

Article

A Standard-Free Calibration Transfer Strategy for a Discrimination Model of Apple Origins Based on Near-Infrared Spectroscopy

Lisha Li, Bin Li, Xiaogang Jiang and Yande Liu *

School of Mechatronics and Vehicle Engineering, East China Jiaotong University, Nanchang 330013, China; lilisha19961030@163.com (L.L.); li_yang202110@163.com (B.L.); jxg_ecjtu@163.com (X.J.)

* Correspondence: 2607@ecjtu.edu.cn

Abstract: The nondestructive discrimination model based on near-infrared is usually established by detected spectra and chemometric methods. However, the inherent differences between instruments prevent the model from being used universally, and calibration transfer is often used to solve these problems. Standard-sample calibration transfer requires additional standard samples to build a mathematical mapping between instruments. Thus, standard-free calibration transfer is a research hotspot in this field. Based on near-infrared spectroscopy (NIRS), the new combined strategy of wavelength selection and standard-free calibration transfer was proposed to transfer the model between two portable near-infrared spectrometers. Three transfer learning (TL) algorithms—transferred component analysis (TCA), balanced distribution adaptation (BDA), and manifold embedded distribution alignment (MEDA)—were applied to achieve standard-free calibration transfer. Moreover, this paper presents a relative error analysis (REA) method to select wavelength. To select the optimal model, the parameters of accuracy, precision, and recall were examined to evaluate the discriminatory capacities of each model. The findings show that the MEDA-REA model is capable of higher prediction accuracy (accuracy = 94.54%) than the other transferring models (TCA, BDA, MEDA, TCA-REA, and BDA-REA), and it is demonstrated that the new strategy has good transmission performance. Moreover, REA shows the potential to filter wavebands for calibration transfer and simplify the transferable model.

Keywords: standard-free calibration transfer; near-infrared spectroscopy; wavelength selection



Citation: Li, L.; Li, B.; Jiang, X.; Liu, Y. A Standard-Free Calibration Transfer Strategy for a Discrimination Model of Apple Origins Based on Near-Infrared Spectroscopy. *Agriculture* **2022**, *12*, 366. <https://doi.org/10.3390/agriculture12030366>

Academic Editor: Maciej Zaborowicz

Received: 26 January 2022

Accepted: 28 February 2022

Published: 4 March 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

At present, China is the global major producer and exporter of apples. To ensure the quality of exported apples, quality testing is essential, so near-infrared non-destructive testing technology needs to be adopted [1,2]. There are many Fuji-producing areas in China, and famous amongst these are Aksu in Xinjiang and Yantai in Shandong, followed by Shaanxi, Sichuan, etc. Red Fuji from different origins have different tastes and internal qualities, but as it is usually difficult to distinguish them based on external morphology, there is a tendency for defective products to be substituted for the quality products. Near-infrared spectroscopy (NIRS) is a commonly used technique to determine the origin [3], variety [4], and optimum harvest time [5] for agricultural products, etc.

However, NIRS has the limitation of “one model for one instrument”, even if these instruments are of the same type. Inherent differences between spectrometers are the main cause of this limitation, such as differences in the durability of hardware, as well as differences in illumination intensities, in the instruments’ environments, and in temperature or humidity [6,7]. Therefore, an instrument can be used in a new environment only by calibrating the model, and this process is called calibration transfer [8].

Generally, calibration transfer methods are split into two categories: standard-sample calibration transfer and standard-free calibration transfer. Standard-sample calibration

transfer is a mathematical method according to which the adaptability of the near-infrared model is improved by computing the mathematical relationship between the spectrum collected by the source and the target instruments [9]. There are commonly used methods of standard-sample calibration transfer, such as the slope/bias (S/B) algorithm [10], Shenk's algorithm [11], the slope/bias correction (SBC) algorithm [12], the direct standardization (DS) algorithm [13], the piece-wise direct standardization (PDS) algorithm [14], the transfer by orthogonal projection (TOP) algorithm [15], the spectral space transformation (SST) algorithm [16], the canonical correlation analysis (CCA) algorithm [17], and the extreme learning machine auto-encoder (TAEM) method [18], etc. Dong et al. [19] studied the calibration transfer between near-infrared models of White Leghorn eggs and Bantam eggs by DS and SBC algorithms, respectively; the best prediction results for albumen pH with an R_p of 0.908 and an RMSEP of 0.133 was obtained by Poerio et al. [9], who demonstrated the effectiveness of TOP in the field of calibration transfer of NIRS, and showed a significant correction effect in the apparent baseline by testing on three near-infrared data sets. Li et al. [20] studied three calibration transfer methods (PDS, SST, and CCA) between two developed portable Vis/NIR devices to establish a robust model to predict the soluble solids content (SSC) of apples. The results indicated that the PDS method had the best calibration performance ($R_p = 0.926$, RMSEP = 0.778).

Even though standard-sample calibration transfer can realize an excellent correction effect with the model, additional standard samples are still needed in the transfer process, which is the method's major flaw. The rapid growth of standard-free calibration transfer is a response to this defect [21]. For this method, standard samples are not needed; instead, waveband screening, spectral signal pre-processing, or other patterns are used to calibrate the deviation between instruments. For example, Zhang et al. [22] and Zheng et al. [23] developed the stability competitive adaptive reweighted sampling (SCARS) algorithm and the double competitive adaptive reweighted sampling (Double CARS) algorithm, respectively—two algorithms that belong to standard-free calibration transfer and which have been found to have a good correction performance. The screening wavelengths with consistent and stable signals (SWCSS) method proposed by Zhang et al. [24] could transfer a model between instruments, and the study in which the calibrated model was predicted with $R_p = 0.959$ and RMSEP = 0.236 achieved a better prediction than before. Xu et al. [25] developed the correlation analysis-based wavelength selection (CAWS) method, which takes a Pearson correlation coefficient as the screening condition, and the transferred model obtained the lowest RMSEP (0.069). In recent years, same transfer learning (TL) algorithms, such as TrAdaBoost [26], transfer component analysis (TCA) [27], and easy transfer learning (EasyTL) [28], were led into the standard-free calibration transfer field. Yu et al. [29] used the TrAdaBoost to simulate verification of ten datasets (fuels and foods) from different instruments, and the prediction accuracy of the cetane number of the fuels with $R^2_p = 0.993$ and RMSEP = 0.438 was significantly improved. Mishra et al. [30] compared dynamic orthogonalization projection (DOP) with TCA in the study of prediction model of fruit's interior quality, and the experimental results showed that the number of latent variables of TCA was lower, indicating that TCA could achieve the underlying subspaces more efficiently than DOP. Zhao et al. [31] used EasyTL to realize calibration transfer from a desktop HSI system to a near-infrared portable spectrometer, and the accuracy of the discrimination model of pollution degree in soil was 69%.

Introducing a TL algorithm into the calibration transfer field is a current research hotspot [21]. Wang et al. successively developed two new TL algorithms named manifold balanced distribution adaptation (BDA) [32] and embedded distribution alignment (MEDA) [33]. The problem of quantitatively estimating the marginal distribution and conditional distribution in TL can be systematically solved by MEDA. The BDA is used to deal with the issue of class imbalance in the TL field and to add weights to each category. To date, the application of MEDA and BDA in the calibration transfer field has not been investigated. Therefore, this research implemented calibration transfer between two near-infrared portable spectrometers, and three TL methods (MEDA, TCA, and BDA) were applied

to calibrate a discrimination model for determining the origin of apples (Aksu Xinjiang, Yantai Shandong, Panzhihua Sichuan, Luochuan Shaanxi) in this work. The wavelength selection process is often lacking in the calibration transfer process because the traditional wavelength selection algorithms are not suitable for the calibration transfer process [34]. Thus, this paper proposed a new wavelength selection method (relative error analysis, REA) for the calibration transfer process. This method was used in combination with TCA, BDA, and MEDA algorithms, respectively, to achieve the calibration of the model.

2. Materials and Methods

2.1. Samples and Pretreatment

Apple samples (Red Fuji) were collected from four regions of China, denoted as Fuji-1, Fuji-2, Fuji-3, and Fuji-4, and stored in a preservation box at 5 °C. The sampling sites were located in the area of Aksu city (80°15'54" E, 41°10'15" N), Panzhihua city (101°42'58" E, 26°34'50" N), Luochuan county (109°25'58" E, 35°45'39" N), and Yantai city (121°23'29" E, 37°32'21" N), respectively. Altogether, 949 samples were selected with a uniform shape (280 samples of Fuji-1, 244 of Fuji-2, 225 of Fuji-3, and 200 of Fuji-4), a single fruit weight of 200 ± 10 g, and intact epidermis. Four circle marks were uniformly made at the equatorial region of each sample. To avoid temperature effects, all samples were placed in the laboratory for 0.5 h at an ambient temperature of 21–23 °C before spectra acquisition. Four Fuji samples are represented in Figure 1.



Figure 1. Fuji apples from four regions.

2.2. Spectra Acquisition

The spectra of the samples were acquired using two F750 handheld near-infrared spectrometers (Felix Instruments, R&D and manufacturing: Camas, WA, USA; Agent: Zhejiang tuopuyunnong Technology Co., Ltd., Zhejiang, China), with a 32 W halogen light and a near-infrared spectrometer module mounted in each device (MMS1, Zeiss, Jena, Germany, 729–975 nm and a resolution of 3 nm). The source instrument was marked as S1 and the target instrument was marked as S2. The physical map of the experimental device is shown in Figure 2a. Each measured spectrum includes 83 wavelength points. For the collection of spectra, the sample's mark was put on the equipment's detection port before pressing the acquisition button. The spectrometer automatically scanned seven times and took the average as the output spectrum. Four spectra were obtained for each sample, and their average was taken as the final spectrum signal of the sample. The average spectra of four kinds of apples are shown in Figure 2b. Although most of the average spectral bands of these four kinds overlap, there are apparent differences in absorbance at 740 nm (Figure 2b-A), 932 nm (Figure 2b-B), and 963 nm (Figure 2b-C) that have been found from the figure. To reflect the standard deviation of each waveband, Figure 2c draws the standard deviation curve of the average spectrum of four types of samples, and this figure corresponds to the wavebands of the four average spectra in Figure 2b. It can be seen that all the standard deviations of the marked bands in Figure 2b are relatively large in Figure 2c. This means that these wavelengths with deviations can be used to identify the origin of apples.

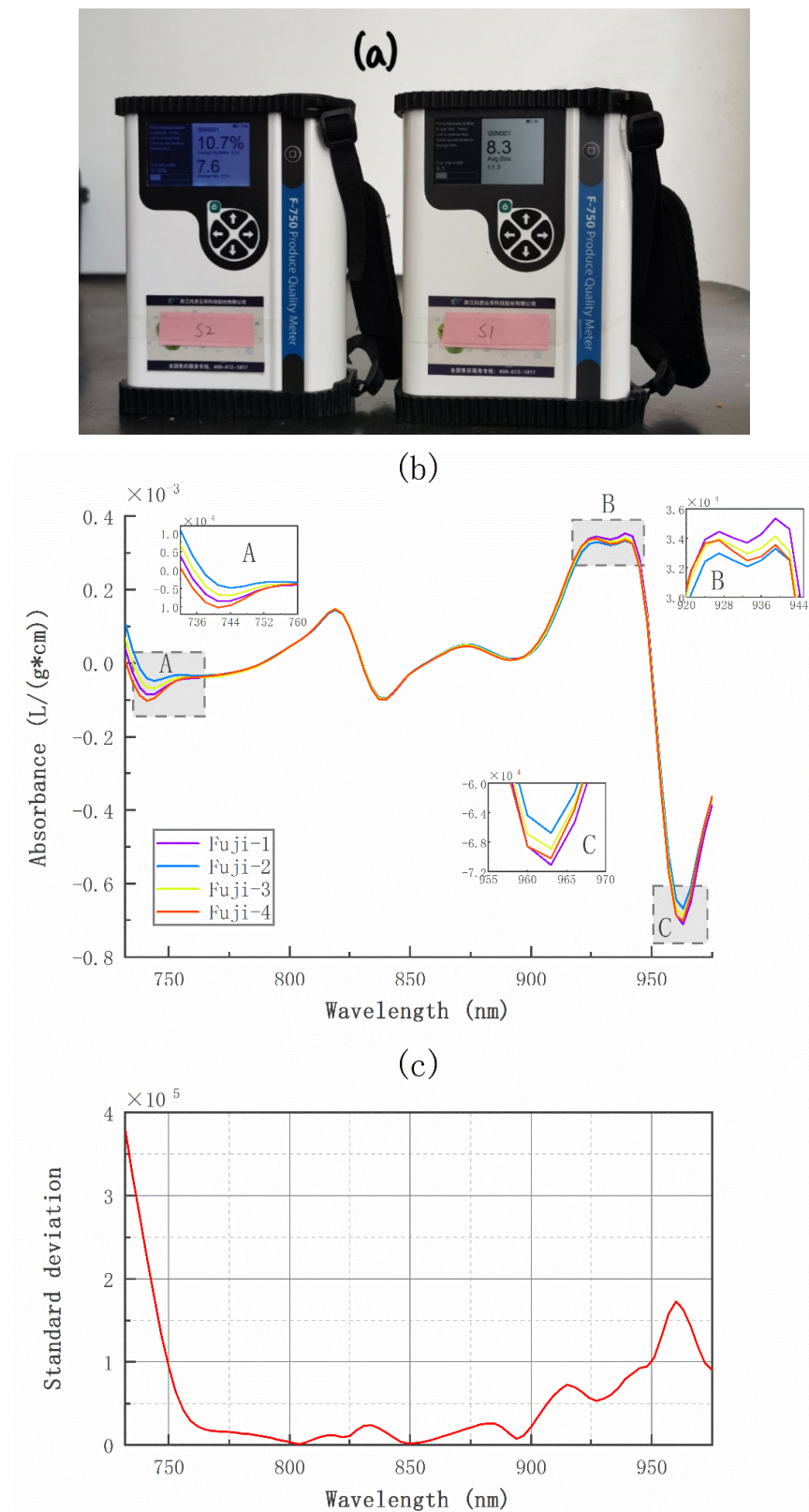


Figure 2. Spectral acquisition of Fuji apples. (a) Experimental instrument. (b) Averaged NIR absorbance spectrum of apples acquired from S1. (c) Standard deviation of average spectra of four categories of Fuji apples.

2.3. Division Training Set and Testing Set

The joint x–y distance (SPXY) algorithm was used to divide the training and testing sets in this study. It is developed based on the Kennard–Stone (KS) algorithm. The x and y variables are taken into account by SPXY when calculating the Euclidean metric between samples [35]. In this work, 238 samples as a testing set and 711 samples as a training set were selected by SPXY in a ratio of 3:1.

2.4. Relative Error Analysis

The REA method was first proposed for application in standard-sample calibration transfer [34]. However, this study required REA to combine with the standard-free calibration transfer method. Therefore, a slight alteration to the former made REA available for standard-free calibration transfer. Figure 3 shows the wavelength selection process of improved REA for standard-free calibration transfer.

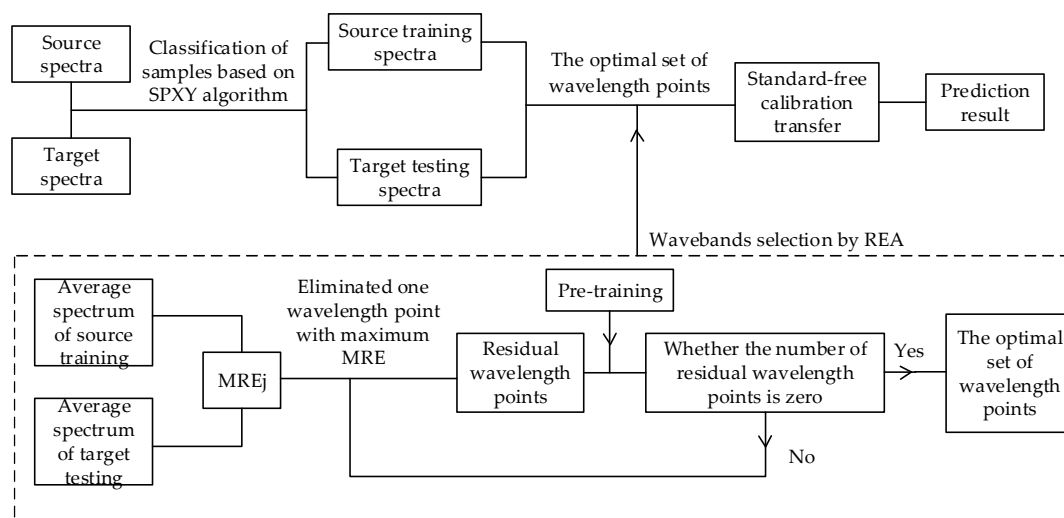


Figure 3. The flowchart of REA's operating principle.

Firstly, the average spectra of source training and target testing spectra are calculated and denoted as X_s and X_t , respectively. Secondly, the mean absolute error (MAE) between X_s and X_t is computed using Equation (1), where j is the number of spectral wavelength points, X_{sj} is the average source spectrum at the j -th wavelength point, and X_{tj} is the average target spectrum at the j -th wavelength point. The MAE value of each wavelength point has been substituted into Equation (2) to calculate the mean relative error (MRE).

$$MAE_j = X_{sj} - X_{tj} \quad (1)$$

$$MRE_j = \left| \frac{MAE_j}{X_{sj}} \times 100\% \right| \quad (2)$$

Finally, wavebands are sorted according to the MRE value, and these wavebands are removed, largest to smallest. When one waveband is eliminated, the accuracy of the built model by remaining bands is calculated one time until all wavebands are rejected. Then, setting the maximum accuracy as the objective to be searched, the optimal set of wavelength points for the full spectrum is derived.

2.5. Model Construction, Model Evaluation, and Software

In terms of multivariate modeling methods, the traditional supporting vector machine (SVM) algorithm can be extensively used for building a near-infrared discrimination model. Therefore, SVM was employed as a fundamental modeling algorithm in this research. To

select a robust model and obtain the best prediction result, the application of evaluation indexes is indispensable. These evaluation indexes were accuracy (Equation (3)), precision (Equation (4)), and recall (Equation (5)), which can be calculated from the confusion matrix (Figure 4):

$$\text{Accuracy} = \frac{TP + TN}{TP + FN + FP + TN} \quad (3)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (4)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (5)$$

where TP, TN, FP, and FN represent the number of true positives, true negatives, false positives, and false negatives, respectively (Figure 4). Accuracy is the proportion of correctly classified samples relative to the total number. Although accuracy is commonly used, it cannot satisfy all the demands of the assignment. Hence, precision and recall are introduced to evaluate the model comprehensively. Precision and recall are contradictory. Generally, the recall rate is often low when the precision rate is high [36]. The reasonable way to judge a model's performance is to consider the extent to which the "double high" of precision and recall is achieved [37].

True class	Predicted class	
	Positive	Negative
Positive	TP	FN
Negative	FP	TN

Figure 4. Confusion matrix of classified results (taking a binary classification task as an example).

3. Results

3.1. Traditional Discrimination Model before Calibration

3.1.1. SVM Model Constructed for S1

The SVM model was established based on 711 training samples and 238 testing samples to examine its predictive performance. All of these spectral signals are measured by S1. The confusion matrix in Figure 5a shows the prediction result of the discrimination model, and its accuracy can reach 96.22%. Recall can intuitively reflect the probability of being detected in a particular class, while precision can intuitively reflect the correctness of a detected class. The precision and recall for the testing samples from four habitats have been illustrated in Figure 5b. These precision and recall values are comparatively high, and the SVM model has optimal predictive performance for Fuji-4.

3.1.2. Transfer the SVM Model from S1 to S2

However, the discrimination model constructed by S1 cannot be directly applied to another spectrometer, even if they are of the same type. This is a common deficiency among portable spectrometers. The data in Figure 6a support this conclusion. When the model established by S1 is transferred to S2, the accuracy decreases from 96.22% to 84.45%. In particular, the model's predictive performance for Fuji-3 declined sharply. Many Fuji-3 samples are misclassified as Fuji-1 or Fuji-2. This phenomenon can also be observed in Figure 6b, in which it can be seen that precision for Fuji-1 dipped from 92.41% to 79.31%,

while precision for Fuji-2 fell from 95% to 68.42%. Therefore, to avoid generating additional human and material resource costs by rebuilding the model, it is essential to calibrate the model established by S1 to enhance its universality and make it applicable to S2.

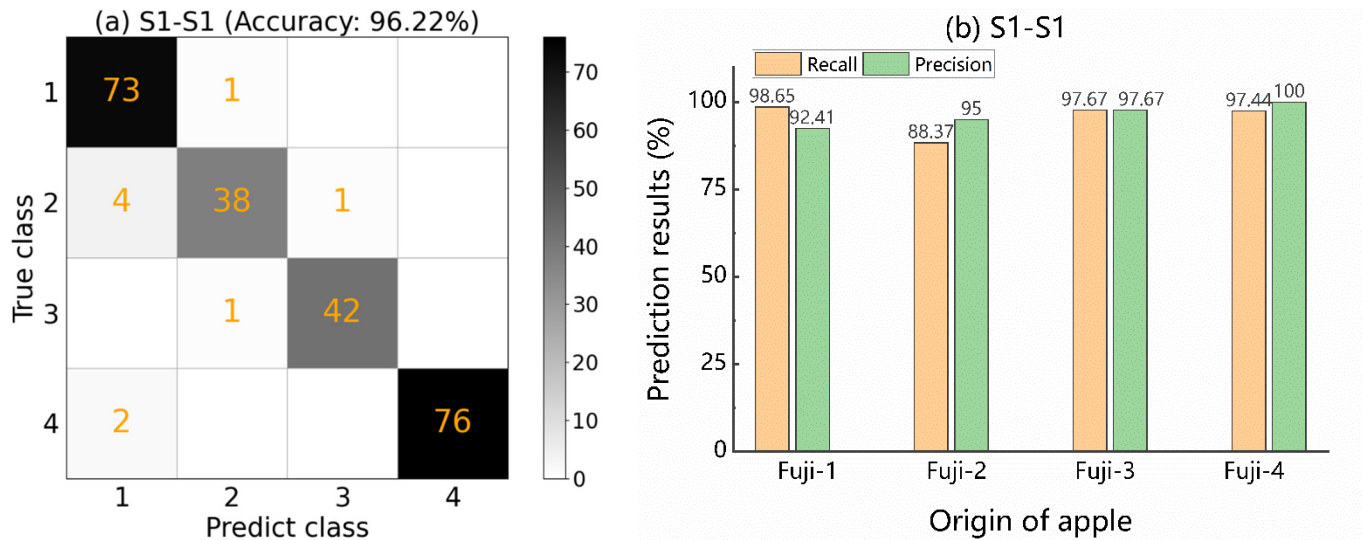


Figure 5. Prediction results of the SVM model established by the S1 instrument. (a) Confusion matrix. (b) Bar chart of precision and recall of various samples.

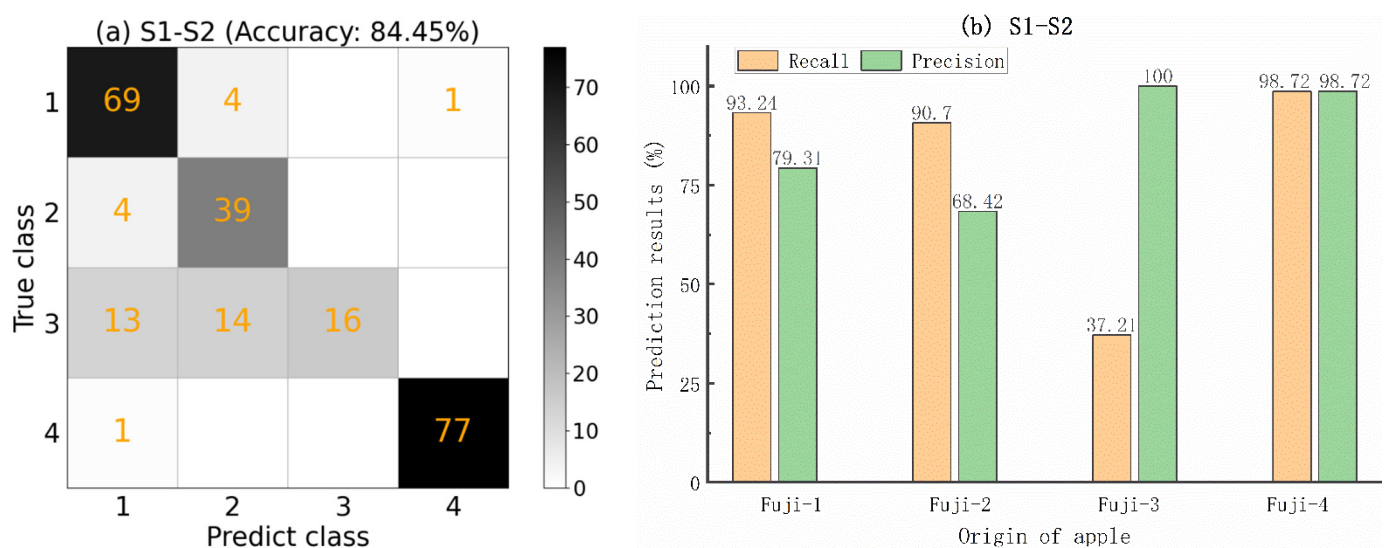


Figure 6. Prediction results of the transferred SVM model from S1 to S2. (a) Confusion matrix. (b) Bar chart of precision and recall of various samples.

3.2. Calibration Transfer from S1 to S2

This study mainly used three TL methods for standard-free calibration transfer: TCA, BDA, and MEDA. TCA is a relatively primitive TL method. It has been applied to the calibration of the near-infrared model and proved its good performance. BDA and MEDA algorithms are extended and evolved based on TCA, and they have not been used in this field. This research compared them with TCA to explore their calibration performance to find a more suitable TL algorithm. Figure 7 shows the confusion matrix of the prediction results of the calibrated models. The data analysis spotted that although TCA has a certain degree of optimization effect on the model (accuracy = 86.13%) and the model's predictive performance for Fuji-3 has been improved, it is not as good as BDA and MEDA. MEDA

has the best calibration performance for the transferred model, and the model's accuracy can reach 92.02%. The main finding in analyzing these confusion matrixes is that the misjudgment rate for Fuji-3 is significantly reduced (Figure 7c).

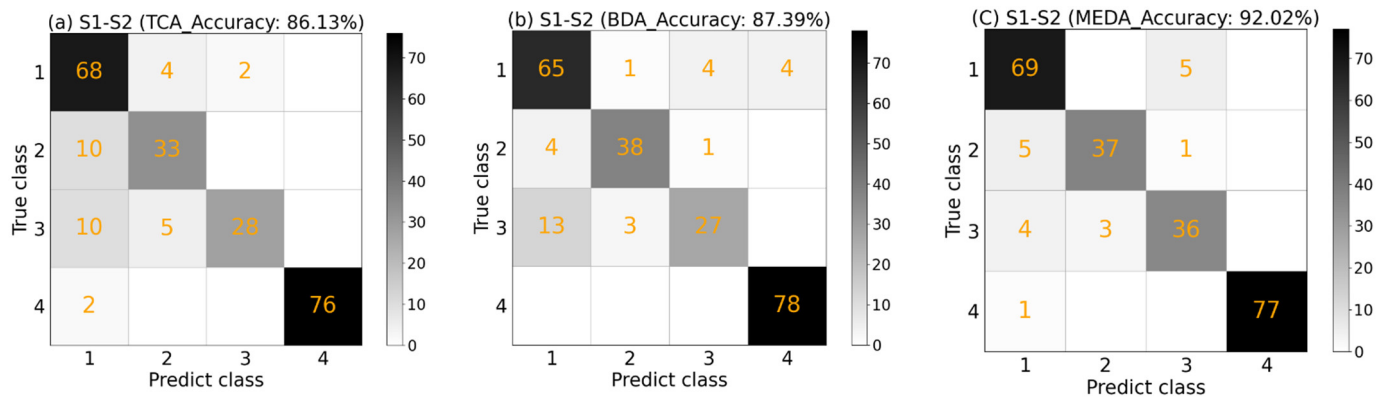


Figure 7. Confusion matrixes of the transferred model. (a) TCA model. (b) BDA model. (c) MEDA model.

Table 1 presents the precision and recall of each class to comprehensively analyze the performance of the optimized model by three TL algorithms. According to forecasted results, there is a large difference between recall and precision before calibration. Researchers have designed performance metrics in machine learning territory to consider precision and recall synthetically. The Break-Even Point (BEP) [38] is such a metric, which is the value when precision is equal to recall. When a class's precision and recall are closer to the BEP, the model has good predictive performance for this class. These data in Table 1 indicate that each class's recall and precision with the MEDA model are close to the BEP, not only recall and precision for Fuji-3 (recall = 83.72%; precision = 85.71%) but others also. However, although MEDA has obtained better results than TCA and BDA, there is still some gap between the accuracy of the calibrated model (92.02%) and the accuracy of the original model (96.22%), so there is room for improvement of the transfer model.

Table 1. Prediction results of these models. Bold figures represent the best results.

Training Set/ Testing Set	Calibration Transfer	Accuracy	Fuji-1		Fuji-2		Fuji-3		Fuji-4	
			Recall	Precision	Recall	Precision	Recall	Precision	Recall	Precision
S1/S2	None	0.9622	0.9865	0.9241	0.8837	0.95	0.9767	0.9767	0.9744	1
	None	0.8445	0.9324	0.7931	0.9070	0.6842	0.3721	1	0.9872	0.9872
	TCA	0.8613	0.9189	0.7556	0.7674	0.7857	0.6512	0.9333	0.9744	1
	BDA	0.8739	0.8784	0.7927	0.8837	0.9048	0.6279	0.8438	1	0.9512
	MEDA	0.9202	0.9324	0.8734	0.8605	0.9250	0.8372	0.8571	0.9872	1

3.3. Visualization of Wavelength Selection Process

3.3.1. Determining the Optimal Wavelength Combinations

To intuitively present the difference in spectra collected by S1 and S2, respectively, the MRE curve of each waveband is drawn in Figure 8a. The boxes with dashed borders superimposed on the curves in the figure represent the deviations of the wavebands corresponding to these curves, which are relatively large. The REA algorithm eliminates the wavebands with a large difference according to the MRE value of the wavelength, so it is requisite to determine the most appropriate filtration range. Figure 8b represents this process. In the iterative process of REA, wavebands were sorted according to the MRE value and these wavebands were removed, largest to smallest. Whenever one waveband was eliminated, the model's accuracy established by the remaining wavebands was calculated to find the optimal waveband set based on the accuracy. For example, the red curve in Figure 8b represents the MEDA-REA model's iterative process, and the model's accuracy

varies with the number of remaining wavelengths. The starred position is the optimal wavelength combination. The number of wavebands corresponding to this position is 77 (the total number of wavebands is 83). This means that when the previous 6 wavebands with large MRE values are removed, the model has achieved optimal performance, and is constructed using the surplus 77 wavebands for the MEDA algorithm.

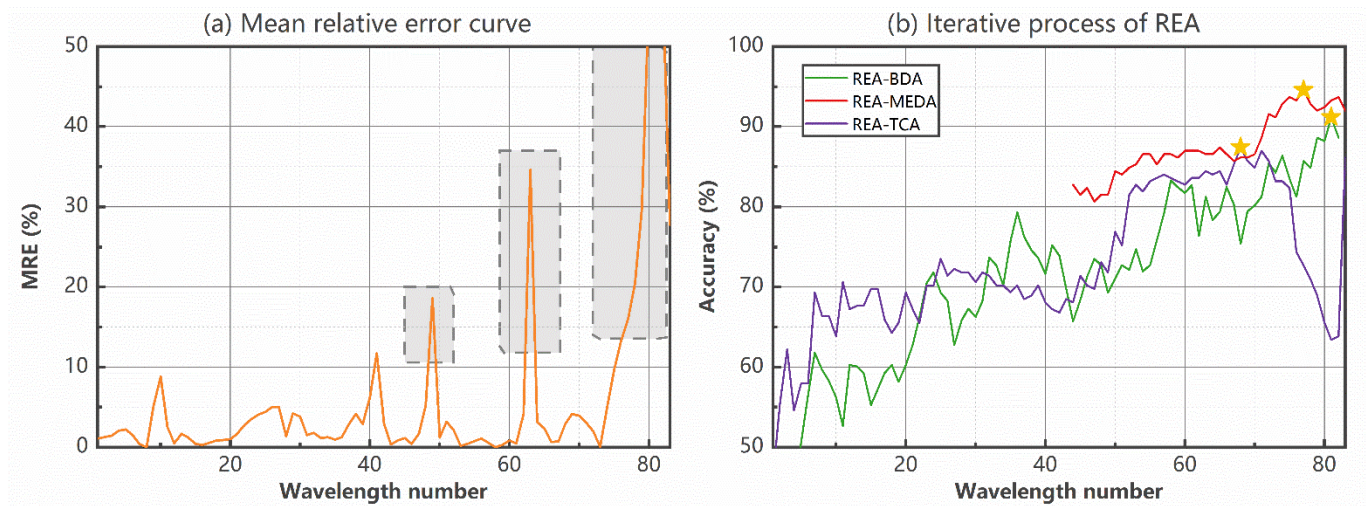


Figure 8. The wavelength selection process of the REA method. (a) Mean relative curve. (b) Iterative process of REA.

3.3.2. Prediction Results of the Optimized Model

Figure 9 shows the confusion matrices of TCA-REA, BDA-REA, and MEDA-REA. These prediction results indicate that the MEDA model's accuracy is improved from 92.02% to 94.54% by REA, and the advantage of this model's low misjudgment rate for each class is maintained. Nevertheless, REA represents the most obvious amelioration of the BDA algorithm; the misjudgment rate of BDA for Fuji-2 and Fuji-3 is reduced and the accuracy is pushed up to 91.60% (an increase of 4.21%). REA has the most negligible influence on the TCA model, its accuracy is only raised by 1.26%, and the high misjudgment rate of the model for Fuji-3 has not been improved.

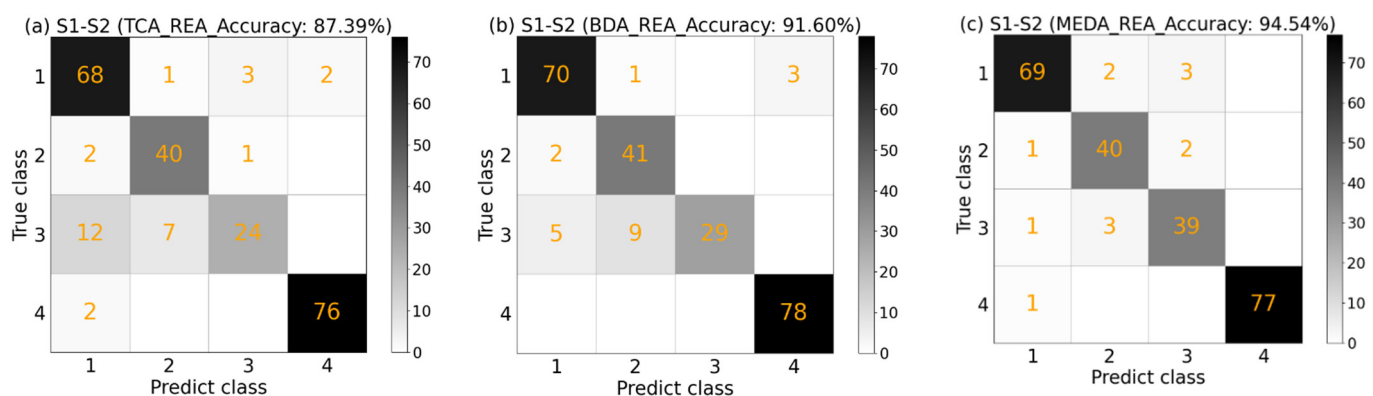


Figure 9. Confusion matrixes of the transferred model. (a) TCA-REA model. (b) BDA-REA model. (c) MEDA-REA model.

According to Table 2, REA has made the most apparent improvement to the BDA model, screening 81 wavebands, while 77 wavebands were selected by the MEDA-REA method, and 68 wavebands were selected by the TCA-REA method. Nevertheless, the predictive performance levels of the TCA-REA and BDA-REA models for Fuji-3 are still on the low side. Therefore, by comprehensively analyzing the accuracy, precision, and recall

of each model, the combined strategy of MEDA and REA is the best, effectively improving and simplifying the discrimination performance of the model.

Table 2. Prediction results of these models. Bold figures represent the best results.

Calibration Transfer	Spectral Variables	Accuracy	Fuji-1		Fuji-2		Fuji-3		Fuji-4	
			Recall	Precision	Recall	Precision	Recall	Precision	Recall	Precision
TCA-REA	68	0.8739	0.9189	0.8095	0.9302	0.8333	0.5581	0.8571	0.9744	0.9744
BDA-REA	81	0.9160	0.9459	0.9091	0.9535	0.8039	0.6744	1	1	0.9630
MEDA-REA	77	0.9454	0.9324	0.9583	0.9302	0.8889	0.9070	0.8864	0.9872	1

4. Discussion

The application of multivariate chemometric models is usually limited by devices. Even slight variations in detection environment, spectral platform, or sample state may lead to deviations in the prediction results and make the model no longer applicable [37]. The model's predictive accuracy decreased from 96.22% to 84.45% when S1 established the discrimination model applied to the testing data gathered by S2 (Figures 5 and 6).

TCA, BDA, and MEDA are used to realize standard-free calibration transfer. The data in Figure 7 has proved that the transfer performance and generality of the model can effectively be improved by these algorithms. Through a comparison of these models' accuracy, it is found that the performance of TCA is inferior to BDA and MEDA. The major reason for this phenomenon is that the distance between the source domain and the target domain data set can be minimized by TCA. However, the difference of edge distribution between the source domain and the target domain has not been considered, and the performance of tasks on unbalanced data sets is limited [27,32]. The central reason for the predominance of MEDA is that the manifold feature transformation is used to reduce the data drift between domains [33]. Therefore, the deviation of spectral signals caused by the inherent differences between instruments can be effectively reduced.

To verify REA's availability, three calibration transfer methods (TCA, BDA, and MEDA) are combined with REA, respectively. However, during the MEDA-REA iteration process, when the number of wavebands is less than 44, an error warning appears in the running code: this concatenation operation contains an empty array with the incorrect number of columns. Therefore, the red curve in Figure 8b disappears at $x \leq 44$. This error is caused by the empty array which appears when the MEDA computes the geodesic flow kernel. This means that the MEDA algorithm cannot effectively build a model with less than 44 variables in the study.

The strategy of combining REA with calibration transfer methods can improve the model's predictive performance, eliminate redundant variables, and simplify the model (Figure 9). However, by analyzing Figure 8b, it can be seen that if eliminating too many information variables will obtain a worse model and the number of removed variables is insufficient, the highest accuracy also cannot be achieved [39]. Therefore, the REA iterative process is required to explore optimal combinations of wavelengths. Nevertheless, most wavebands are still retained after wavelength selection (Table 2). The primary reason is that the two portable instruments used in this study have the same type and measurement environment, so there are relatively fewer significant deviations in the wavebands. Li et al. [13] used the REA algorithm for calibration transfer between two instruments with the same model and two instruments with different models, respectively. The research results demonstrated that when REA was applied for wavelength selection between spectrometers with different models, more wavebands were eliminated (22.98% of wavebands had been removed), and when REA was used for waveband screening between spectrometers with the same model, only 2.99% of the wavebands were filtered.

In recent years, there has been increasing study of calibration transfer, and researchers want to achieve better transfer performances than can be obtained with traditional algorithms [40,41]. Previous studies of the calibration transfer method mainly focused on calibrating spectral signals, and there has been no special treatment of wavelength selection

for the calibration transfer process. However, reasonable waveband screening plays a significant role in improving a model's versatility. Moreover, spectral response can change with various detection conditions and the wavebands determined during modeling may not be suitable for the new conditions, so the selected wavelengths for the source domain have required adjustment according to the target domain. This experiment proves that the strategy of combining wavelength selection with standard-free calibration transfer is feasible and that it can be used as part of a new direction for further of study calibration transfer methods in the future.

5. Conclusions

For the discrimination of the origin of Fuji apples, a combined strategy of wavelength selection and the TL algorithm was developed to realize standard-free calibration transfer between two near-infrared portable spectrophotometers. To unravel the issue of standard-free calibration transfer, this study combined REA with three TL algorithms (TCA, BDA, and MEDA), and seven calibration transfer models (SVM, TCA, BDA, MEDA, TCA-REA, BDA-REA, and MEDA-REA) were constructed for comparison. The results showed the discrimination performance of the MEDA-REA model to have the highest accuracy. The REA method efficiently demonstrates that it can eliminate wavebands with significant deviations and the transmission ability of the model is thereby improved. The REA method makes it possible to select the optimal wavelength combinations during the transference model step, which means that the model can be simplified and the working efficiency raised. The overall results indicate that the combined strategy of wavelength selection and the TL method can provide an available and low-cost approach to decrease modeling investment. Furthermore, more waveband screening methods for calibration transfer should be explored to simplify the transference model.

Author Contributions: Resources, Y.L.; data curation, editing and writing—original draft preparation, L.L.; writing—review and supervision, B.L.; project administration, X.J. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China, grant number 31760344.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to the request of funding scientific research projects.

Acknowledgments: This work was supported by the National Natural Science Foundation of China (grant number 31760344); the National Science and Technology Award Reserve Project Cultivation Program (grant number 20192AEI91007); and the Science and Technology Research Project of Education Department of Jiangxi Province (grant numbers GJJ200615, GJJ190306). The authors are grateful to anonymous reviewers for their comments.

Conflicts of Interest: The authors declare that they have no known competing financial interest or personal relationship that could have appeared to influence the work reported in this paper.

References

1. Giovanelli, G.; Sinelli, N.; Beghi, R.; Guidetti, R.; Casiraghi, E. NIR spectroscopy for the optimization of postharvest apple management. *Postharvest Biol. Technol.* **2014**, *87*, 13–20. [[CrossRef](#)]
2. Guo, Z.; Huang, W.; Peng, Y.; Chen, Q.; Ouyang, Q.; Zhao, J. Color compensation and comparison of shortwave near infrared and long wave near infrared spectroscopy for determination of soluble solids content of “Fuji” apple. *Postharvest Biol. Technol.* **2016**, *115*, 81–90. [[CrossRef](#)]
3. Kabir, M.H.; Guindo, M.L.; Chen, R.; Liu, F. Geographic origin discrimination of millet using vis-nir spectroscopy combined with machine learning techniques. *Foods* **2021**, *10*, 2767. [[CrossRef](#)] [[PubMed](#)]

4. Zhou, L.; Zhang, C.; Taha, M.F.; Wei, X.; He, Y.; Qiu, Z.; Liu, Y. Wheat kernel variety identification based on a large near-infrared spectral dataset and a novel deep learning-based feature selection method. *Front. Plant Sci.* **2020**, *11*, 1682. [\[CrossRef\]](#)
5. Pezzei, C.; Schönbichler, S.; Kirchler, C.; Schmelzer, J.; Hussain, S.; Huck-Pezzei, V.; Popp, M.; Krolitzek, J.; Bonn, G.; Huck, C. Application of benchtop and portable near-infrared spectrometers for predicting the optimum harvest time of *Verbena officinalis*. *Talanta* **2017**, *169*, 70–76. [\[CrossRef\]](#)
6. Feudale, R.N.; Woody, N.A.; Tan, H.; Myles, A.J.; Brown, S.D.; Ferré, J. Transfer of multivariate calibration models: A review. *Chemom. Intell. Lab. Syst.* **2002**, *64*, 181–192. [\[CrossRef\]](#)
7. Fan, S.; Li, J.; Xia, Y.; Tian, X.; Guo, Z.; Huang, W. Long-term evaluation of soluble solids content of apples with biological variability by using near-infrared spectroscopy and calibration transfer method. *Postharvest Biol. Technol.* **2019**, *151*, 79–87. [\[CrossRef\]](#)
8. Malli, B.; Birlutiu, A.; Natschläger, T. Standard-free calibration transfer—An evaluation of different techniques. *Chemom. Intell. Lab. Syst.* **2017**, *161*, 49–60. [\[CrossRef\]](#)
9. Poerio, D.V.; Brown, S.D. Dual-domain calibration transfer using orthogonal projection. *Appl. Spectrosc.* **2018**, *72*, 378–391. [\[CrossRef\]](#)
10. OSBORNE, B.G.; FEARN, T. Collaborative evaluation of universal calibrations for the measurement of protein and moisture in flour by near infrared reflectance. *Int. J. Food Sci. Technol.* **1983**, *18*, 453–460. [\[CrossRef\]](#)
11. Bouveresse, E.; Massart, D.L. Standardisation of near-infrared spectrometric instruments: A review. *Vib. Spectrosc.* **1996**, *11*, 3–15. [\[CrossRef\]](#)
12. Bin, C.; Hao, W. Calibration Transfer Between Near-infrared Spectrometric Instrument for the Determination of Wine Alcoholicity Using Shenk's Algorithm. *Infrared Technol.* **2006**, *28*, 245–248. [\[CrossRef\]](#)
13. Wang, Y.; Veltkamp, D.J.; Kowalski, B.R. Multivariate instrument standardization. *Anal. Chem.* **1991**, *63*, 2750–2756. [\[CrossRef\]](#)
14. Sulub, Y.; Small, G.W. Spectral simulation methodology for calibration transfer of near-Infrared spectra. *Appl. Spectrosc.* **2007**, *61*, 406–413. [\[CrossRef\]](#)
15. Andrew, A.; Fearn, T. Transfer by orthogonal projection: Making near-infrared calibrations robust to between-instrument variation. *Chemom. Intell. Lab. Syst.* **2004**, *72*, 51–56. [\[CrossRef\]](#)
16. Du, W.; Chen, Z.P.; Zhong, L.J.; Wang, S.X.; Yu, R.Q.; Nordon, A.; Littlejohn, D.; Holden, M. Maintaining the predictive abilities of multivariate calibration models by spectral space transformation. *Anal. Chim. Acta* **2011**, *690*, 64–70. [\[CrossRef\]](#)
17. Fan, W.; Liang, Y.; Yuan, D.; Wang, J. Calibration model transfer for near-infrared spectra based on canonical correlation analysis. *Anal. Chim. Acta* **2008**, *623*, 22–29. [\[CrossRef\]](#)
18. Chen, W.R.; Bin, J.; Lu, H.M.; Zhang, Z.M.; Liang, Y.Z. Calibration transfer via an extreme learning machine auto-encoder. *Analyst* **2016**, *141*, 1973–1980. [\[CrossRef\]](#)
19. Dong, X.; Dong, J.; Li, Y.; Xu, H.; Tang, X. Maintaining the predictive abilities of egg freshness models on new variety based on VIS-NIR spectroscopy technique. *Comput. Electron. Agric.* **2019**, *156*, 669–676. [\[CrossRef\]](#)
20. Li, L.; Huang, W.; Wang, Z.; Liu, S.; He, X.; Fan, S. Calibration transfer between developed portable Vis/NIR devices for detection of soluble solids contents in apple. *Postharvest Biol. Technol.* **2022**, *183*, 111720. [\[CrossRef\]](#)
21. Mishra, P.; Nikzad-Langerodi, R.; Marini, F.; Roger, J.M.; Biancolillo, A.; Rutledge, D.N.; Lohumi, S. Are standard sample measurements still needed to transfer multivariate calibration models between near-infrared spectrometers? The answer is not always. *TrAC Trends Anal. Chem.* **2021**, *143*, 116331. [\[CrossRef\]](#)
22. Zhang, X.; Li, Q.; Zhang, G. Calibration transfer without standards for spectral analysis based on stability competitive adaptive reweighted sampling. *Spectrosc. Spectr. Anal.* **2014**, *68*–70. [\[CrossRef\]](#)
23. Zheng, K.; Feng, T.; Zhang, W.; Huang, X.; Li, Z.; Zhang, D.; Yao, Y.; Zou, X. Variable selection by double competitive adaptive reweighted sampling for calibration transfer of near infrared spectra. *Chemom. Intell. Lab. Syst.* **2019**, *191*, 109–117. [\[CrossRef\]](#)
24. Zhang, L.; Li, Y.; Huang, W.; Ni, L.; Ge, J. The method of calibration model transfer by optimizing wavelength combinations based on consistent and stable spectral signals. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **2020**, *227*, 117647. [\[CrossRef\]](#)
25. Xu, Z.; Fan, S.; Cheng, W.; Liu, J.; Zhang, P.; Yang, Y.; Xu, C.; Liu, B.; Liu, J.; Wang, Q.; et al. A correlation-analysis-based wavelength selection method for calibration transfer. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **2020**, *230*, 118053. [\[CrossRef\]](#)
26. Li, Z.; Liu, B.; Xiao, Y. Cluster and dynamic-TrAdaBoost-based transfer learning for text classification. In Proceedings of the 13th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD), Guilin, China, 29–31 July 2017; pp. 2291–2295. [\[CrossRef\]](#)
27. Pan, S.J.; Tsang, I.W.; Kwok, J.T.; Yang, Q. Domain adaptation via transfer component analysis sinno. *IEEE Trans. Neural. Netw.* **2017**, *14*, 164–169. [\[CrossRef\]](#)
28. Po, P.; Boe, V. Easy transfer learning by exploiting intra-domain structures. In Proceedings of the IEEE International Conference on Multimedia and Expo (ICME), Shanghai, China, 8–12 July 2019.
29. Yu, Y.; Huang, J.; Liu, S.; Zhu, J.; Liang, S. Cross target attributes and sample types quantitative analysis modeling of near-infrared spectroscopy based on instance transfer learning. *Meas. J. Int. Meas. Confed.* **2021**, *177*, 109340. [\[CrossRef\]](#)
30. Mishra, P.; Roger, J.M.; Rutledge, D.N.; Woltering, E. Two standard-free approaches to correct for external influences on near-infrared spectra to make models widely applicable. *Postharvest Biol. Technol.* **2020**, *170*, 111326. [\[CrossRef\]](#)
31. Zhao, S.; Qiu, Z.; He, Y. Transfer learning strategy for plastic pollution detection in soil: Calibration transfer from high-throughput HSI system to NIR sensor. *Chemosphere* **2021**, *272*, 129908. [\[CrossRef\]](#)

32. Wang, J.; Chen, Y.; Hao, S.; Feng, W.; Shen, Z. Balanced distribution adaptation for transfer learning. In Proceedings of the IEEE International Conference on Data Mining (ICDM), New Orleans, LA, USA, 18–21 November 2017; pp. 1129–1134. [[CrossRef](#)]
33. Wang, J.; Feng, W.; Chen, Y.; Huang, M.; Yu, H.; Yu, P.S. Visual domain adaptation with manifold embedded distribution alignment. In Proceedings of the 26th ACM international conference on Multimedia, Seoul, Korea, 22–26 October 2018; Volume 1, pp. 402–410. [[CrossRef](#)]
34. Li, L.; Jang, X.; Li, B.; Liu, Y. Wavelength selection method for near-infrared spectroscopy based on standard-sample calibration transfer of mango and apple. *Comput. Electron. Agric.* **2021**, *190*, 106448. [[CrossRef](#)]
35. Galvão, R.K.H.; Araujo, M.C.U.; José, G.E.; Pontes, M.J.C.; Silva, E.C.; Saldanha, T.C.B. A method for calibration and validation subset partitioning. *Talanta* **2005**, *67*, 736–740. [[CrossRef](#)] [[PubMed](#)]
36. Lancaster, F.W. *Information retrieval systems; Characteristics, testing and evaluation*; John Wiley & Sons: Hoboken, NJ, USA, 1979; pp. 104–138.
37. Hao, Y.; Wang, Q.M.; Zhang, S.M. Study on online detection method of “Yali” pear black heart disease based on vis-near infrared spectroscopy and adaboost integrated model. *Spectrosc. Spectr. Anal.* **2021**, *41*, 2764–2769.
38. Nakache, D.; Metais, E.; Timsit, J. Evaluation and NLP. In *International Conference on Database and Expert Systems Applications*; Springer: Berlin/Heidelberg, Germany, 2005; Volume 3588, pp. 626–632.
39. Wang, H.; Peng, J.; Xie, C.; Bao, Y.; He, Y. Fruit quality evaluation using spectroscopy technology: A review. *Sensors* **2015**, *15*, 11889–11927. [[CrossRef](#)] [[PubMed](#)]
40. Shan, P.; Zhao, Y.; Wang, Q.; Ying, Y.; Peng, S. Principal component analysis or kernel principal component analysis based joint spectral subspace method for calibration transfer. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **2020**, *227*, 117653. [[CrossRef](#)]
41. Workman, J.J. A review of calibration transfer practices and instrument differences in spectroscopy. *Appl. Spectrosc.* **2018**, *72*, 340–365. [[CrossRef](#)]