbon with an excessive potassium dichromate sulfuric acid solution, and then titrate the excess potassium dichromate with a ferrous sulfate standard solution. The quantity of organic carbon was determined by converting the amount of potassium dichromate consumed into organic carbon using the oxidation correction coefficient and then multiplying the result by the constant 1.724 to obtain the amount of organic matter in the soil. The organic matter content of each sampling area is shown in Figure 1.

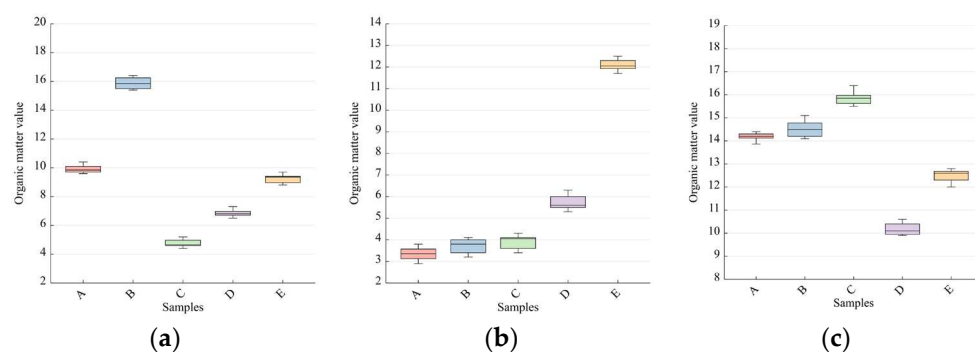


Figure 1. Soil organic matter content in each sampling area. (a) Soil samples of tea plantations in Rizhao; (b) Soil samples of tea plantations in Qingdao; (c) Soil samples of tea plantations in Linyi.

2.2. Hyperspectral Image Acquisition

ISpecHyper-VS1000-Lab hyperspectral imaging equipment from Lyson Optics was used to gather hyperspectral data. The imaging system consists of an integrated hyperspectral dark box, a hyperspectral camera, a high-performance homogeneous light source, a diffuse reflectance standard version (3%/50%) and a linear displacement sample stage. The wavelength acquisition range is 400~1000 nm with a spectral resolution of 2.5 nm. Before sampling, the machine was warmed for 30 minutes to establish a stable atmosphere. After preheating was completed, the camera lens was covered with a lens cap for dark current collection. For consistent tiling, 50 g of each soil sample was weighed and placed in a test dish. The images were placed within the dark box of the hyperspectral image capture system. Before each sample, the system obtains a calibration picture of the reference plate within the dark box. The soil samples were eventually subjected to 150 hyperspectral pictures.

2.3. Extraction of Spectral Features and Picture Features

Using the aforementioned collected hyperspectral pictures, the spectral information of the soil samples was retrieved. On the hyperspectral picture of each soil sample, 10 ROI sections were chosen at random. The spectral data inside the regions are extracted, and the average of the ten sets of data is used as the soil sample's spectral data. The ENVI 5.3 program was used to carry out this procedure. The collected hyperspectral images were adjusted with the following formula to produce precise reflectance profiles of the samples:

$$I_{cor} = \frac{I_{raw} - I_{dark}}{I_{white} - I_{dark}} \quad (1)$$

where I_{cor} is the hyperspectral image that has been rectified, I_{raw} is the picture of the sample that was acquired, I_{dark} is the image of the dark current that was acquired, and I_{white} is the image of the diffuse reflective plate.

The image processing program created by the MATLAB GUI module was used to extract the soil's image characteristics (software copyright number: 2014SR149549). Each hyperspectral image contained 10 randomly chosen zones of interest for the soil, from which the color and texture attributes were retrieved. The Dong et al. [14] algorithm was then used to convert the RGB color system to the HSV color space. To produce 9 color features and 5 texture features for each image, the data from the 10 regions of interest of each image were averaged. The red channel mean (R), green channel mean (G), blue channel mean (B), color point mean (H), saturation mean (S), luminance mean (V), supergreen transform (2G-R-B), ratio of red channel mean to green channel mean (R/G), and color (hab*) are among the color features; the mean gray value (m), standard deviation (δ), smoothness (r), consistency (U), and entropy (e) are among the texture features.

2.4. Preprocessing of the Spectral Data and Characteristic Band Screening

The system may be impacted by external conditions during the measurement process. To improve the spectrum features and remove superfluous information such as baseline drift and noise generated during the measurement process, the spectral data are preprocessed. Three preprocessing methods are used, including standard normalized variate (SNV), multisource scattering correction (MSC) and smoothing.

The wavelength range of the hyperspectral imaging equipment employed in this investigation is 400~1000 nm. The wavelength range of the gathered soil spectra is 344~986 nm, with a resolution of less than 2.5 nm. There are 300 wavelength points in all, which is a significant quantity of duplicated data. As a result, distinctive bands were searched for in the gathered spectral data. In this article, three feature band screening techniques, random frog (RF) [15], variable combination population analysis (VCPA) [16] and variable combination population analysis and iterative retained

information variable algorithm (VCPA-IRIV) [17], are chosen to filter out irrelevant data and further enhance the model accuracy.

In modeling, data overfitting can occur often, although feature band screening does not always completely prevent it. As a result, further feature extraction from the data is needed using feature extractions. After feature wavelength screening, the principal component analysis (PCA) dimensionality reduction approach was utilized in this work to further reduce the dimensionality of the data. Using mapping, PCA rearranged the data while keeping the primary spectral characteristics and condensing the many feature variables into a handful of key features [18,19].

2.5. Data Fusion

To improve the connectivity between several datasets and provide better modeling outcomes, data fusion is used. In this work, the soil's spectral and color texture data are extracted from the obtained hyperspectral pictures of the soil, and the two separate data are combined using a data fusion technique to produce superior prediction outcomes. Low-level fusion, middle-level fusion, and high-level fusion are the three different types of data fusion methodologies. Among these, low-level fusion merely splices the two data matrices and does not yield superior outcomes. If the high-level fusion is not handled carefully, some crucial information may be lost [20]. As a result, the modeling dataset in this work is processed using the intermediate-level fusion approach. The resultant color texture feature matrix is spliced with the spectral data after it has undergone feature wavelength screening and PCA dimensionality reduction, and the newly formed matrix is then normalized to obtain a new data matrix. The newly acquired matrix will be applied to other modeling tasks. The broad idea of data fusion is shown in Figure 2.

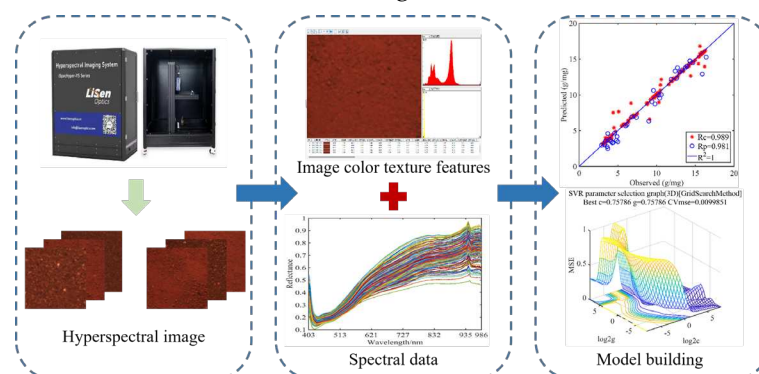


Figure 2. Flowchart of data fusion concepts.

2.6. Model Building and Evaluation Criteria

The dataset was split into training and validation sets using the KS technique at a ratio of 3:1 after acquiring the fusion data of soil spectral characteristics and color texture features. The dataset consists of 150 sets of fusion data, including 112 sets of fusion data in the training set and 38 sets in the validation set. A quantitative prediction model for soil organic matter using linear PLSR and nonlinear SVR models was created on the basis of this information. To see how fused data affected the forecast of soil organic matter content, the model of single source data and the model of fused data were compared. The correlation coefficient (RC), correlation root mean square error (RMSEC), prediction set correlation coefficient (RP), prediction root mean square error (RMSEP), and relative percentage of deviation (RPD) were employed in this work as assessment indices for the correction and prediction sets [21]. The more closely the RC and RP approach 1, the more accurate the model is; similarly, the more closely the RMESC and RMESP converge, and the more closely they diverge, the more generalizable the model is. Typically, it is believed that if RPD is less than 1.4, the model is subpar and cannot be used for prediction research, that if RPD is between 1.4 and 1.8, the model can achieve its prediction objective, but accuracy still needs to be improved, and that if RPD is larger than 2, the model has good prediction performance [22].

3. Results

3.1. Spectral Preprocessing Results

The soil spectral data were collected at 300 wavelength locations between 344 and 986 nm. The spectral data for the wavelength range of 344–402 nm were discarded once it was discovered via observation that the spectral data in this region had more significant noise. Three preprocessing methods, MSC, SNV and smooth, were used to preprocess the spectral data to compare the modeling effect under different preprocessing methods. The spectrum data were preprocessed using three preprocessing techniques, MSC, SNV, and smooth, to examine the modeling impact under various preprocessing techniques. The spectral profiles following the three preprocesses are shown in Figure 3, together with the soil sample's original spectral profile.

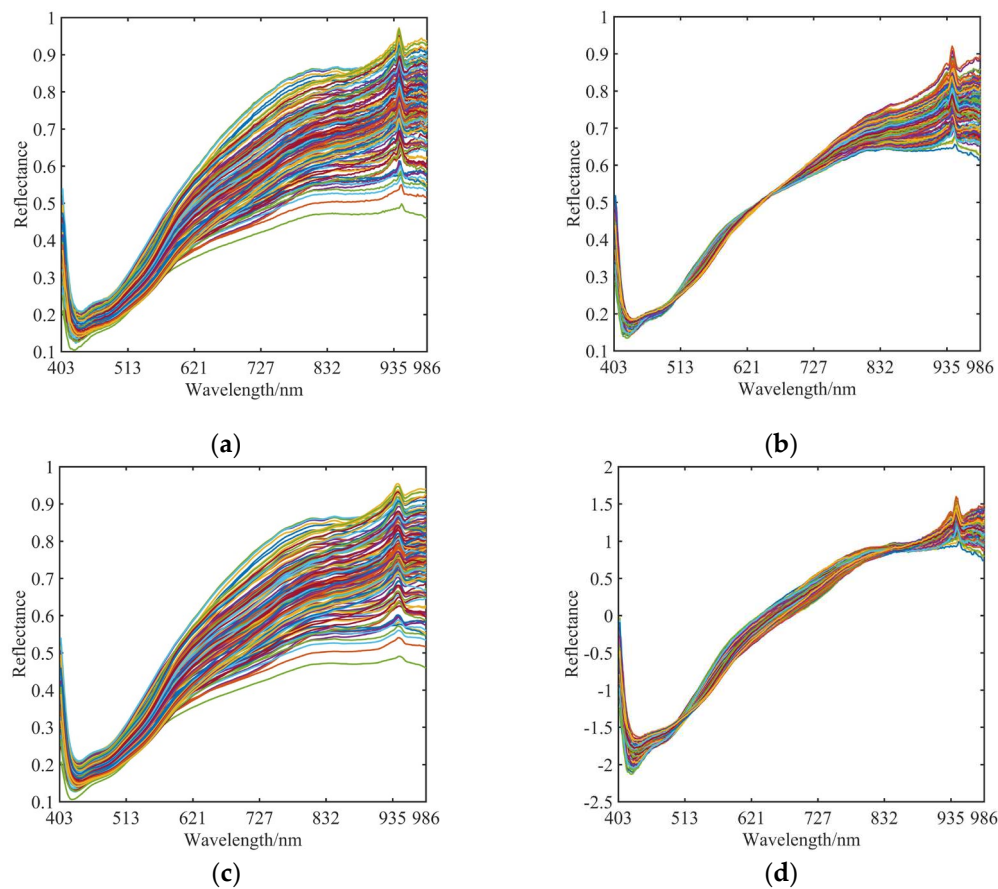


Figure 3. Raw spectra and various preprocessed spectra. (a) Raw spectra; (b) Spectra after MSC processing; (c) Spectra after SMOOTH processing; (d) Spectra after SNV processing.

3.2. Characteristic Wavelength Screening Results

The RF, VCPA, and VCPA-IRIV algorithms were used for the screening of the unique wavelengths based on various preprocessing of the spectral data. The stochastic selectivity of the variables and the effectiveness of the RF algorithm in selecting spectral feature variables result in a competitive selection mechanism for the variables that in turn ensures their validity [23]. The VCPA technique makes use of the exponentially declining function (EDF), model population analysis (MPA), and binary matrix sampling (BMS), which may effectively account for potential interactions between random variables [24]. A novel approach to variable selection that enables the selective omission of unimportant variables is the VCPA-IRIV method [25]. The number of iterations in the VCPA-IRIV algorithm is set at 50 for the EDF function and 1000 for the BMS function.

The number of wavelengths screened by the three methods after MSC pretreatment was 10, 11, and 43; the number of wavelengths screened by the three methods after SNV pretreatment was 10, 11, and 45; and the number of wavelengths screened by the three methods after smooth pretreatment was 10, 10, and 28. The results of the typical wavelength screening for various preprocessing steps are displayed in Figure 4a–c. It is clear that the predominant concentration of the distinctive wavelengths is in the range of 410 nm, 621–832 nm, and 940 nm. The average spectra of the 15 sampling regions are shown in Figure 4d, and it is clear that there is an obvious absorption peak at 940 nm. This peak is caused by the OH stretching vibration, which creates a slight water vapor absorption band and can be used to characterize the amount of soil organic matter. The primary matching bands of soil organic matter were found to be between 620–810 nm in the investigation by, and there was only a weak link between soil organic matter and the wavelength at 440 nm [26]. The typical wavelength distribution from this study is consistent.

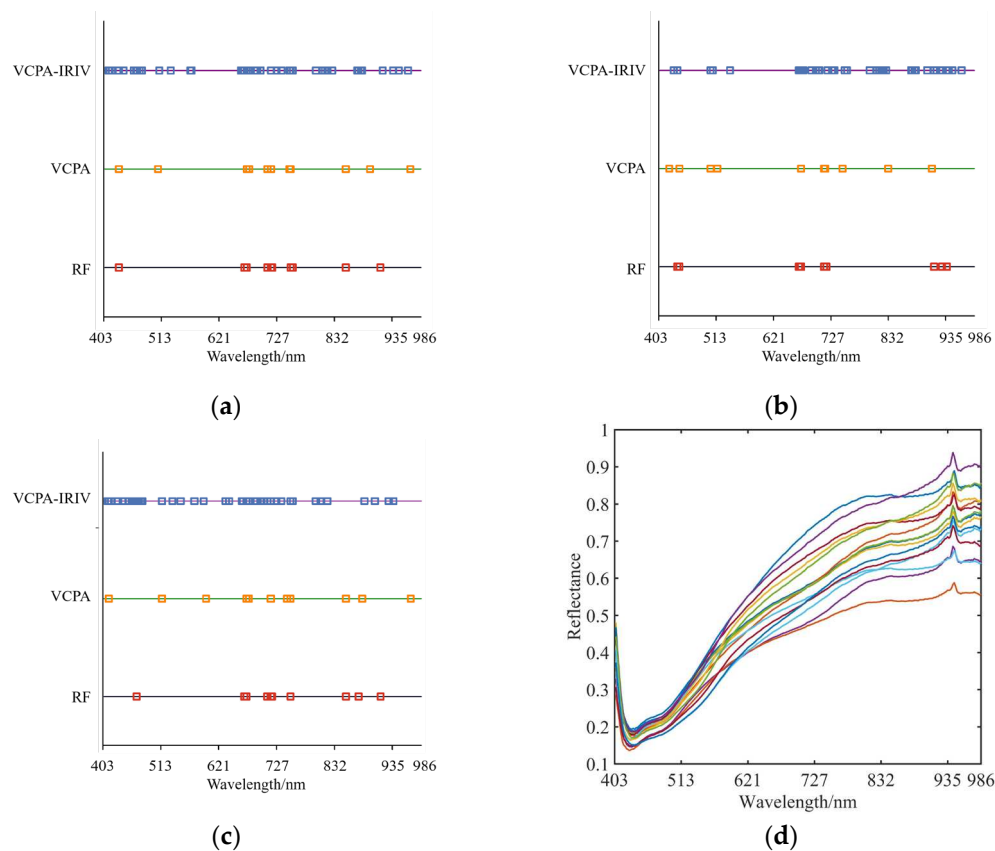


Figure 4. Characteristic wavelength screening results and average soil spectra for each region. (a) Distribution of characteristic wavelength screening after MSC preprocessing; (b) Distribution of characteristic wavelength screening after SMOOTH preprocessing; (c) Distribution of characteristic wavelength screening after SNV preprocessing; (d) Distribution of characteristic wavelength screening after SNV preprocessing.

3.3. Predictive Modeling of Single Spectral Data

Following feature wavelength extraction from the spectral data and PCA downscaling, PLSR and SVR prediction models for soil organic matter were created. Ten sets of main components were utilized as inputs to the models, and Table 1 displays the results for each model.

Table 1. Predictive modeling results for single spectral data.

Preprocessing Methods	Data Dimensionality Reduction	Models	PCs	Calibration Set		Prediction Set		RPD
				R ² _C	RMSEC	R ² _P	RMSEP	
MSC	RF	SVR	5	0.978	0.677	0.957	0.822	4.847
		PLSR	9	0.868	1.658	0.799	1.766	2.261
	VCPA	SVR	5	0.972	0.756	0.943	0.672	5.981
		PLSR	7	0.925	1.233	0.950	0.887	4.533
	VCPA-IRIV	SVR	7	0.994	0.350	0.973	0.693	6.119
		PLSR	10	0.926	1.234	0.892	1.205	3.083
SNV	RF	SVR	4	0.901	1.419	0.914	1.219	3.456
		PLSR	7	0.859	1.663	0.885	1.443	2.999
	VCPA	SVR	7	0.965	0.846	0.964	0.807	5.273
		PLSR	9	0.937	1.118	0.953	0.895	4.711
	VCPA-IRIV	SVR	6	0.984	0.565	0.960	0.768	5.071
		PLSR	9	0.901	1.401	0.909	1.200	3.376
SMOOTH	RF	SVR	4	0.960	0.895	0.935	1.079	3.940
		PLSR	8	0.827	1.854	0.761	2.100	2.076
	VCPA	SVR	5	0.926	1.213	0.909	1.337	3.345
		PLSR	9	0.946	1.017	0.940	1.082	4.168
	VCPA-IRIV	SVR	4	0.901	1.419	0.914	1.219	3.456
		PLSR	10	0.921	1.267	0.907	1.258	3.335

The preprocessing techniques MSC and SNV are generally superior to the SMOOTH procedure, according to a comparison of all models. In a side-by-side comparison of MSC and SNV, the SVR model under the MSC approach comes out on top overall, while some of the PLSR models under the SNV method perform better than MSC. Additionally, all SVR models outperformed PLSR models, showing how effective nonlinear spectral data modeling was in predicting soil organic matter. The best nonlinear model combination overall had R²_P, RSMEP, and RPD values of 0.973, 0.693, and 6.119, respectively, and was MSC+VCPA-IRIA+SVR. The best linear model combination had R²_P, RSMEP, and RPD values of 0.953, 0.895, and 4.711, respectively, and was SNV+VCPA+PLSR. Figure 5 depicts the connection between these two models' training and prediction sets. The findings show that it is possible to anticipate the amount of soil organic matter using hyperspectral technology.

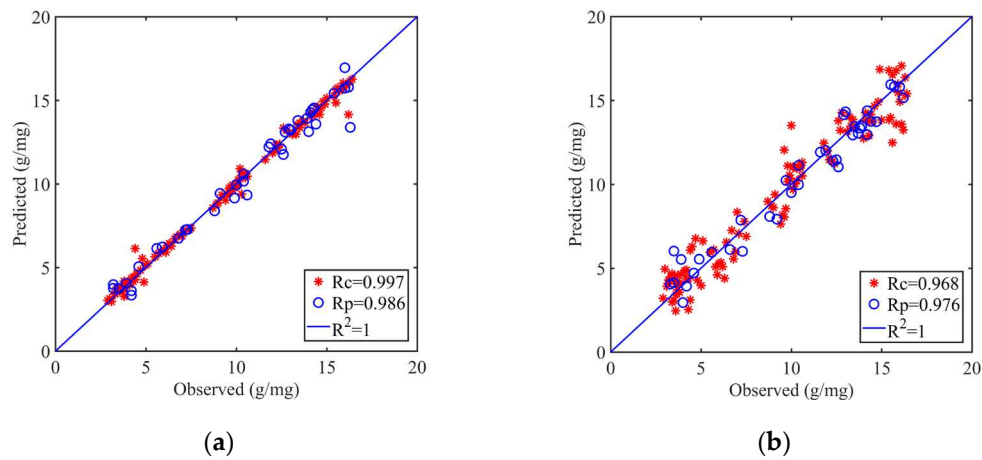


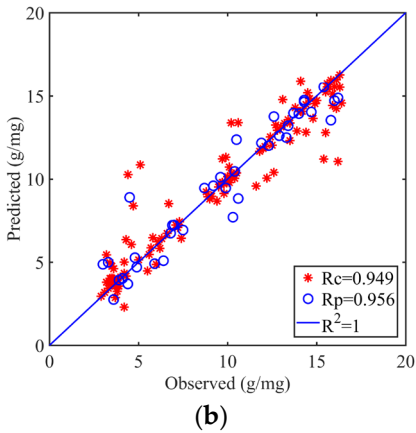
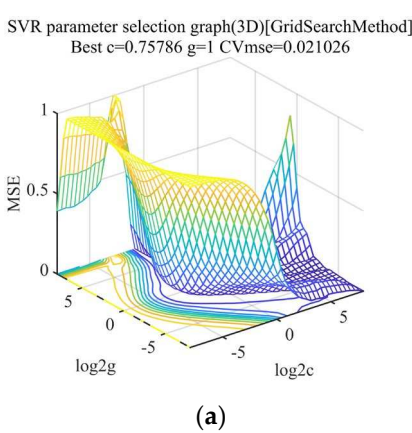
Figure 5. Comparison of model training set and prediction set under single spectral data. (a) MSC+VCPA-IRIA+SVR; (b) SNV+VCPA+PLSR.

3.4. Predictive Modeling of Fusion Data

SVR and PLSR models were created based on the fused data of spectral and image characteristics, respectively, to investigate the predictive modeling impact of the fused data. Table 2 displays the accuracy of each model's predictions for soil organic matter. It is evident from the data in the table that the performance of the models with fused data is always better than the performance of the models with single spectral data. With an improvement in R^2_c from 0.901 to 0.980, R^2_p from 0.914 to 0.962, and RPD from 3.456 to 4.563, SMOOTH+VCPA-IRIV+SVR is the combination of modeling approaches that has the greatest change in impact. In Figure 6b,d, the training set-prediction set connections using spectral data and fused data are compared. Figure 6a,c show how spectral data and fused data are optimized for parameters using the SMOOTH+VCPA-IRIV+SVR combination.

Table 2. Performance results of each model for fusion data.

Preprocessing Methods	Data Dimensionality Reduction	Models	PCs	Calibration Set		Prediction Set		RPD
				R2C	RMSEC	R2P	RMSEP	
MSC	RF	SVR	9	0.990	0.622	0.959	0.948	4.938
		PLSR	9	0.888	1.442	0.889	1.547	3.051
	VCPA	SVR	9	0.983	0.579	0.976	0.660	6.459
		PLSR	10	0.951	0.987	0.950	0.941	4.534
	VCPA-IRIV	SVR	10	0.995	0.312	0.986	0.558	8.155
		PLSR	10	0.947	1.020	0.921	1.252	3.600
SNV	RF	SVR	9	0.989	0.480	0.962	0.851	5.008
		PLSR	10	0.912	1.327	0.923	1.191	3.650
	VCPA	SVR	9	0.995	0.323	0.970	0.761	5.854
		PLSR	10	0.954	0.956	0.965	0.818	5.448
	VCPA-IRIV	SVR	9	0.992	0.406	0.982	0.639	6.957
		PLSR	10	0.903	1.373	0.925	1.233	3.704
SMOOTH	RF	SVR	8	0.981	0.623	0.950	0.942	4.530
		PLSR	10	0.894	1.442	0.904	1.306	3.267
	VCPA	SVR	8	0.976	0.710	0.942	0.972	4.051
		PLSR	10	0.965	0.831	0.941	0.988	4.156
	VCPA-IRIV	SVR	8	0.980	0.639	0.962	0.951	4.563
		PLSR	10	0.940	1.095	0.925	1.132	3.693



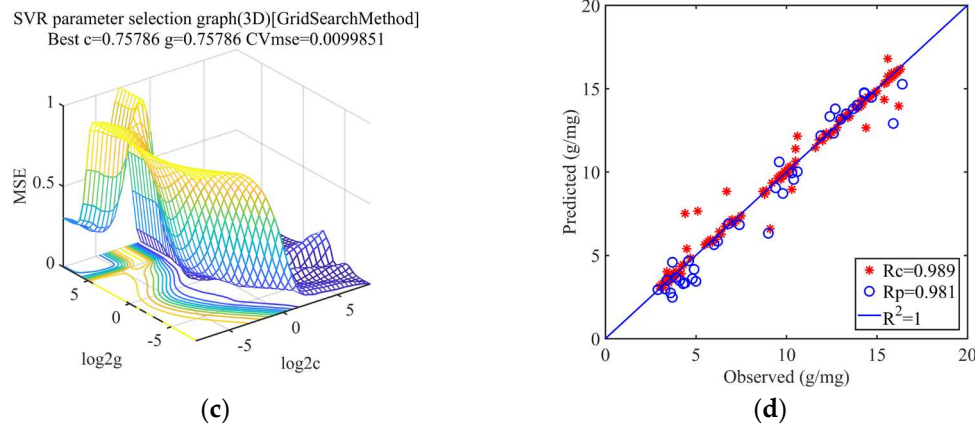


Figure 6. Comparison of non-fusion and fusion data models under SNV+VCPA-IRIV approach. (a) Parameter optimization process for a single spectral data model; (b) Comparison of training and prediction sets of single spectral data models; (c) Parameter optimization process for fusion data models; (d) Comparison of training and prediction sets of fusion data models.

Comparing the modeling performance of the combined data reveals that the best linear model combination is SNV+VCPA-PLSR, with R^2_c , R^2_p , and RPD values of 0.954, 0.965, and 5.448, respectively, while the best nonlinear model combination is MSC+VCPA-IRIV+SVR, with R^2_c , R^2_p , and RPD values of 0.995, 0.986, and 8.155, respectively. This is in line with the modeling findings from the spectral data, showing that these two combinations can provide the most accurate predictions of soil organic matter and are thus likely to be used. The model performance results for both combinations are displayed in Figure 7a,b.

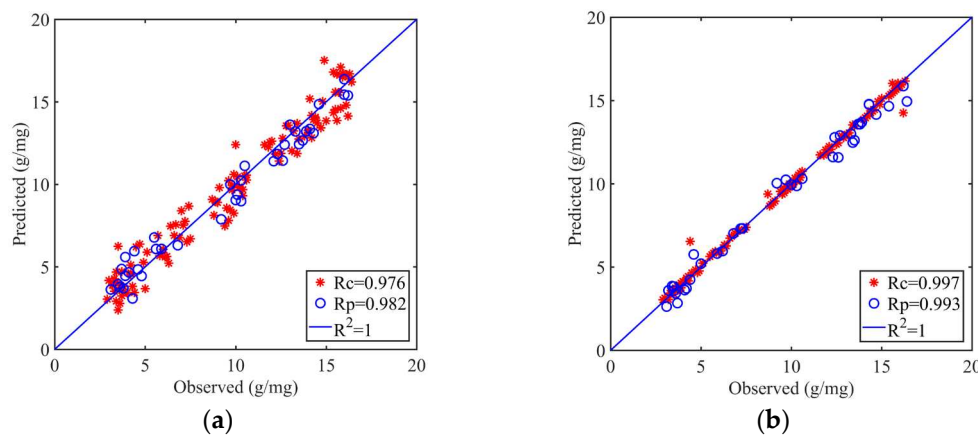


Figure 7. Effectiveness of linear and nonlinear optimal models under fusion data. (a) SNV+VCPA-PLSR; (b) MSC+VCPA-IRIV+SVR.

4. Conclusions

In this work, the SVR and PLSR models were built using the intermediate layer fusion approach to combine the soil hyperspectral and picture texture characteristic data. The findings demonstrate that the performance of the fused data models is adequate and that both have outperformed the single spectral data models in the same manner. Out of the three, the nonlinear model performs better at estimating the amount of soil organic matter. The most accurate SVR model is the one built using the MSC and VCPA-IRIV algorithms. It had an R^2_c of 0.995, an R^2_p of 0.986, and an RPD of 8.155. The results of this study demonstrate the feasibility of soil organic matter content prediction by fusion data, which improves the accuracy of the prediction model and leads to better modeling results compared to single data. It is beneficial to promote accurate control of soil organic matter content in tea plantations. Targeted irrigation and sowing follow from this.

Supplementary Materials: The following supporting information can be downloaded at the website of this paper posted on Preprints.org, Figure S1: title; Table S1: title; Video S1: title.

Author Contributions: Conceptualization, C.D.; methodology, H.Z. and Z.L.; software, C.D.; validation, C.Y. and X.Z.; formal analysis, H.Z.; investigation, H.Z.; resources, C.D. and X.Z.; data curation, H.Z.; writing—original draft preparation, H.Z.; writing—review and editing, C.D.; visualization, H.Z.; supervision, M.L.; project administration, C.D.; funding acquisition, X.Z. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Acknowledgments: This study was supported by The Key R&D Projects in Zhejiang Province (2022C02010, 2023C02043), the Research start-up funds-TRI-SAAS (CXGC2023F18, CXGC2023G33) and the Key Projects of Science and Technology Cooperation in Jiangxi Province (20212BDH80025).

Conflicts of Interest: No conflict of interest exists in the submission of this manuscript, and manuscript is approved by all authors for publication. All the authors listed have approved the manuscript that is enclosed.

References

- Chen, J., Wu, S., Dong, F., Li, J., Zeng, L., Tang, J., and Gu, D. (2021). Mechanism Underlying the Shading-Induced Chlorophyll Accumulation in Tea Leaves. *Frontiers in Plant Science* 12.
- Dong, C., Ye, Y., Yang, C., An, T., Jiang, Y., Ye, Y., Li, Y., and Yang, Y. (2021). Rapid detection of catechins during black tea fermentation based on electrical properties and chemometrics. *Food Bioscience* 40.
- Rahimi-Ajdadi, F., Abbaspour-Gilandeh, Y., Mollazade, K., and Hasanzadeh, R. P. R. (2018). Development of a novel machine vision procedure for rapid and non-contact measurement of soil moisture content. *Measurement* 121, 179-189.
- Hoffland, E., Kuyper, T. W., Comans, R. N. J., and Creamer, R. E. (2020). Eco-functionality of organic matter in soils. *Plant and Soil* 455, 1-22.
- İnik, O., İnik, Ö., Öztas, T., Demir, Y., and Yüksel, A. (2023). Prediction of Soil Organic Matter with Deep Learning. *Arabian Journal for Science and Engineering* 48, 10227-10247.
- Zwiazek, J. J., Kyaw, T. Y., Siegert, C. M., Dash, P., Poudel, K. P., Pitts, J. J., and Renninger, H. J. (2022). Using hyperspectral leaf reflectance to estimate photosynthetic capacity and nitrogen content across eastern cottonwood and hybrid poplar taxa. *Plos One* 17.
- Yu, F., Bai, J., Jin, Z., Zhang, H., Yang, J., and Xu, T. (2023). Estimating the rice nitrogen nutrition index based on hyperspectral transform technology. *Frontiers in Plant Science* 14.
- Chen, D., Wu, P., Wang, K., Wang, S., Ji, X., Shen, Q., Yu, Y., Qiu, X., Xu, X., Liu, Y., and Tang, G. (2022). Combining computer vision score and conventional meat quality traits to estimate the intramuscular fat content using machine learning in pigs. *Meat Science* 185.
- Chen, F., Xu, J., Wei, Y., and Sun, J. (2019). Establishing an eyeball-weight relationship for *Litopenaeus vannamei* using machine vision technology. *Aquacultural Engineering* 87.
- Yang, C., Feng, M., Song, L., Jing, B., Xie, Y., Wang, C., Qin, M., Yang, W., Xiao, L., Sun, J., Zhang, M., Song, X., and Kubar, M. S. (2022). Hyperspectral monitoring of soil urease activity under different water regulation. *Plant and Soil* 477, 779-792.
- Rahimi-Ajdadi, F., Abbaspour-Gilandeh, Y., Mollazade, K., and Hasanzadeh, R. P. R. (2018). Development of a novel machine vision procedure for rapid and non-contact measurement of soil moisture content. *Measurement* 121, 179-189.
- An, T., Huang, W., Tian, X., Fan, S., Duan, D., Dong, C., Zhao, C., and Li, G. (2022). Hyperspectral imaging technology coupled with human sensory information to evaluate the fermentation degree of black tea. *Sensors and Actuators B: Chemical* 366.
- Yin, Y., Li, J., Ling, C., Zhang, S., Liu, C., Sun, X., and Wu, J. (2023). Fusing spectral and image information for characterization of black tea grade based on hyperspectral technology. *Lwt* 185.
- Dong, C., Liang, G., An, T., Wang, J., Zhu, H. (2018). Near-infrared spectroscopy detection model for sensory quality and chemical constituents of black tea, *Transactions of the Chinese Society of, Agricultural Engineering* 34, 306-313.
- Sun, J., Yang, W., Feng, M., Liu, Q., and Kubar, M. S. (2020). An efficient variable selection method based on random frog for the multivariate calibration of NIR spectra. *RSC Advances* 10, 16245-16253.
- Jiang, H., Xu, W., Ding, Y., and Chen, Q. (2020). Quantitative analysis of yeast fermentation process using Raman spectroscopy: Comparison of CARS and VCPA for variable selection. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* 228.
- Han, M., Wang, X., Xu, Y., Cui, Y., Wang, L., Lv, D., and Cui, L. (2021). Variable selection for the determination of the soluble solid content of potatoes with surface impurities in the visible/near-infrared range. *Biosystems Engineering* 209, 170-179.

18. Cebi, N., Yilmaz, M. T., and Sagdic, O. (2017). A rapid ATR-FTIR spectroscopic method for detection of sibutramine adulteration in tea and coffee based on hierarchical cluster and principal component analyses. *Food Chemistry* 229, 517-526.
19. Cruz-Tirado, J. P., Oliveira, M., de Jesus Filho, M., Godoy, H. T., Amigo, J. M., and Barbin, D. F. (2021). Shelf life estimation and kinetic degradation modeling of chia seeds (*Salvia hispanica*) using principal component analysis based on NIR-hyperspectral imaging. *Food Control* 123.
20. Liu, Z., Zhang, R., Yang, C., Hu, B., Luo, X., Li, Y., Dong, C. (2022). Research on moisture content detection method during green tea processing based on machine vision and near-infrared spectroscopy technology. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* 272, 1386-1425.
21. Ren, G., Wang, Y., Ning, J., and Zhang, Z. (2020). Highly identification of keemun black tea rank based on cognitive spectroscopy: Near infrared spectroscopy combined with feature variable selection. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* 230.
22. Dong, C., Ye, Y., Yang, C., An, T., Jiang, Y., Ye, Y., Li, Y., and Yang, Y. (2021). Rapid detection of catechins during black tea fermentation based on electrical properties and chemometrics. *Food Bioscience* 40.
23. Ouyang, Q., Wang, L., Zareef, M., Chen, Q., Guo, Z., and Li, H. (2020). A feasibility of nondestructive rapid detection of total volatile basic nitrogen content in frozen pork based on portable near-infrared spectroscopy. *Microchemical Journal* 157.
24. Jiang, H., He, Y., Xu, W., and Chen, Q. (2021). Quantitative Detection of Acid Value During Edible Oil Storage by Raman Spectroscopy: Comparison of the Optimization Effects of BOSS and VCPA Algorithms on the Characteristic Raman Spectra of Edible Oils. *Food Analytical Methods* 14, 1826-1835.
25. Yang, C., Zhao, Y., An, T., Liu, Z., Jiang, Y., Li, Y., and Dong, C. (2021). Quantitative prediction and visualization of key physical and chemical components in black tea fermentation using hyperspectral imaging. *Lwt* 141.
26. Liu, H., Zhang, Y., and Zhang, B. (2008). Novel hyperspectral reflectance models for estimating black-soil organic matter in Northeast China. *Environmental Monitoring and Assessment* 154, 147-154.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.