

## Research Article

# Calibration Transfer of Soil Total Carbon and Total Nitrogen between Two Different Types of Soils Based on Visible-Near-Infrared Reflectance Spectroscopy

Xue-Ying Li <sup>1,2,3</sup>, Yan Liu,<sup>1,2,3</sup> Mei-Rong Lv,<sup>1,2,3</sup> Yan Zou,<sup>1,2,3</sup> and Ping-Ping Fan <sup>1,2,3</sup>

<sup>1</sup>Institute of Oceanographic Instrumentation, Qilu University of Technology (Shandong Academy of Sciences), Jinan, China

<sup>2</sup>Shandong Provincial Key Laboratory of Ocean Environmental Monitoring Technology, Qingdao, China

<sup>3</sup>National Engineering and Technological Research Center of Marine Monitoring Equipment, Qingdao, China

Correspondence should be addressed to Ping-Ping Fan; fanpp\_sdioi@126.com

Received 23 April 2018; Accepted 31 July 2018; Published 25 September 2018

Academic Editor: Khaliq Ahmed

Copyright © 2018 Xue-Ying Li et al. 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.

Although visible-near-infrared spectroscopy can rapidly and accurately determine soil nutrients without sample destruction, some problems remain unresolved, such as the mismatch of the established spectral model with different types of samples, limiting the wide application of this technology. Here, we took riverside and mountain soils as examples to explore the calibration transfer between two different types of soils by the WMPDS-S/B algorithm (wavelet multiscale piecewise direct standardization combined with Slope/Bias correction method) and by adding new samples. The predicted TN and TC concentrations improved significantly after being transformed. Compared with adding new samples, the WMPDS-S/B algorithm obtained more accurate results. The average relative errors dropped from 440.2% (without transformation) to approximately 6% for TN and from 342.0% to approximately 7% for TC. The maximum relative errors were reduced from 538.1% to less than 20% for TN and from 403.7% to less than 20% for TC. The RMSEP decreased from 2.42 to approximately 0.04 for TN and from 15.74 to approximately 0.4 for TC. The WMPDS-S/B algorithm had advantages in selecting fewer known samples and obtaining better prediction results. In contrast to past studies, which resolved the calibration transfer between different spectrometers and the measurement environment for the same samples, our study resolved the calibration transfer between different types of samples under the same spectrometer and the measurement environment. The former could only be used for correction among instruments, while the latter fundamentally solved the problem of model sharing across different samples.

## 1. Introduction

Spectral analysis is fast, accurate, and nondestructive, and it can provide rich material information [1–3]; thus, it is widely used in many fields [4–9]. In agriculture, soil nutrients have been successfully predicted by the spectral model, which is established by the known soils sampled in the same sites and under the same environment [10–14]. That is, the spectral model can work only within a specific range. If the unknown samples are different from the model samples due to altering sampling sites, seasons, or other factors, the spectral model will lead to inaccurate results. These questions have hindered the progress of spectroscopy technology; therefore, a resolution for the problem is highly needed.

Currently, studies are concentrated on the similar questions for the same samples derived from different instruments, measurement conditions, or other aspects (called calibration transfer) [15–19]. Many algorithms have successfully resolved these calibration transfer, such as direct standardization (DS) [20, 21], piecewise direct standardization (PDS) [22, 23], orthogonal signal method (OSC) [24, 25], wavelet transform (WT) algorithm [26, 27], and canonical correlation analysis (CCA) [28, 29]. Successful calibration transfer between the same type of samples could keep the accuracy of the same set of spectrometers; however, it could not resolve how to make a spectrometer share a stable spectral model for different types of samples.

Studies on calibration transfer between different types of samples are very scarce. The main solution was to add new samples in the original spectral model. Then, by using the enriched sample set, a new spectral model was rebuilt and was used for predicting the different samples. This method needed to rebuild the spectral model; thus, it was labor-intensive and time-consuming, and the prediction capability of the new spectral model was reduced. Therefore, it was urgent to explore other methods to predict unknown samples of different types quickly and accurately.

The piecewise direct standardization (PDS) algorithm was a calibration transfer method proposed by Wang and Kowalski in 1991 [30]. It required few samples, transferred the correction model faster, and improved the robustness of calibration effectively, so PDS was the most popular method for calibration transfer. However, a fixed transformation matrix was used for the whole spectral data and might not be suitable for every specific spectrum region [31, 32]. The disadvantages of PDS could be avoided by wavelet analysis. Wavelet analysis could disassemble a signal by layers and dimensions; therefore, it had good localization properties and could provide frequency information in the target signal. Its multiresolution analysis combined with PDS had been successfully used in calibration transfer for the same samples derived from different instruments [33–35]. We hypothesized that it could also be used for calibration transfer between different types of samples.

Concentrations of total nitrogen (TN) and total carbon (TC) in soils were better predicted than other soil nutrients by spectral models and were thus selected for the study of calibration transfer [36, 37]. The Slope/Bias algorithm (S/B) was often used to correct the result of the chemical values in calibration transfer, which rebuilt the linear relationships based on abundant data [38]. For instance, Cooper et al. provided a simple slope and bias correction to transfer calibration with different instruments [39]. We would also use S/B to reduce the differences between concentrations of soil nutrients.

Here, we explored the calibration transfer of TN and TC concentrations between riverside and mountain soil samples using wavelet multiscale piecewise direct standardization combined with the Slope/Bias correction method (WMPDS-S/B). Additionally, we compared the results with those by adding new samples. To reduce the influence of other factors on calibration transfer, we sampled spectral data from the same equipment and the same test environment.

## 2. Experimental

**2.1. Samples and Data Acquisition.** Two types of upper soils (0–20 cm) were collected from the riverside of Licun River and the foot of Fushan Mountain in Qingdao, China. Licun River soil (riverside soil) was silt loam rich in nutrients, while Fushan Mountain soil (mountain soil) was sandy loam poor in nutrients. We removed foreign bodies such as stones from the soil samples. After drying to a constant weight, all soil samples were filtered through a 0.45 mm nylon screen. Soil TN and TC concentrations were measured with a carbon and nitrogen analyzer (Perkin Elemental Analyzer,

USA). The concentrations of TN and TC between two types of soils have a large difference (Table 1).

Soil spectral data were determined by an Ocean Optics QE65000 spectrometer. The sampling interval was 1 nm, and the integral time was 1000 ms. The spectral range was 200–1100 nm, mainly the visible and near-infrared spectrum, including a small amount of the ultraviolet spectrum. The optical fiber probe was inserted into the hole of the probe bracket at a 45° angle, and this made the probe stick to the bracket and just exposed the bracket.

Soil (3–5 g) was gently flattened in a homemade sample box, whose size was the same as the probe bracket and whose cell overlapped the hole of the probe bracket. The schematic diagram of soil spectral data measurements is shown in Figure 1. Each soil sample was measured five times. To reduce the influence of the noise in the reflectance spectrum, the spectral data of 226–975 nm were retained. The reflectance spectrum of the river and mountain soils is shown in Figure 2.

The principal component space was constructed by the scores of the spectral principal component (Figure 3), and the distribution of the principal components in the samples could be used to determine whether the unknown samples were fit to the original model. As shown in Figure 3, the river soil and mountain soil samples were divided into two regions in the principal component space. This outcome indicated that the spectrum of the two types of soil samples was obviously different. Therefore, the model established by the river soil was hardly suitable for predicting the mountain soil directly.

### 2.2. Data Treatment and Analysis

**2.2.1. Pretreatment and Modeling.** The reflectance spectral data (226–975 nm) of the river and mountain soils were pretreated using Savitzky–Golay smoothing and differentiation. The calibration set and test set of the two types of soil samples were divided into 3:1 proportions by the Kennard–Stone algorithm, with 45 soil samples as the calibration set and the rest of the samples as the test set. The spectral model of the calibration set was built by partial least squares (PLS) [40, 41] (Table 2). Results showed that a better spectral model was built by the river soil (Table 2); therefore, it was set as the First Sample and mountain soil was set as the Second Sample.

**2.2.2. Calibration Transfer Algorithm.** Here, we use two methods: wavelet multiscale piecewise direct standardization combined with the Slope/Bias correction method (WMPDS-S/B) and the method of adding new samples. WMPDS-S/B is more complicated. Its specific steps are described below.

- (1) Wavelet decomposition is used to obtain the coefficients of each wavelet.

First, the wavelet generating function and the number of wavelet decomposition layers  $N$  are selected. Then, the spectrum of the First Sample  $X_m$  and the spectrum of the

TABLE 1: Concentrations of TC and TN in the river and mountain soils.

Samples	Number	TN concentrations (mg/kg)		TC concentrations (mg/kg)	
		Range	Mean $\pm$ SD	Range	Mean $\pm$ SD
River soil	60	4.14–5.66	4.70 $\pm$ 0.40	3.8–13.4	10.58 $\pm$ 2.24
Mountain soil	58	0.44–0.69	0.55 $\pm$ 0.05	0.72–1.74	1.40 $\pm$ 0.26

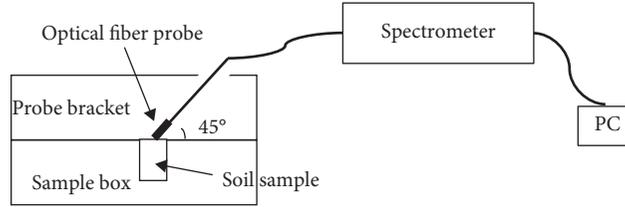


FIGURE 1: The schematic diagram of soil spectral data measurements.

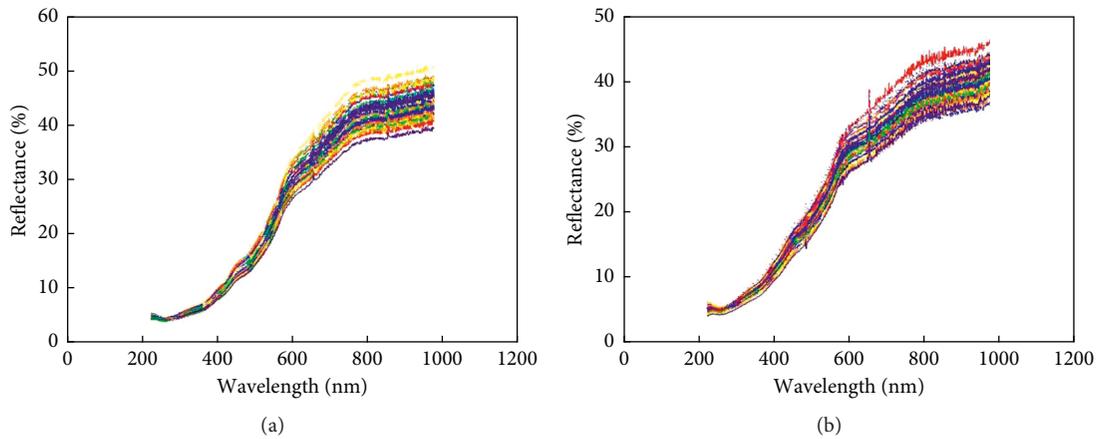


FIGURE 2: Visible-near-infrared reflectance spectrum of the river soil (a) and mountain soil (b).

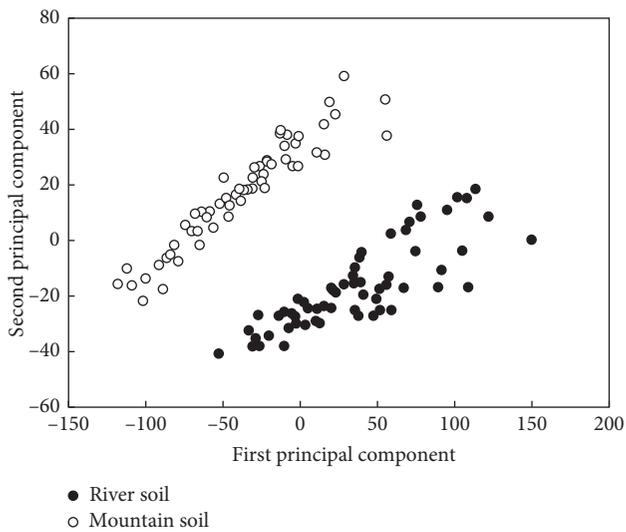


FIGURE 3: Principal component score diagram of the river and mountain soils.

Second Sample  $X_s$  are decomposed by wavelet analysis, and the approximate coefficients  $Xm_a$  and  $Xs_a$  and detail coefficients  $Xm_{dn}$  and  $Xs_{dn}$  ( $n = 1, 2, \dots, N$ ) of the First Sample and the Second Sample are obtained. The correlation formula is as follows:

$$\begin{aligned} Xm &\xrightarrow{\text{waveletdecomposition}} Xm_a + Xm_{d1} + \dots + Xm_{dN}, \\ Xs &\xrightarrow{\text{waveletdecomposition}} Xs_a + Xs_{d1} + \dots + Xs_{dN}. \end{aligned} \quad (1)$$

- (2) The transformation matrices  $F_n$  ( $n = 0, 1, \dots, N$ ) of each wavelet coefficients are calculated using the PDS algorithm.

A part of the soil samples from the Second Sample are taken as standard samples for calculating the transformation matrices, denoted as standard set  $Xs_{\text{standard}}$ . The remainders of the Second Sample are unknown samples, denoted as unknown set  $Xs_{\text{unknown}}$ . If we want to calculate the transformation matrices  $F_0$  of the approximate coefficients between the calibration set of the First Sample and the standard

TABLE 2: Spectral models built by the river and mountain soils.

Samples	Number	TN concentrations (mg/kg)					TC concentrations (mg/kg)				
		$r_c$	$r_p$	RMSEC	RMSEP	RPD	$r_c$	$r_p$	RMSEC	RMSEP	RPD
River soil	60	0.97	0.85	0.064	0.142	1.49	0.93	0.85	0.869	1.095	1.64
Mountain soil	58	0.40	0.36	0.046	0.031	1.09	0.86	0.60	0.194	0.323	1.28

set of the Second Sample  $Xm_{a,training}$ ,  $Xs_{a,standard}$ , the average spectral  $Mm_a$ ,  $Ms_a$  of  $Xm_{a,training}$ ,  $Xs_{a,standard}$  are first obtained, and the formulas are as follows:

$$Mm_a(j) = \frac{1}{n_1} \sum_{i=1}^{n_1} Xm_{a,training}(i, j), \quad (2)$$

$$Ms_a(j) = \frac{1}{n_2} \sum_{i=1}^{n_2} Xs_{a,standard}(i, j),$$

where  $(i, j)$  is the  $j$ th wavelength point of the  $i$ th sample and  $n_1$  and  $n_2$  are the number of samples for  $Xm_{a,training}$  and  $Xs_{a,standard}$ .

The spectral band of the window width  $(j - k \sim j + k)$  is intercepted near the  $j$ th wavelength point. Let  $Z_{a,j} = [Ms_{a,j-k}, \dots, Ms_{a,j}, \dots, Ms_{a,j+k}]$ , make  $Mm_a(j) = Z_{a,j} \times f_{a,j}$ , and then the regression coefficients  $f_{a,j}$  are calculated. Loop  $j$ , and find all the  $f_{a,j}$ . Put  $f_{a,j}$  on the main diagonal of the transformation matrix  $F_0$ , and set other elements as 0, and finally get the transformation matrix  $F_0$ . According to the above algorithms, we can obtain the transformation matrices  $F_1, \dots, F_N$  of the detail coefficients between the calibration set of the First Sample and the standard set of the Second Sample.

- (3) The transfer spectrum of the standard set and unknown set in the Second Sample is obtained by using transformation matrices.

According to the transformation matrix  $F_0$ , the transfer approximation coefficients  $XsF_{a,standard}$  and  $XsF_{a,unknown}$  are achieved by the transforming approximation coefficients in the standard set and unknown set of the Second Sample  $Xs_{a,standard}$  and  $Xs_{a,unknown}$ .

$$\begin{aligned} XsF_{a,standard} &= Xs_{a,standard} \cdot F_0, \\ XsF_{a,unknown} &= Xs_{a,unknown} \cdot F_0. \end{aligned} \quad (3)$$

Then, the transfer detail coefficients of the standard set and unknown set in the Second Sample  $XsF_{dn,standard}$  and  $XsF_{dn,unknown}$  ( $n = 1, 2, \dots, N$ ) are calculated sequentially. Finally, the transfer spectrum  $XsF_{all,standard}$  and  $XsF_{all,unknown}$  are obtained by reconstructing each wavelet coefficient of the Second Sample.

- (4) The Slope/Bias correction method is used to calculate the final prediction value.

The least squares solution is obtained by fitting the transfer spectrum  $XsF_{all,standard}$  and the measured nutrient values  $Y_{standard}$  in a standard set of the Second Sample with linear regression, providing the slope  $slope$  and intercept  $bias$  of the linear model. By substituting  $XsF_{all,unknown}$  into

the formula, we can find the prediction values of the unknown set  $Y_{unknown}$  by the following equation:

$$Y_{unknown} = slope \cdot XsF_{all,unknown} + bias. \quad (4)$$

In this paper, the wavelet generating function chooses "db3," and the decomposition layer number  $N$  is 3.

**2.2.3. Evaluation Standard.** The evaluation standard used the average relative error ( $error_a$ ), the maximum relative error ( $error_m$ ), and the prediction root mean square error (RMSEP), which offered a comprehensive analysis. The smaller the average relative error, the maximum relative error and RMSEP were, the better the calibration transfer was.

### 3. Results and Discussion

**3.1. Calibration Transfer of TN Concentrations between Two Different Types of Soils.** The calibration transfer of TN concentrations between two different types of soils was realized by WMPDS-S/B. The PDS window width and the number of the standard set in the Second Sample affected the results of calibration transfer (Figure 4). We made the number of the standard set 15 by the Kennard–Stone algorithm, the window width ranged from 3 to 19, and the interval was 2. The influence of different window widths on the results was studied. We studied the effects of the number of the standard set on the results. Setting the window width to 15, the number of the standard set ranged from 11 to 20 by the Kennard–Stone algorithm, and the interval was 1. When the window width was 15 and the number of the standard set was 14, the average relative error was the smallest (Figure 4). Therefore, we chose this parameter for the subsequent calibration transfer.

Using the WMPDS-S/B algorithm, the average relative error ( $error_a$ ), the maximum relative error ( $error_m$ ), and RMSEP were obviously decreased (Table 3). The  $error_a$  dropped from 440.2% to 5.77%. The  $error_m$  were reduced from 538.1% to 16.67%, and the RMSEP was reduced from 2.416 to approximately 0.041 (Table 3).

We used the optimum number of the standard set to divide the standard set and unknown set in the Second Sample by the Kennard–Stone algorithm; the result of the calibration transfer was accidental (Table 3). To avoid the contingency, we used interval sampling by proportion (Table 4) and random sampling (Table 5) to divide the standard set and unknown set. The two methods of sample classification led to similar results of calibration transfer (Tables 4 and 5).

The different proportions between the standard set and unknown set in the Second Sample by the Kennard–Stone

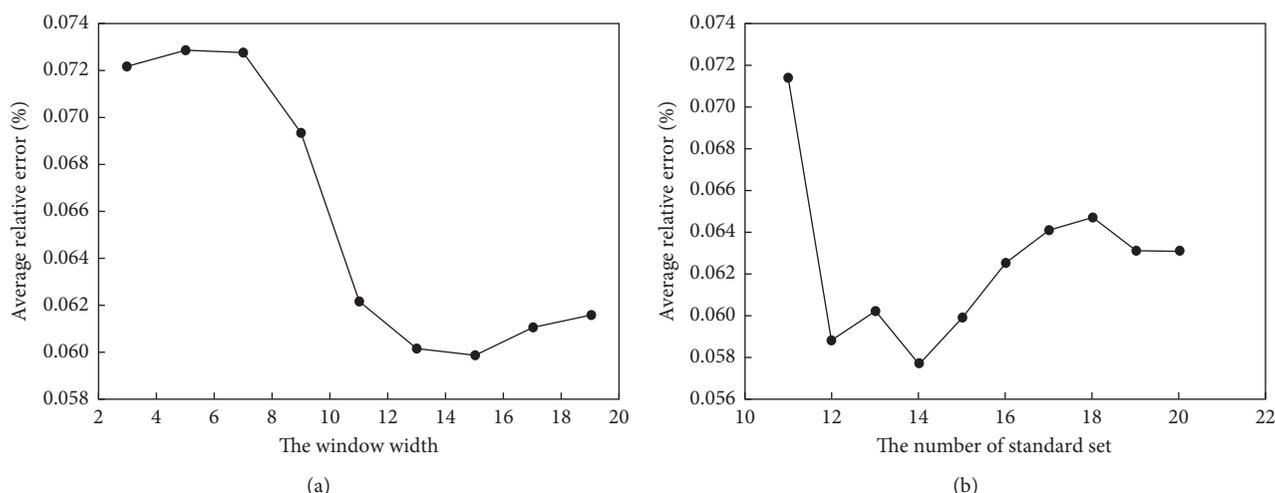


FIGURE 4: Influence of the window width (a) and the number of the standard set (b) on the transfer results of soil TN concentration.

TABLE 3: Comparison of TN concentrations statistical results with and without calibration transfer.

Calibration transfer algorithm	Error <sub>a</sub> (%)	Error <sub>m</sub> (%)	RMSEP
Without algorithm	440.2	538.1	2.416
WMPDS-S/B	5.77	16.67	0.041

TABLE 4: Comparison of TN concentrations statistical results from dividing standard sets and unknown sets using different proportions.

Proportion	Error <sub>a</sub> (%)	Error <sub>m</sub> (%)	RMSEP
1:2	5.58	18.93	0.041
1:3	6.05	17.55	0.041
1:4	6.52	15.87	0.044

TABLE 5: Comparison of TN concentrations statistical results from dividing standard sets and unknown sets by using random sample.

Random sampling number	Error <sub>a</sub> (%)	Error <sub>m</sub> (%)	RMSEP
Random 1	5.34	18.59	0.039
Random 2	5.83	19.76	0.044
Random 3	6.12	17.87	0.044
Random 4	5.93	17.38	0.042
Random 5	5.50	17.79	0.040

algorithm could bring different results of calibration transfer (Table 4). Here, the results showed that, with increased proportion between the standard set and unknown set (from 1:2 to 1:4), the error<sub>a</sub> increased, the error<sub>m</sub> decreased, and the RMSEP increased (Table 4). It seemed that the proportion of 1:2 was the best, perhaps because the number of the standard set was proper and contained enough information.

Random sampling did not have a big influence on the results of calibration transfer (Table 5). The values of error<sub>a</sub>, error<sub>m</sub>, and RMSEP derived from random sampling were similar with those derived from interval sampling by the Kennard–Stone algorithm (Tables 4 and 5).

3.2. Calibration Transfer of TC Concentrations between Two Different Types of Soils. Using the same method, the PDS window width and the number of the standard set in the Second Sample were also found to affect the results of calibration transfer (Figure 5). When the window width was 11 and the number of the standard set was 18, the calibration transfer of the TC concentrations model was the best (Figure 5). Therefore, we chose these parameters for the subsequent calibration transfer of TC concentrations.

Under the best condition, the standard set and unknown set were classified by the Kennard–Stone algorithm, proportion, and random sampling, respectively (Table 6). The accuracy of prediction was significantly improved after the WMPDS-S/B algorithm was used for the calibration transfer (Table 6). The error<sub>a</sub> decreased from 341.97 to approximately 6%, the error<sub>m</sub> decreased from 403.70% to <20%, and the RMSEP decreased from 15.737 to <0.46 (Table 6).

Among all the prediction results, the best was that from the 1:3 proportion between the standard set and unknown set (Table 6), whose error<sub>a</sub>, error<sub>m</sub>, and RMSEP were the smallest. The results of the 1:2 proportion were relatively poor, maybe because the number of the standard set was too large and some unrelated information was included in the calibration transfer.

Some division approaches could lead to spectral over-correction and reduce the prediction effect of the calibration transfer. But in this study, when the number of standard set was fixed to 18, the values of error<sub>a</sub>, error<sub>m</sub>, and RMSEP were similar across different classification methods (Table 6). This shows that different classification methods produce similar prediction results, suggesting the division approaches in this study were reasonable.

3.3. Comparison between the WMPDS-S/B Algorithm and Adding Samples. Under the condition of selecting the same samples of the standard set by the Kennard–Stone algorithm, we then study the effects of the two methods on the results of calibration transfer in TN and TC concentrations (Tables 7 and 8). The number of the standard set is 10, 15, 20, 25, and

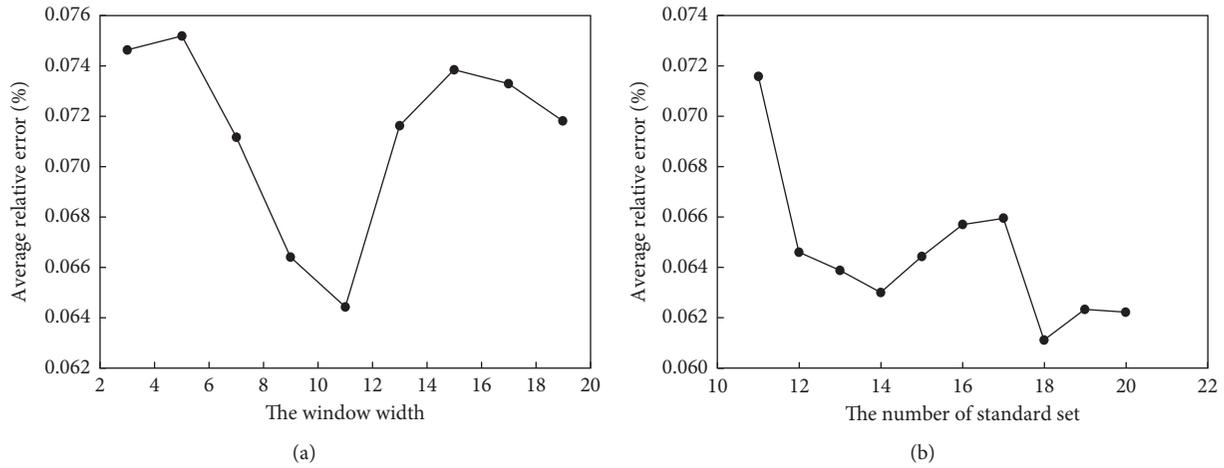


FIGURE 5: Influence of the window width (a) and the number of the standard set (b) on the transfer results of soil TC concentrations.

TABLE 6: Comparison of TC concentrations statistical results with and without calibration transfer.

Calibration transfer algorithm	Standard set division method	Error <sub>a</sub> (%)	Error <sub>m</sub> (%)	RMSEP
Without algorithm	Kennard–Stone	341.97	403.70	15.737
WMPDS-S/B	Kennard–Stone	6.11	15.37	0.351
WMPDS-S/B	Proportion 1 : 2	7.17	18.16	0.453
WMPDS-S/B	Proportion 1 : 3	6.79	15.96	0.380
WMPDS-S/B	Proportion 1 : 4	6.92	16.27	0.406
WMPDS-S/B	Random 1	5.78	15.75	0.349
WMPDS-S/B	Random 2	6.25	16.54	0.384
WMPDS-S/B	Random 3	6.67	18.57	0.394
WMPDS-S/B	Random 4	6.08	19.06	0.430
WMPDS-S/B	Random 5	6.24	16.50	0.357

TABLE 7: Comparison of different calibration transfer algorithms for the modified result of TN concentrations.

The number of standard set	Result	First Sample		Second Sample		
		WMPDS-S/B	Add samples	Result	WMPDS-S/B	Add samples
10	$r_c$	0.971	1.000	Error <sub>a</sub> (%)	7.52	33.28
	$r_p$	0.851	0.421	Error <sub>m</sub> (%)	20.34	81.10
	RPD	1.49	1.031	RMSEP	0.050	0.214
15	$r_c$	0.971	0.975	Error <sub>a</sub> (%)	5.99	27.81
	$r_p$	0.851	0.656	Error <sub>m</sub> (%)	15.89	76.06
	RPD	1.49	1.069	RMSEP	0.042	0.176
20	$r_c$	0.971	1.000	Error <sub>a</sub> (%)	6.38	25.30
	$r_p$	0.851	0.594	Error <sub>m</sub> (%)	17.18	63.36
	RPD	1.49	1.090	RMSEP	0.043	0.166
25	$r_c$	0.971	1.000	Error <sub>a</sub> (%)	6.21	24.87
	$r_p$	0.851	0.674	Error <sub>m</sub> (%)	16.60	65.58
	RPD	1.49	1.186	RMSEP	0.042	0.168
30	$r_c$	0.971	1.000	Error <sub>a</sub> (%)	6.55	23.68
	$r_p$	0.851	0.707	Error <sub>m</sub> (%)	16.57	64.82
	RPD	1.49	1.195	RMSEP	0.044	0.165

30. The window width of the WMPDS-S/B algorithm in TN concentrations is 15 and in TC concentrations is 11. The results of the First Sample include the relative coefficients of calibration set ( $r_c$ ), the relative coefficients of test set ( $r_p$ ), and

the RPD. The results of the unknown set of the Second Sample include the error<sub>a</sub>, error<sub>m</sub>, and RMSEP.

The prediction performance (the value of the  $r_c$ ,  $r_p$ , and RPD) of the original model did not change by using the

TABLE 8: Comparison of different calibration transfer algorithms for the modified result of TC concentrations.

The number of standard set	Result	First Sample		Second Sample		
		WMPDS-S/B	Add samples	Result	WMPDS-S/B	Add samples
10	$r_c$	0.926	0.962	Error <sub>a</sub> (%)	6.43	34.69
	$r_p$	0.853	0.699	Error <sub>m</sub> (%)	18.24	86.51
	RPD	1.64	1.016	RMSEP	0.380	1.821
15	$r_c$	0.926	0.963	Error <sub>a</sub> (%)	6.44	31.51
	$r_p$	0.853	0.733	Error <sub>m</sub> (%)	18.80	75.83
	RPD	1.64	1.036	RMSEP	0.378	1.714
20	$r_c$	0.926	0.966	Error <sub>a</sub> (%)	6.22	26.59
	$r_p$	0.853	0.697	Error <sub>m</sub> (%)	15.21	62.72
	RPD	1.64	0.990	RMSEP	0.358	1.403
25	$r_c$	0.926	0.938	Error <sub>a</sub> (%)	6.25	24.98
	$r_p$	0.853	0.676	Error <sub>m</sub> (%)	15.66	59.33
	RPD	1.64	1.096	RMSEP	0.364	1.242
30	$r_c$	0.926	0.941	Error <sub>a</sub> (%)	6.72	26.07
	$r_p$	0.853	0.679	Error <sub>m</sub> (%)	15.95	59.33
	RPD	1.64	1.106	RMSEP	0.391	1.262

WMPDS-S/B in the First Sample model of soil TN and TC concentrations (Tables 7 and 8). With the addition of the number of the standard set, the prediction performance of the model by adding new samples was unstable (Tables 7 and 8), with high  $r_c$  but low  $r_p$  and RPD. So, WMPDS-S/B could keep the stability of the original model.

Among all calibration transfer algorithms, the best was the WMPDS-S/B algorithm, while the method of adding new samples was the worst. By adding new samples, the error<sub>a</sub>, error<sub>m</sub>, and RMSEP of TN and TC concentrations were all high. Therefore, it could not be selected for calibration transfer between two different types of soils. The WMPDS-S/B algorithm was based on the S/B algorithm and combined with the WMPDS algorithm to eliminate spectral differences. It could improve the accuracy of nutrient concentration prediction between two different types of soil.

The number of the standard set had a relatively small impact on the results by the WMPDS-S/B algorithm (Tables 7 and 8). With the increase in the number of the standard set in the Second Sample, the error<sub>a</sub>, error<sub>m</sub>, and RMSEP decreased by adding new samples, showing that the addition of the samples was beneficial to the improvement of the prediction. When the number of the standard set increased to 30, which was the largest samples in both the TN and TC concentrations, the result was the best. However, when the number of the standard set was a certain value in the middle, which was 15 and 20 in the TN and TC concentrations, respectively, the result had the best prediction effect by the WMPDS-S/B algorithm. Compared to the first method, the number of the standard set had a relatively small impact on the results by the WMPDS-S/B algorithm, and this method could get better results in selecting as few a number of standard set as possible. In practical applications, it was impossible to select many standard set samples, which would increase workload and be time-consuming. The new algorithm had advantages in selecting fewer standard samples and obtaining better prediction results.

**3.4. Differences from Previous Studies.** The calibration transfer studied in this paper was different from previous studies, which were mainly for the calibration transfer between two different types of samples. Previous studies had solved the change of the same sample in different instruments and testing environments, thus ensuring the consistency of the instruments [13–18]. Fernandez et al. used four different calibration transfer techniques (Direct Standardization, Piecewise Direct Standardization, Orthogonal Signal Correction, and Generalized Least Squares Weighting) to offset the effect of temperature change on the prediction of gas sensors [13]. Yahaya et al. compared the calibration transfer effect of the reflectance spectrum of pH for mango among different spectrometers (QE65000, Jaz, and ASD FieldSpec 3) through direct calibration transfer [18]. These studies are all based on the calibration transfer of the same samples under different test conditions or different instruments.

At present, many studies were based on spectroscopy for the prediction of different types of soil nutrients. For example, Kuang et al. used near-infrared spectroscopy to detect the content of soil organic carbon and total nitrogen, and soil samples were collected on five farms. The result for the RPD of organic carbon OC was from 2.66 to 3.39 and of total nitrogen TN was from 2.85 to 3.45 [42]. The soil nutrient concentrations model was established through many soil samples at different locations, and this method still needed a large number of known information about different types of soil nutrient chemical values, so it was not the soil calibration transfer in real sense.

In this study, we solved calibration transfer between two different types of samples in the same instrument and the test environment using one type of soil nutrient concentrations model to predict another type of soil nutrient concentration. If the concentrations model of riverside soil was used to predict the concentrations of mountain soil directly, there was a large error (Tables 3 and 6). After we did the calibration transfer, the concentrations of mountain soil were predicted more accurately. The error significantly

declined, and the prediction results could be accepted. The calibration transfer that we studied in this work could fundamentally solve the problem of model sharing between different types of samples and eliminate the technical bottleneck restricting the application of the spectral method. This solution would make it possible for one spectral instrument to rapidly share a spectral model to determine soil nutrients.

**3.5. Future Studies on Calibration Transfer with Different Samples.** Wavelet analysis has multiscale and multi-resolution characteristics. It can process signals and analyze spectral information from coarse to fine. The spectral data can be decomposed into two subspaces by multiscale wavelet processing, that is, low frequency and high frequency. Calibration transfer between low-frequency and high-frequency information is carried out. Different transfer matrices can be obtained according to different frequency domains. The transfer matrices between different regions and different varieties are further refined. However, there are some problems in the WMPDS-SB algorithm.

- (1) The PDS window width and the number of standard samples still need to be selected artificially. In the future, we can select the optimal PDS window and the number of samples automatically through algorithms.
- (2) The entire band of the spectrum is used in this study. To reduce redundant information and improve prediction performance, we will extract the diagnostic spectral wavelength in the future.
- (3) In this study, the method of calibration transfer in soil nutrients is labeled samples, which still require a portion of the chemical value of unknown samples. In the future, we intend to achieve the calibration transfer among different regions or different varieties with smaller labeled samples or without labeled samples.
- (4) The WMPDS-S/B algorithm is applied to the calibration transfer in two different types of soil and has good results. In the future, this algorithm can be applied to calibration transfer in different types of soil samples.

## 4. Conclusions

This paper took mountain and riverside soils as examples to explore the calibration transfer of TC and TN concentrations between two different types of soils using WMPDS-S/B. To reduce the influence of calibration transfer in the same samples, we sampled spectral data with the same equipment and in the same test environment. The riverside soil was set as the First Sample, and the mountain soil was regarded as the Second Sample. First, a calibration model was established by the First Sample. Then, concentrations of total carbon (TC) and total nitrogen (TN) for the Second Sample were predicted by the transferred calibration model. In the calibration transfer of TN concentrations with two

different types of soil, when the PDS window width was 15 and the number of standard set in the Second Sample was 14, the prediction accuracy of TN concentrations improved significantly compared to the accuracy without using calibration transfer. The average relative error ( $error_a$ ) dropped from 440.20% to approximately 6%, the maximum relative error ( $error_m$ ) was reduced from 538.1% to less than 20%, and RMSEP was reduced from 2.416 to approximately 0.04. In calibration transfer of TC concentrations between two different types of soils, when the PDS window width was 11 and the number of standard set was 18, the prediction accuracy of TC concentrations in the Second Sample showed significant improvement compared to the accuracy without using calibration transfer. The  $error_a$  dropped from 341.97% to approximately 7%, the  $error_m$  was reduced from 403.70% to less than 20%, and the RMSEP was reduced from 15.737 to approximately 0.4.

Under the condition of selecting the same samples of standard set, we studied the effects of the three methods on the results of calibration transfer in TN and TC concentrations. Compared to the traditional method of adding new samples, the WMPDS-S/B algorithm obtained better forecasting results. With the increase of the number of standard set in the Second Sample,  $error_a$ ,  $error_m$ , and RMSEP were decreased by adding new samples. However, the number of the standard set had a relatively small impact on the results from the WMPDS-S/B algorithm. This method had advantages in selecting fewer standard samples and obtaining better prediction results. It was feasible to realize the calibration transfer of soil nutrients between two different types of soils rapidly and accurately, providing new ideas and methods for calibration transfer among different types of samples.

## Data Availability

All data in the paper are fully available without restriction at <https://figshare.com/s/6c301c6479dc3f1a554c>, or from the corresponding author upon request.

## Conflicts of Interest

The authors report there are no conflicts of interest.

## Acknowledgments

This work was supported by the National Key Research and Development Program of China (2017YFC1404802), Shandong Provincial Natural Science Foundation, China (ZR2018LD007), the National Key Research and Development Program of China (no. SQ2016YFSF070090), and the National Key Research and Development Program of China (no. SQ 2017YFC1403702).

## References

- [1] K. S. Hong and H. Santosa, "Decoding four different sound-categories in the auditory cortex using functional near-infrared spectroscopy," *Hearing Research*, vol. 333, pp. 157–166, 2016.

- [2] A. N. Sen, S. P. Gopinath, and C. S. Robertson, "Clinical application of near-infrared spectroscopy in patients with traumatic brain injury: a review of the progress of the field," *Neurophotonics*, vol. 3, no. 3, p. 031409, 2016.
- [3] Y. Guo, Y. Ni, and S. Kokot, "Evaluation of chemical components and properties of the jujube fruit using near infrared spectroscopy and chemometrics," *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, vol. 153, pp. 79–86, 2016.
- [4] M. Schmutzler, A. Beganovic, G. Böhler et al., "Modern safety control for meat products: near infrared spectroscopy utilised for detection of contaminations and adulterations of premium veal products," *NIR News*, vol. 27, no. 4, pp. 11–13, 2016.
- [5] I. Marquetti, J. V. Link, and E. Bona, "Partial least square with discriminant analysis and near infrared spectroscopy for evaluation of geographic and genotypic origin of arabica coffee," *Computers and Electronics in Agriculture*, vol. 121, pp. 313–319, 2016.
- [6] A. M. Rady, D. E. Guyer, W. Kirk, and I. R. Donis-González, "The potential use of visible/near infrared spectroscopy and hyperspectral imaging to predict processing-related constituents of potatoes," *Journal of Food Engineering*, vol. 135, pp. 11–25, 2014.
- [7] H. Li, Q. Chen, J. Zhao, and M. Wu, "Nondestructive detection of total volatile basic nitrogen (TVB-N) content in pork meat by integrating hyperspectral imaging and colorimetric sensor combined with a nonlinear data fusion," *LWT-Food Science and Technology*, vol. 63, no. 1, pp. 268–274, 2015.
- [8] U. Khulal, J. Zhao, W. Hu, and Q. Chen, "Nondestructive quantifying total volatile basic nitrogen (TVB-N) content in chicken using hyperspectral imaging (HSI) technique combined with different data dimension reduction algorithms," *Food Chemistry*, vol. 197, pp. 1191–1199, 2016.
- [9] F. Kutsanedzie, Q. Chen, M. M. Hassan, M. Yang, H. Sun, and M. H. Rahman, "Near infrared system coupled chemometric algorithms for enumeration of total fungi count in cocoa beans neat solution," *Food Chemistry*, vol. 240, pp. 231–238, 2018.
- [10] F. Feyziyev, M. Babayev, S. Priori, and G. L'Abate, "Using visible-near infrared spectroscopy to predict soil properties of Mugan Plain, Azerbaijan," *Open Journal of Soil Science*, vol. 06, no. 3, pp. 52–58, 2016.
- [11] L. Zhang and R. B. Zhang, "Measurement of available phosphorus and potassium contents in soil using visible-near-infrared spectroscopy in conjunction with SPA-LS-SVM methods," *Advance Journal of Food Science and Technology*, vol. 10, no. 12, pp. 934–941, 2016.
- [12] X. Yu, Q. Liu, Y. Wang, X. Liu, and X. Liu, "Evaluation of MLSR and PLSR for estimating soil element contents using visible/near-infrared spectroscopy in apple orchards on the Jiaodong peninsula," *CATENA*, vol. 137, pp. 340–349, 2016.
- [13] N. K. Niazi, B. Singh, and B. Minasny, "Mid-infrared spectroscopy and partial least-squares regression to estimate soil arsenic at a highly variable arsenic-contaminated site," *International Journal of Environmental Science and Technology*, vol. 12, no. 6, pp. 1–10, 2015.
- [14] A. M. Mouazen, B. Kuang, J. De Baerdemaeker, and H. Ramon, "Comparison among principal component, partial least squares and back propagation neural network analyses for accuracy of measurement of selected soil properties with visible and near infrared spectroscopy," *Geoderma*, vol. 158, pp. 23–31, 2010.
- [15] L. Fernandez, S. Guney, A. Gutierrez-Galvez, and S. Marco, "Calibration transfer in temperature modulated gas sensor arrays," *Sensors and Actuators B Chemical*, vol. 231, pp. 276–284, 2016.
- [16] J. H. Kalivas, G. G. Siano, E. Andries, and H. C. Goicoechea, "Calibration maintenance and transfer using Tikhonov regularization approaches," *Applied Spectroscopy*, vol. 63, no. 7, pp. 800–809, 2009.
- [17] S. Amat-Tosello, N. Dupuy, and J. Kister, "Contribution of external parameter orthogonalisation for calibration transfer in short waves--near infrared spectroscopy application to gasoline quality," *Analytica Chimica Acta*, vol. 642, no. 1–2, pp. 6–11, 2009.
- [18] S. Guo, R. Heinke, S. Stöckel, P. Rösch, T. Bocklitz, and J. Popp, "Towards an improvement of model transferability for Raman spectroscopy in biological applications," *Vibrational Spectroscopy*, vol. 91, pp. 111–118, 2017.
- [19] O. K. M. Yahaya, M. Z. Matjafri, A. A. Aziz, and A. F. Omar, "Visible spectroscopy calibration transfer model in determining pH of Sala mangoes," *Journal of Instrumentation*, vol. 10, no. 5, article T05002, 2015.
- [20] A. L. B. Brito, A. V. P. Santos, K. D. T. M. Milanez, M. J. C. Pontes, and L. F. B. L. Pontes, "Calibration transfer of flour NIR spectra between benchtop and portable instruments," *Analytical Methods*, vol. 9, no. 21, pp. 3184–3190, 2017.
- [21] C. F. Pereira, M. F. Pimentel, R. K. Galvão, F. A. Honorato, L. Stragevitch, and M. N. Martins, "A comparative study of calibration transfer methods for determination of gasoline quality parameters in three different near infrared spectrometers," *Analytica Chimica Acta*, vol. 611, no. 1, pp. 41–47, 2008.
- [22] L. S. A. Pereira, M. F. Carneiro, B. G. Botelho, and M. M. Sena, "Calibration transfer from powder mixtures to intact tablets: a new use in pharmaceutical analysis for a known tool," *Talanta*, vol. 147, pp. 351–7, 2016.
- [23] Y. Sulub and G. W. Small, "Spectral simulation methodology for calibration transfer of near-infrared spectra," *Applied Spectroscopy*, vol. 61, no. 4, pp. 406–413, 2007.
- [24] L. Laghi, A. Versari, G. P. Parpinello, D. Y. Nakaji, and R. B. Boulton, "FTIR spectroscopy and direct orthogonal signal correction preprocessing applied to selected phenolic compounds in red wines," *Food Analytical Methods*, vol. 4, no. 4, pp. 619–625, 2011.
- [25] J. Zhang, Z. Zhang, Y. Xiang, Y. Dai, and P. B. Harrington, "An emphatic orthogonal signal correction-support vector machine method for the classification of tissue sections of endometrial carcinoma by near infrared spectroscopy," *Talanta*, vol. 83, no. 5, pp. 1401–1409, 2011.
- [26] J. Yoon, B. Lee, and C. Han, "Calibration transfer of near-infrared spectra based on compression of wavelet coefficients," *Chemometrics and Intelligent Laboratory Systems*, vol. 64, no. 1, pp. 1–14, 2002.
- [27] C. Tan and M. Li, "Calibration transfer between two near-infrared spectrometers based on a wavelet packet transform," *Analytical Sciences*, vol. 23, no. 2, pp. 201–206, 2007.
- [28] J. Bin, X. Li, W. Fan, J.-H. Zhou, and C.-W. Wang, "Calibration transfer of near-infrared spectroscopy by canonical correlation analysis coupled with wavelet transform," *Analyst*, vol. 142, no. 12, pp. 2229–2238, 2017.
- [29] W. Fan, Y. Liang, D. Yuan, and J. Wang, "Calibration model transfer for near-infrared spectra based on canonical correlation analysis," *Analytica Chimica Acta*, vol. 623, no. 1, pp. 22–29, 2008.
- [30] H. Li, J.-X. Wang, Z.-N. Xing et al., "Influence of improve Kennard/Stone algorithm on the calibration transfer in

- near-infrared spectroscopy," *Spectroscopy and Spectral Analysis*, vol. 31, no. 2, pp. 362–365, 2011.
- [31] W. Zen, W. Bei, W. Hao et al., "Study on near infrared spectroscopy model transfer by using piecewise direct standardization at different temperature," *Journal of instrumental analysis*, vol. 27, no. 12, pp. 1330–1333, 2008.
- [32] X. Chen, F. Yan, and H. Chang, "Evaluation of piecewise direct standardization algorithm for near infrared quantitative model updating," *Chinese Journal of Analytical Chemistry*, no. 09, pp. 1307–1313, 2014.
- [33] W. Ju, L. Hua, X. Zhi et al., "Application of wavelet multi-scale piecewise direct standardization on near infrared analysis calibration," *Chinese Journal of Analytical Chemistry*, vol. 39, no. 06, pp. 846–850, 2011.
- [34] T. Gao, C. Xiao, Y. Hong et al., "Application of wavelet transform-piecewise direct standardization on the near infrared analysis model transfer Chinese," *Journal of Analytical Chemistry*, vol. 34, no. 07, pp. 927–932, 2006.
- [35] C. Da, L. Fan, and L. Qi, "Development of multi-scale modeling methods for calibration transfer in near infrared spectroscopy," *Nanotechnology and Precision Engineering*, vol. 15, no. 2, pp. 121–126, 2017.
- [36] R. A. V. Rossel and R. Webster, "Predicting soil properties from the Australian soil visible–near infrared spectroscopic database," *European Journal of Soil Science*, vol. 63, no. 6, pp. 848–860, 2012.
- [37] J. Wetterlind, S. Bo, and M. Söderström, "Increased sample point density in farm soil mapping by local calibration of visible and near infrared prediction models," *Geoderma*, vol. 156, no. 3–4, pp. 152–160, 2010.
- [38] L. Jiao, L. Xiao, and G. Xiao, "Transfer method among water content detection models for different breeds of pork by hyperspectral imaging technique," *Transactions of the Chinese Society of Agricultural Engineering*, vol. 30, no. 17, pp. 276–284, 2014.
- [39] J. Cooper, C. Larkin, and M. Abdelkader, "Virtual standard slope and bias calibration transfer of partial least squares jet fuel property models to multiple near infrared spectroscopy instruments," *Journal of Near Infrared Spectroscopy*, vol. 19, no. 2, pp. 139–150, 2011.
- [40] Q. Chen, J. Zhao, M. Liu et al., "Determination of total polyphenols content in green tea using FT-NIR spectroscopy and different PLS algorithms," *Journal of Pharmaceutical and Biomedical Analysis*, vol. 46, no. 3, pp. 568–73, 2008.
- [41] L. J. Janik, S. T. Forrester, and A. Rawson, "The prediction of soil chemical and physical properties from mid-infrared spectroscopy and combined partial least-squares regression and neural networks (PLS-NN) analysis," *Chemometrics and Intelligent Laboratory Systems*, vol. 97, no. 2, pp. 179–188, 2009.
- [42] B. Kuang and A. M. Mouazen, "Non-biased prediction of soil organic carbon and total nitrogen with vis–NIR spectroscopy, as affected by soil moisture content and texture," *Biosystems Engineering*, vol. 114, no. 3, pp. 249–258, 2013.