

Non-Destructive Testing of Mechanical Properties of Solid Wood Panel Based on Partial Least Squares Structural Equation Modeling Transfer Method

Dapeng Jiang,^a Yizhuo Zhang,^{a,*} and Chen Jinhao^b

Calibration transfer between near infrared (NIR) spectrometers is a subtle issue in the chemometrics and process industry. Similar instruments may generate strongly different spectral responses, and regression models developed on a first NIR system can rarely be used with spectra collected by a second apparatus. In this work, two novel methods based on Structural Equation Modeling (SEM), called Enhanced Feature Extraction Approaches for factor analysis (EFEA-FA) and Enhanced Feature Extraction Approaches for spectral space transformation (EFEA-SST), were proposed to perform calibration transfer between NIR spectrometers. They were applied to a NIR nondestructive testing model for solid wood panels mechanical properties. Four different standardization algorithms were evaluated for transferring solid wood panels quality databases between a portable NIRS (InGaAs)-array spectrometer (NIRquest512) and a HSI Camera (SPECIM FX17). The results showed that EFEA-SST yielded the best model evaluation metrics (R^2 and Root Mean Square Error of Prediction (RMSEP)) values for tensile strength (RMSEP=11.309, $R^2=0.865$) parameters, while EFEA-FA gave the best fit for flexural strength (RMSEP=10.653, $R^2=0.912$). These results suggest the potential of two novel quality parameters prediction methods based on spectral databases transferred between diverse NIRS spectrometers.

DOI: 10.15376/biores.18.2.3620-3641

Keywords: Solid wood panel; Calibration transfer; Near infrared; Hyperspectral; Mechanical property

Contact information: a: College of Computer Science and Artificial Intelligence, Changzhou University, 1 gehu middle Rd., Changzhou, 213164, China; b: College of Mechanical and Electrical Engineering, Northeast Forestry University, 26 Hexing Rd., Harbin, 150040, China;

* Corresponding author: nefuzyz@163.com

INTRODUCTION

The forest products industry utilizes near infrared spectroscopy (NIRS) for quantitative examination of solid wood panels, such as tensile strength and flexural strength. NIRS offers just the mean spectrum of a sample (Tuncer 2022), regardless of the scanned area of the sample. As a result of averaging the gathered spectra to produce a single spectrum, information regarding the spatial distribution of constituents inside the sample is lost. The development of NIR hyperspectral imaging (HSI), combines NIR spectroscopy with digital imaging (Lima *et al.* 2022; Yakubu *et al.* 2022). For each pixel in the imaging plane, hyperspectral pictures can be collected spanning the whole visible and NIR wavelength range of a material (Vidal and Pasquini 2021). Consequently, using this stack of wavelength images or spectral cube, the average intensity and local changes of the intensity pixels at each spectral image may be analyzed and used for pattern identification. However, expensive and specialized hardware is required to capture hyperspectral images

(Nakawajana *et al.* 2021; Tunny *et al.* 2022). In general, hyperspectrometers with somewhat higher resolution cost over a million dollars. To solve this problem, calibration transfer technology strikes a balance between price and resolution, and near-infrared spectrum transfer technology was used to transfer the calibration model from the high-precision near-infrared spectrometer equipment platform in the laboratory to the low-precision hyperspectral equipment (Wang *et al.* 2023). A high-precision detection model derived from a laboratory NIR spectrometer was transferred to an industrial-grade online HSI spectrometer to improve the model's accuracy and reduce the overhead of the industrial pipeline.

Several standardisation approaches address this crucial issue and permit the transfer of calibration models (X. Li *et al.* 2021). There are methods for resolving transfer problems that do not require standardization (multiplicative scatter correction (MSC), orthogonal signal correction (OSC), *etc.*) (Shan *et al.* 2020). However, when the issue is not caused by spectral intensity variations but rather by wavelength shifts, various standardization approaches might be utilized. According to Qiao *et al.* (2021), a calibration transfer can be performed in a few ways: *a priori* correction involves correcting the spectra prior to applying the existing calibration model; model correction involves adapting the calibration model; and *a posteriori* correction involves correcting the predictions of the existing calibration model. Some modern transfer techniques are based on factor analysis, which separates spectral information related elements from noise to enhance transfer outcomes.

In the framework of transfer between instruments, *a priori* correction and model correction are based on multivariate spectrum correction. In the first mode, secondary spectra are matched to primary spectra and entered the existing model. In the second mode, the spectra of the primary spectra database are adjusted to match those of the secondary database, and the model is recalibrated. Spectra multivariate correction may use a large number of techniques, such as direct standardisation (DS) (Tian *et al.* 2022), or piecewise direct standardisation (PDS) (Sun *et al.* 2021; Chen *et al.* 2022). In *a posteriori* correction, existing primary spectra are applied to secondary spectra whose responses are known. After calibrating a model of the prediction error, its inverse is used to make future predictions. Typically, a simple univariate technique, such as bias/slope correction (BSC) of the projected values (Salguero-Chaparro *et al.* 2013), is used to implement this model.

Factor analysis (FA) is a highly effective method for establishing relationships between two sets of measurements (spectrum of two instruments). Transfer methods based on factor analysis, such as the Spectral Space Transformation algorithm (SST) (L. Li *et al.* 2022; Du *et al.* 2011), the alternating trilinear decomposition (ATLD) algorithm (Yap *et al.* 2022), the Principal Component Analysis (PCA) algorithm (Rehman *et al.* 2022), and the Canonical Correlation Analysis (CCA) algorithm (Fan *et al.* 2008; Zheng *et al.* 2014), have been widely applied and are frequently compared to traditional methods. FA, in contrast to PCR or PLS (Mendoza *et al.* 2018), utilizes correlation rather than covariance. A substantial covariance between two instruments NIR spectrum is not always indicative of a significant connection. A pair of those spectrum may have perfect correlation but low covariance. In such instances, correlation (FA) should be utilized rather than covariance (PCR or PLS).

PCA algorithm, CCA algorithm, and ATLD algorithm are statistically classified as exploratory factor analysis (EFA) approaches. In addition, the succession of peaks and troughs that show in an NIR spectrum represent the sample's molecular vibrations. The relationship between molecular overtones and combined bands in the NIR spectrum and the stretching vibrations of hydrocarbon bonds in solid wood panels is linear. By

comparing the NIR spectra of wood and pure chemical components such as lignin, cellulose, extractives, and xylan, researchers estimate the contribution of the chemical components to the mechanical performance of wood, in accordance with the premise of Confirmatory Factor Analysis (CFA) (Dash and Paul 2021). CFA assesses *a priori* hypotheses derived from earlier research on the link between the mechanical characteristics of solid wood panels and conducts validated testing on proposed models.

Lignin content and the lignin and extractives are strongly correlated with the absorbances near 1668 nm and 1684 nm (Horvath *et al.* 2011). The microfibril angle of wood samples were correlated with the absorbances near 1,150 nm. The absorbance peaks between 1075 and 1250 nm had been identified to be related to the lignin content (Watanabe *et al.* 2012). The band at 1143 nm belonged to aromatic groups, and the absorbances near 1130 nm were closely related to the lignin content after 2-d preprocess. Fujimoto *et al.* (2012) believed that the absorbances near 1428.5 nm were highly correlated with cellulose, the bands at 1366 nm, 1400 nm, and 1428 nm were associated with the cellulose and the moisture content, the band at 1396.6 nm was highly correlated with freewater, and the band at 1366.1 nm was highly correlated with freewater. The bands are highly correlated with bound water. Therefore, it is feasible to hypothesize, in advance, the number of factors affected the mechanical properties of wood, whether these factors are correlated, and which band in NIR load onto and reflect which factors, according to the above literature.

CFA is utilized by Partial Least Squares Structural Equation Modeling (PLS-SEM) (Sarstedt *et al.* 2022; Smith *et al.* 2022) to evaluate the measurement model. It differs from EFA in that it evaluates using empirical data with existing factor specification. For CFA, model fit is examined to validate the measurement. After the model is fitted, the route models between the latent variables are evaluated (Jonckere and Rosseel 2022). The PLS-SEM approach is intriguing because it allows them to estimate complicated models with numerous constructs, indicator variables, and structural routes without putting distributional assumptions on the data (Hair *et al.* 2020). PLS-SEM is a causal-predictive approach to SEM that stresses prediction in estimating statistical models, the structures of which are aimed to give causal explanations (Hult *et al.* 2021). Thus, the method avoids the seeming conflict between explanation and prediction, which is normally stressed in academic research. In addition, user-friendly software packages are available, such as PLS-Graph (Tenenhaus and Hanafi 2010) and SmartPLS (Sarstedt and Cheah 2019), that need minimum technical knowledge of the method.

In addition to introducing the CFA theory into the NIR transfer model, the purpose of this study is to present a unified calibration transfer framework based on the PLS-SEM model. Based on this framework, two existing algorithms, CCA and SST, were unified into a new process, named EFEA-FA and EFEA-SST. This work introduces Structural Equation Modeling in statistics and improves model transfer methods from a novel perspective. The band is selected according to the fundamentals of NIR spectroscopy. Using EFEA-FA and EFEA-SST for the prediction of the Solid Wood Panels mechanical character could reach great performance over both source and target domain. This work determined the number of factors and the relationship between factors and spectral bands. These components were chosen in accordance with knowledge from previous research as well as the PLS-SEM model's unified structure. The goal was to improve the interpretability of the transfer model.

EXPERIMENTAL

Materials

Logs went through the initial wood working process, for instance peeling or debarking. Next, wood was cut and processed into solid wood panels. A set of 90 Sheets of solid wood panels samples were processed. In the lab, the solid wood boards were kept at 5 °C and 90% relative humidity. Before spectroscopic measurements, the samples were equilibrated to room temperature (25 °C).

Spectra were collected from samples in reflectance mode (log 1/R) using two NIRS-instruments, and 2 batches of NIR samples were established: (1) portable NIRS (InGaAs)-array spectrometer (NIRquest512, Orlando, FL, USA) and (2) HSI Camera (SPECIM FX17, Oulu, Finland). The Smart PLS (3.2.8) statistical tool was used to examine the data through partial least square equation modeling (PLS-SEM), and MATLAB, a proprietary multi-paradigm programming language and numeric computing environment developed by MathWorks, was used to propose a novel model to achieve calibration transfer.

Figure 1 illustrates the HSI camera. A line-scan mode SPECIM FX17 provided with a transport module was used to measure all spectra in the scanning range of 900 to 1700 nm, at 8 nm intervals. A rectangular natural product cell with a window surface of 94.9 cm² was used. Each spectrum was the average of 32 scans. Spectra were captured using the CameraLink and GigE Vision Camera Interface in conjunction with the LUMO Software Development Kit (LUMO Development Kit, Oulu, Finland).

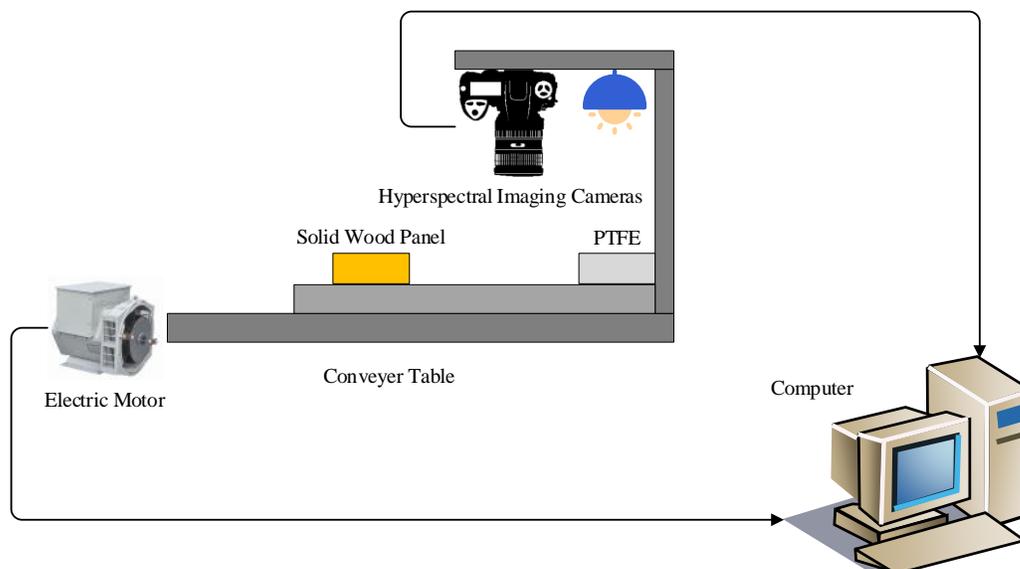


Fig. 1. SPECIM FX17 Camera

Figure 2 shows the NIRS Spectrometer. The second instrument was a NIRQuest512 spectrometer based on an array detector. The spectrometer with an InGaAs array work in the range from 900 to 1700 nm, with a spectral wavelength interval of 3.1 nm. The distance from the sample surface to the sensing head was approximately 13 mm. With an integration time of 5 s, 10 scans were averaged for each measurement. A total of 30 spectra of each sample were acquired, and the mean spectrum was used for data processing. All spectra were recorded using SpectraSuite 2.0 (Quadrangle Blvd, Orlando, FL, USA).

This study focused on the restoration and maintenance of ancient wooden structures, specifically the ancient bucket arch structures found in China. In analyzing these structures, it is equally important to consider both their ultimate strength index and their stiffness index. The study focused on the wood strength both under flexural and tension stresses. The mechanical testing material used in this study was birch wood grown in Northeast China, which was collected from the Chonghe Forest Farm of the Forestry Bureau of Wuchang City, Heilongjiang Province. The forest farm is characterized by continuous mountains, criss-crossing terrain, high elevation in the east and low elevation in the west, a dense river network, and abundant water sources. The geographical location and environmental conditions are detailed in Table 1.

Table 1. Climatic Conditions of Log Collection Areas

Latitude	Longitude	Average Altitude (m)	Max. Temp. (°C)	Minimum Temp. (°C)	Annual Average Temp. (°C)	Annual Precipitation (mm)	Annual Evaporation (mm)
44°37'~44°47'N	127°35'~127°55'E	350	35	-34	2.3	750	340

Within the birch forest, three groups of sample trees were selected based on their elevation, with each group containing four trees for a total of 12 sample trees. The trees were 20 years old, with a height of 12 to 14 m and a diameter at breast height of 16 to 17 cm. After marking the growth direction of each tree, the sample logs were felled, and the logs were cut at chest height (about 1.3 m above the ground). To highlight the differences in mechanical properties of the test materials, interval sampling was adopted during cutting to account for the distribution characteristics of mechanical properties in the vertical direction. Specifically, wood sections with a length of 1 m were cut for processing the bending test materials, while wood sections with a length of 2 m were cut for processing the tensile test materials, as shown in Fig. 2.

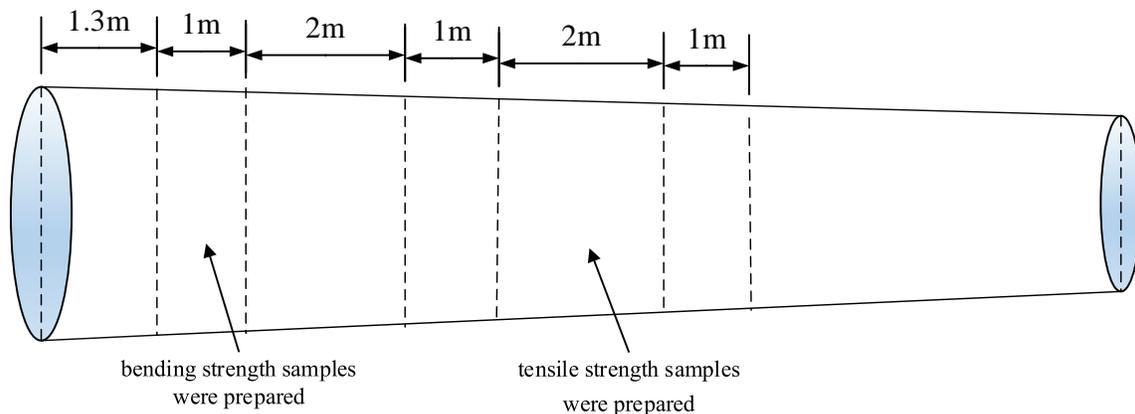


Fig. 2. Schematic diagram of wood cutting process

After air-drying, the wood section was sawn, excluding the pithwood and without distinguishing between the heartwood and sapwood. First, rough strips for bending resistance were made, and then bending strength values of samples of 300 mm × 20 mm × 20 mm were prepared in accordance with the Chinese National Standard (GB/T 1928-2009). Out of the total samples, 90 flawless samples were selected. These were numbered

from 1 to 90, and placed in a drying oven to reduce moisture content to 12%. To prevent moisture from affecting the samples, each one was placed in an airtight plastic bag. Reference values for flexural strength and tensile strength in solid wood panels samples were determined using official analysis methods. On solid wood panel samples, flexural strength was determined using the method of testing in bending strength of wood procedure (GB 1936.1-2009), which was proposed by the Standardization Administration of China (SAC, 2003). The tensile strength parallel to grain of wood (GB/T 1938.1-2009) was used to determine tensile strength.

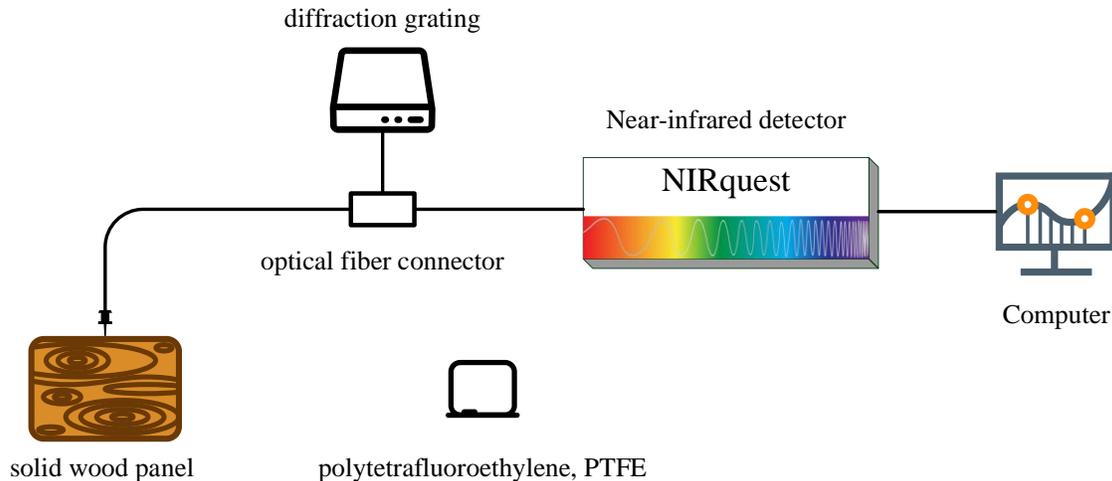


Fig. 3. NIR Quest 512 Spectrometer

Source instrument and target instrument matrices are used to denote the primary and secondary spectra. After preprocessing spectral data, 90 samples were obtained from each instrument. The source domain dataset X_{source} (90×512), target domain dataset X_{target} (90×224), and Y (90×2), were divided into three groups: transfer (X_t , Y_t), calibration (X_c , Y_c) and prediction (X_p , Y_p), in the following proportions: 60% (54 samples), 10% (9 samples) and 30% (27 samples), respectively. The transfer set, calibration set, and prediction set of source spectra are denoted by X_{st} , X_{sc} , and X_{sp} , while the analogous sets of target spectra are denoted by X_{tt} , X_{tc} , and X_{tp} . The parameter X_{new} represents the spectra corrected from X_{pt} by calibration transfer, whereas y_s and y_t represent the sample concentrations that correspond to the source and target spectra, respectively. In this article, the values of y_s and y_t are the same.

Calibration Transfer Based on CCA

X_{st} and X_{tt} can be executed by CCA:

$$P_t = T_{tt} \times X_{tt} \quad (1)$$

$$P_s = T_{st} \times X_{st} \quad (2)$$

where T_{st} and T_{tt} stand for the canonical weights, while P_s and P_t indicate the corresponding canonical scores of X_{st} and X_{tt} , respectively. The transfer matrix T was computed as follows,

$$F_1 = P_t \times P_s^+ \quad (3)$$

$$F_2 = X_{st} \times P_t^+ \quad (4)$$

$$F = F_2 \times F_1 \times T_{st} \quad (5)$$

where superscript “+” represents the pseudo-inverse, and both F_1 and F_2 are interim matrices for calculating F . Next, the X_{new} corrected from X_{sp} can be obtained by right multiplying F .

$$X_{\text{new}} = F \times X_{\text{sp}} \quad (6)$$

Finally, substituting B_{new} into the calibration model of A_c can yield the predicted values directly.

EFEA-FA Modeling

Generalizing: The PLS-SEM modeling and relationship with the wood biological function

PLS-SEM is a statistical technique that bears some resemblance to principal components analysis; however, instead of finding hyper-planes of maximum variance between the response and independent variables, it includes a method for assessing measurement model quality known as CFA. Researchers have begun referring to the measurement model assessment stage in PLS-SEM as CCA in recent years (Henseler *et al.* 2014; Schubert *et al.* 2018). CCA is a methodical procedure for systematically validating measurement models in PLS-SEM.

When conducting CCA with formative composite measurement models, the formative composite measurement models are linear combinations of the construct's indicators. The indicators are deemed causal and do not always co-vary, pointing from the measured variables to the composite concept. Therefore, the internal consistency principles underlying reflective measurement models cannot be applied to formative measurement models.

Due to its oriented fibers, wood has been viewed as an anisotropic material. The tensile strength of wood has a high correlation with the fiber angle, whereas the tensile strength of wood along the grain direction is more dependent on the fiber angle and the strength of bundles of molecular chains that combine in groups to form the cellulose fibres; the flexural strength of wood is largely a function of lignin. Lignin, a crucial structural component in the supporting tissues of most plants, is a reliable predictor of flexural strength and stiffness.

When the SEM is based on secondary data, a reflective or formative secondary measure of the same construct should be identified and used as a proxy variable for assessing the convergent validity of formative constructs. It is possible to identify acceptable endogenous reflectively assessed items for use as proxy variables in convergent validity testing by analyzing established scales from prior research. We chose three variables for this study: lignocellulose, moisture content, and lignin. This study selected as observation variables from the 900 nm to 1700 nm spectral bands 1668 nm to 1684 nm, 1075 nm to 1250 nm, 1070 nm to 1150 nm, 1366 nm, 1400 nm, and 1428 nm, based on previous research on the link between NIR bands and wood strength. Utilizing three variables as latent variable factors. The path dependence between latent variable factors (lignin, etc.) and observable variables (spectral bands) is established based on past research findings, and a PLS-SEM structural equation model is created.

EFEA-FA modeling steps

The EFEA-FA model is a new regression model designed by the PLS-SEM structural. As a statistical method that bears some resemblance to principal components analysis, rather than finding hyper-planes of maximum variance between the Mechanical

properties of wood and the near infrared wavelength spectrum, PLS-SEM computes latent variables from linear combinations of sets of specific wavelengths that correspond to previous research results to represent the Concept of wood Biological Functions. Based on the PLS-SEM calculated factors or composites, we developed the NIR spectral factor analysis transfer model.

The measuring model and the structural model make up SEM. A measurement model measures latent variables or composite variables, but a structural model examines all conceivable dependencies using path analysis. This is how the measurement model is expressed:

$$y = \Lambda_y \eta + \varepsilon \quad (7)$$

$$\xi = \Lambda_x x + \delta \quad (8)$$

Where y is the dependent variable, for this study, y is the measured value of Mechanical properties for wood; Where x is the independent variable, for this study, x is the measured value of NIR for wood; Λ_y and Λ_x are PLS-SEM loadings matrices; δ and ε are errors; η and ξ are PLS-SEM factors; the structural model is expressed as:

$$\eta = B\eta + \Gamma\xi + \zeta \quad (9)$$

In Eq. 9, B and Γ are relationship matrices; ζ is residual matrix. According to approximate estimation and iterative optimization of EM method, the optimal solution of PLS-SEM relationship matrix and factors is obtained. The simplified flow chart of the model is as follows:

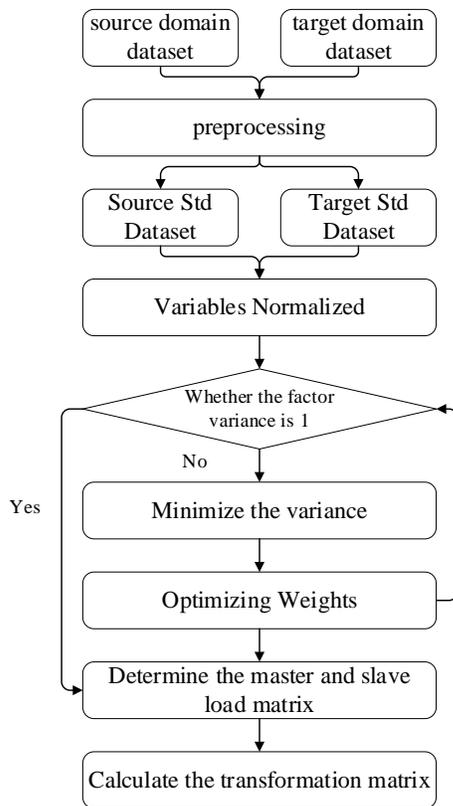


Fig. 4. PLS-SEM Framework Flowchart

Step 1: Establishing the PLS-SEM path model of wood mechanical properties.

Step 2: Iterating PLS-SEM model to optimize the relationship matrix, latent variables and parameters, the algorithm including the apparent variables normalized, the loading matrix approximation estimation, the relationship matrix approximation estimation and weight estimation.

The apparent variables normalized

The NIR spectral band and the mechanical properties of solid wood panels are normalized:

$$E(X_{tt}) = E(X_{st}) = E(y) = 0 \quad (10)$$

The loading matrix approximation estimation

Estimated Loading Matrix:

$$\xi_i^{t+1} = f_i^{t+1} \sum (w_{ij}^t \square x_{ij}) \quad (11)$$

$$\eta_i^{t+1} = g_i^{t+1} \sum (w_{ij}^{*t} \square y_{ij}) \quad (12)$$

Minimize the variance of ξ_i and η_i , let: $Var(\xi_i) = Var(\eta_i) = 1$

The relationship matrix approximation estimation

Approximate Estimation of Relationship Matrix

$$\xi_i^{t+1} = f_i^{*t+1} \sum (\theta_{ia}^{t+1} \square \xi_a^{t+1} + \lambda_{ia}^{t+1} \square \eta_a^{t+1}) \quad (13)$$

$$\xi_i^{t+1} = g_i^{*t+1} \sum (\theta_{ia}^{*t+1} \square \xi_a^{t+1} + \lambda_{ia}^{*t+1} \square \eta_a^{t+1}) \quad (14)$$

θ_{ia}^{t+1} , λ_{ia}^{t+1} , λ_{ia}^{*t+1} and θ_{ia}^{*t+1} is factor relationship matrix:

Minimize the variance of ξ_i and η_i , let: $Var(\xi_i) = Var(\eta_i) = 1$

Weight estimation

Iterative results of load matrix and score matrix for PLS-SEM

$$y_{ij} = w_{ij}^{*t+1} \eta_i^{t+1} + \xi_{y,ij}^{t+1} \quad (15)$$

$$\xi_1^{t+1} = \sum (w_{1j}^{t+1} \square x_{1j}) + \delta_{\xi,1}^{t+1} = (w_{11}^{t+1}, w_{12}^{t+1}, \dots, w_{1n}^{t+1}) \begin{pmatrix} x_{11} \\ x_{12} \\ \vdots \\ x_{1n} \end{pmatrix} + \delta_{\xi,1}^{t+1} \quad (16)$$

The augmented matrix between primary and secondary is X_{comb} , The above formula is transformed into:

$$X_{comb} = \begin{pmatrix} x_{11} \\ x_{12} \\ \vdots \\ x_{1n} \end{pmatrix} = (w_{11}^{t+1}, w_{12}^{t+1}, \dots, w_{1n}^{t+1})^+ (\xi_1^{t+1} - \delta_{\xi,1}^{t+1}) \quad (17)$$

The loading matrices P_t and P_s are expressed as:

$$(P_t, P_s) = (w_{11}^{t+1}, w_{12}^{t+1}, \dots, w_{1n}^{t+1})^+ \quad (18)$$

The score matrix T_{tt} of the host spectrum and the score matrix T_{st} of the secondary spectrum are expressed as:

$$T_{comb} = (T_{tt}, T_{st})^T = (\xi_1 - \delta_{\xi,1}) \quad (19)$$

Step 3: Calculate the conversion matrix T between T_{st} and T_{tt} and transfer the

spectral matrix X_{sp} from the machine to X_{new} . The formula is as follows:

$$F_1 = P_t \times P_s^+ \quad (20)$$

$$F_2 = X_{st} \times P_t^+ \quad (21)$$

$$F = F_2 \times F_1 \times T_{st} \quad (22)$$

Here, superscript “+” represents the pseudo-inverse, and both F_1 and F_2 are interim matrices for calculating F . Next, the X_{new} corrected from X_{sp} can be obtained by right multiplying F .

$$X_{new} = F \times X_{sp} \quad (23)$$

Finally, substituting X_{new} into the calibration model can yield the predicted values directly.

In summary, the algorithm block diagram of EFEA-FA is shown in Fig. 5:

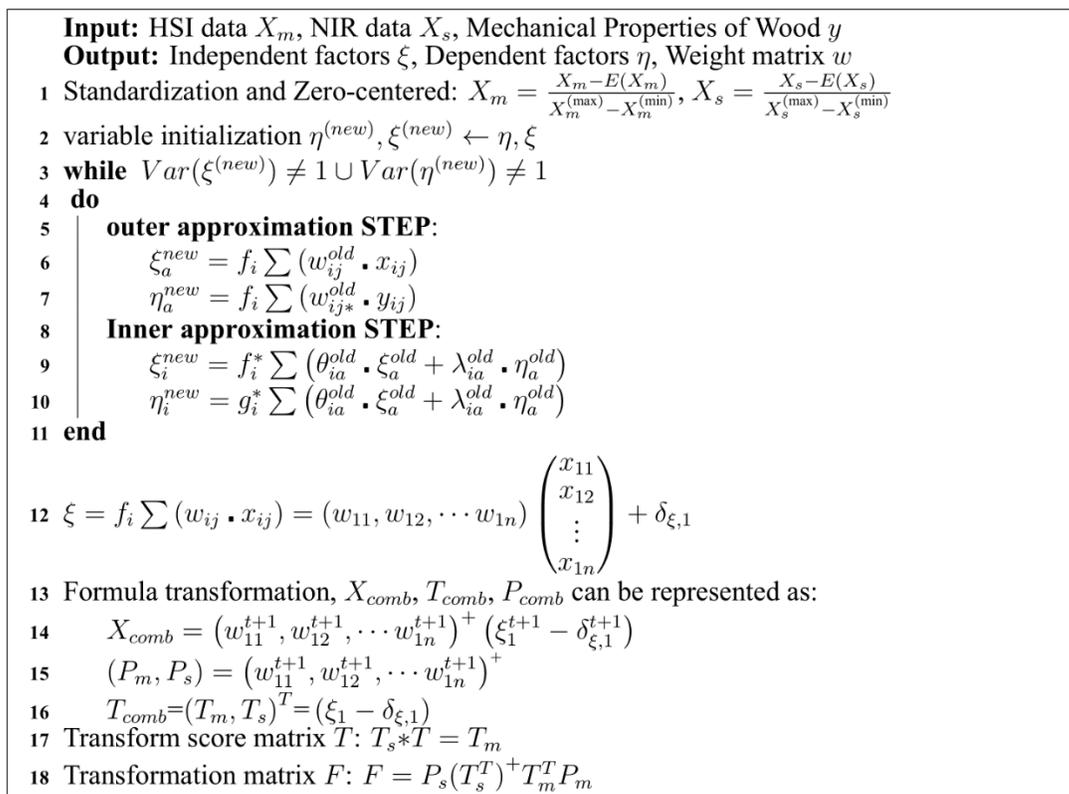


Fig. 5. EFEA-FA Algorithm

Calibration Transfer Based on SST

Assume the rows of the spectral matrices X_{it} and X_{st} are the corresponding spectra of the same subset of standardization samples measured on the primary and secondary instruments (or under the initial calibration and the modified test conditions), respectively. Let the singular value decomposition of X_{comb} as follows:

$$X_{comb} = [X_{it}, X_{st}] = TP = T[P_t^T, P_s^T] \quad (24)$$

Here, the scoring matrix is T , The primary load matrix is P_t , The secondary load matrix is P_s . The relationship between the primary matrix and the secondary matrix is:

$$X_{it} = X_{st} + X_{st} (P_s^T)^+ (P_t^T - P_s^T) \quad (25)$$

Arrange above formula, the translation matrix F can be calculated as follows:

$$F = I + (P_s^T)^+ (P_t^T - P_s^T) \quad (26)$$

translate the secondary spectrum X_{sp} to the primary spectrum X_{tnew} . the primary spectrum X_{tnew} can be obtained by:

$$X_{tnew} = F \times X_{sp} \quad (27)$$

EFEA-SST Modeling

The second model in our paper is EFEA-SST used for source domain data and target domain data respectively. As well as EFEA-FA, the EFEA-SST also implement calibration transfer based on the PLS-SEM framework. The specific algorithm flow is as follows:

Step 1: Establishing the PLS-SEM path model of wood mechanical properties.

Step 2: iterating PLS-SEM model to optimize the relationship matrix, latent variables and parameters, the algorithm including the apparent variables normalized, the loading matrix approximation estimation, the relationship matrix approximation estimation and weight estimation.

The apparent variables normalized

The NIR spectral band and the mechanical properties of solid wood panels are normalized:

$$E(X_{tt}) = E(X_{st}) = E(y) = 0 \quad (28)$$

The loading matrix approximation estimation

Estimated Loading Matrix:

$$\xi_i^{t+1} = f_i^{t+1} \sum (w_{ij}^t \square x_{ij}) \quad (29)$$

$$\eta_i^{t+1} = g_i^{t+1} \sum (w_{ij}^{*t} \square y_{ij}) \quad (30)$$

Minimize the variance of ξ_i and η_i , let: $Var(\xi_i) = Var(\eta_i) = 1$

The relationship matrix approximation estimation

Approximate Estimation of Relationship Matrix

$$\xi_i^{t+1} = f_i^{*t+1} \sum (\theta_{ia}^{t+1} \square \xi_a^{t+1} + \lambda_{ia}^{t+1} \square \eta_a^{t+1}) \quad (31)$$

$$\xi_i^{t+1} = g_i^{*t+1} \sum (\theta_{ia}^{*t+1} \square \xi_a^{t+1} + \lambda_{ia}^{*t+1} \square \eta_a^{t+1}) \quad (32)$$

θ_{ia}^{t+1} , λ_{ia}^{t+1} , λ_{ia}^{*t+1} and θ_{ia}^{*t+1} is factor relationship matrix:

Minimize the variance of ξ_i and η_i , let: $Var(\xi_i) = Var(\eta_i) = 1$

Weight estimation

Iterative results of load matrix and score matrix for PLS-SEM

$$y_{ij} = w_{ij}^{*t+1} \eta_i^{t+1} + \xi_{y,ij}^{t+1} \quad (33)$$

$$\xi_1^{t+1} = \sum (w_{1j}^{t+1} \square x_{1j}) + \delta_{\xi,1}^{t+1} = (w_{11}^{t+1}, w_{12}^{t+1}, \dots, w_{1n}^{t+1}) \begin{pmatrix} x_{11} \\ x_{12} \\ \vdots \\ x_{1n} \end{pmatrix} + \delta_{\xi,1}^{t+1} \quad (34)$$

The augmented matrix between primary and secondary is X_{comb} , The above formula is transformed into:

$$X_{comb} = \begin{pmatrix} x_{11} \\ x_{12} \\ \vdots \\ x_{1n} \end{pmatrix} = (w_{11}^{t+1}, w_{12}^{t+1}, \dots, w_{1n}^{t+1})^+ (\xi_1^{t+1} - \delta_{\xi,1}^{t+1}) \quad (35)$$

The loading matrices P_t and P_s are expressed as:

$$(P_t, P_s) = (w_{11}^{t+1}, w_{12}^{t+1}, \dots, w_{1n}^{t+1})^+ \quad (36)$$

The score matrix T_{tt} of the host spectrum and the score matrix T_{st} of the secondary spectrum are expressed as:

$$T_{comb} = (T_{tt}, T_{st})^T = (\xi_1 - \delta_{\xi,1}) \quad (37)$$

Step 3: Calculate the transformation matrix F . Next, the X_{new} corrected from X_{sp} can be obtained by right multiplying F .

$$F = I + (P_s^T)^+ (P_t^T - P_s^T) \quad (38)$$

$$X_{new} = F \times X_{sp} \quad (39)$$

In summary, the algorithm block diagram of EFEA-FA is shown in Fig. 6:

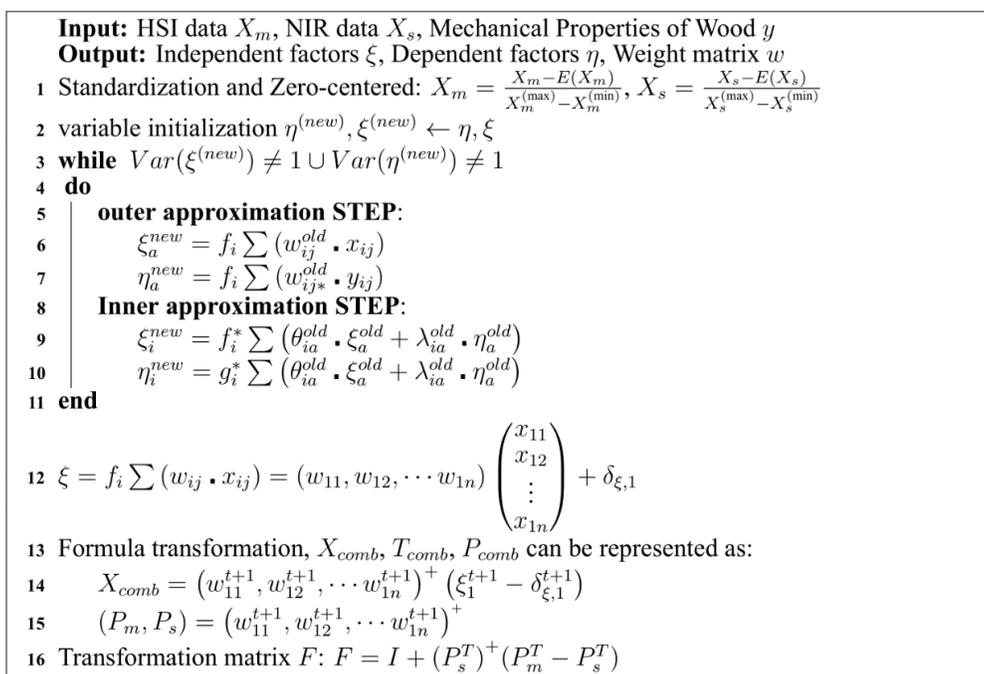


Fig. 6. EFEA- SST Algorithm

RESULTS AND DISCUSSION

NIR Spectral Preprocessing Results

The raw spectra from the original database as measured by the NIRquest512 (Fig. 2) and the SPECIM FX17 (Fig. 1) spectrometers are depicted in Fig. 7 and Fig. 8, respectively. Some major bands were observed on both figures. It is important that the light scattering has a significant impact on the spectra collected by both instruments. In fact, a significant proportion of photons that are not caught by the sensors and are so assumed to be absorbed are multiplicative scattered. After the log transform, this is additive scattered.

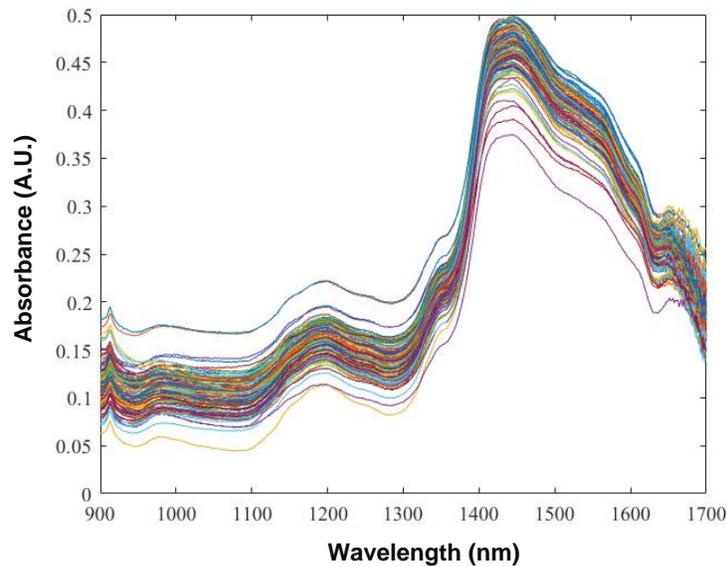


Fig. 7. Unprocessed primary spectra

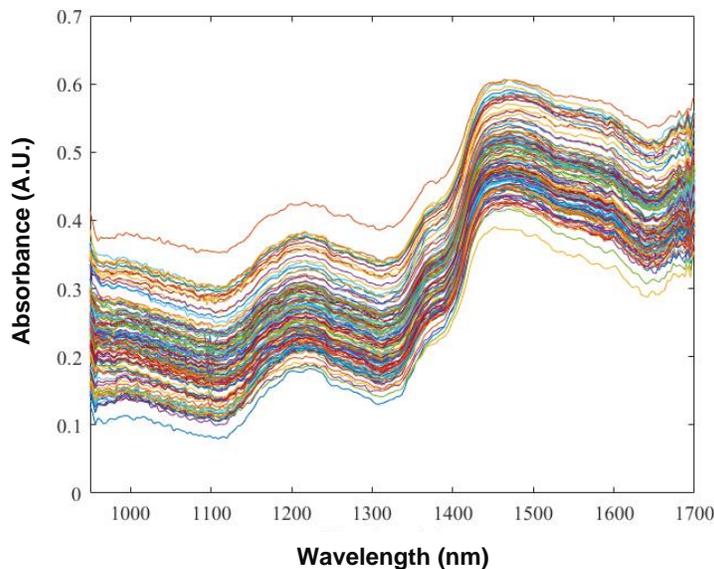


Fig. 8. Unprocessed secondary spectra

In this experiment, we used two classical spectrum preprocessing methods, standard normal variate (SNV) and Savitzky Golay Filter (S-G), to identify the spectral preprocessing approach with the greatest generalization capability. For each spectrum preprocessing approach, we first applied the identical spectral preprocessing to the source domain and the target domain, then used the processed source domain to create the calibration transfer models and the processed source domain and target domain to evaluate the performance. As depicted in Figs. 9-12, the spectra of each dataset were drawn following SNV-SG spectral preprocessing. Compared to Figs. 7, 8, The spectral gaps between the datasets were substantially decreased, and the spectral signal was noticeably smoothed.

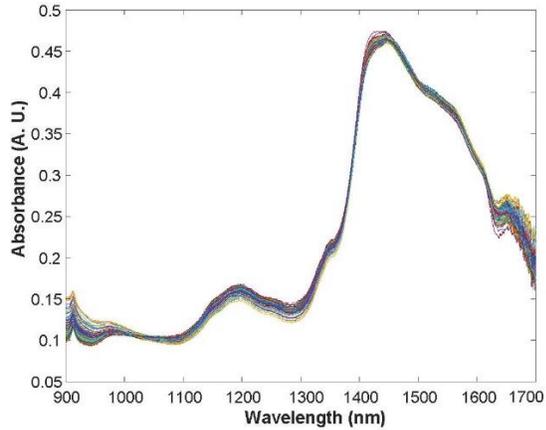


Fig. 9. Primary spectra pre-processed using the SNV method

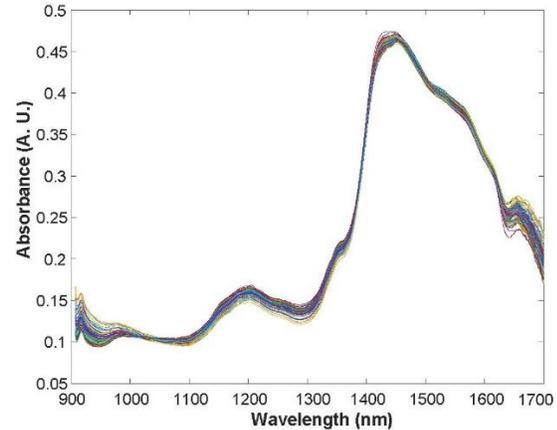


Fig. 10. Primary spectra pre-processed using the SNV-SG method

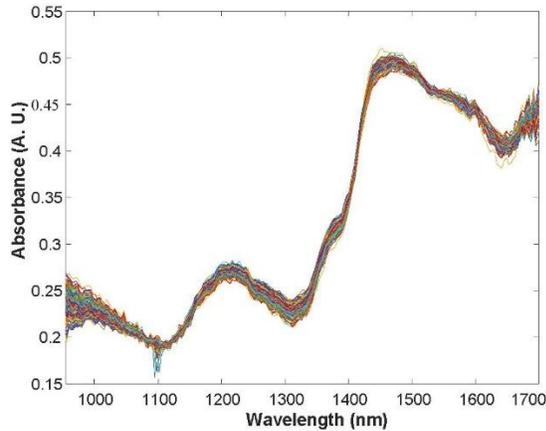


Fig. 11. Secondary spectra pre-processed using the SNV method

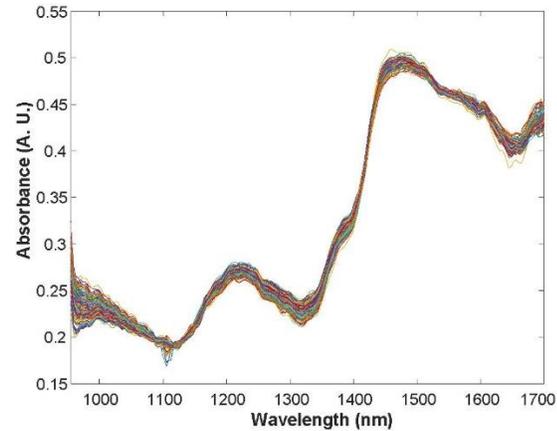


Fig. 12. Secondary spectra pre-processed using the SNV-SG method

PLS-SEM Modeling Process

Figures 13 and 14 depict statistically significant relationships between lignocellulose, lignin, moisture content and mechanical characteristics, with roundel symbols representing the pertinent latent variables. The inner model displayed non-normalized path coefficients, then the outer model displayed rescaled outer weights. Numbers along paths correspond to the weighted correlation coefficients that represent the strength of the relationship between two connected variables; overall impacts can be determined by multiplying path coefficients along one or more segments and summing across all feasible paths.

PLS-SEM was used to examine the scale's validity and reliability. The composite reliability (CR) and outer loading values were used to evaluate the scale's dependability. The average variance extracted (AVE) values were used to evaluate the convergent validity of this scale. Convergent validity assesses the degree of correlation between two measures of the same notion. The acceptance criteria for the reported values for the scale's validity and dependability are listed in Table 1. Values for composite reliability, which measure how well the construct indicators reveal the latent variable, should be higher than 0.70. More frequently, CR is expressed as follows:

$$CR = \frac{\left(\sum_i \lambda_i\right)^2}{\left(\sum_i \lambda_i\right)^2 + \sum_i e_i} \tag{40}$$

The AVE loading value higher than 0.50 was advised to justify the use of the construct as it measures the variance captured by the indicators relative to measurement error. More frequently, AVE is written as follows:

$$AVE = \frac{\sum_i (\lambda_i)^2}{\sum_i (\lambda_i)^2 + \sum_i e_i} \tag{41}$$

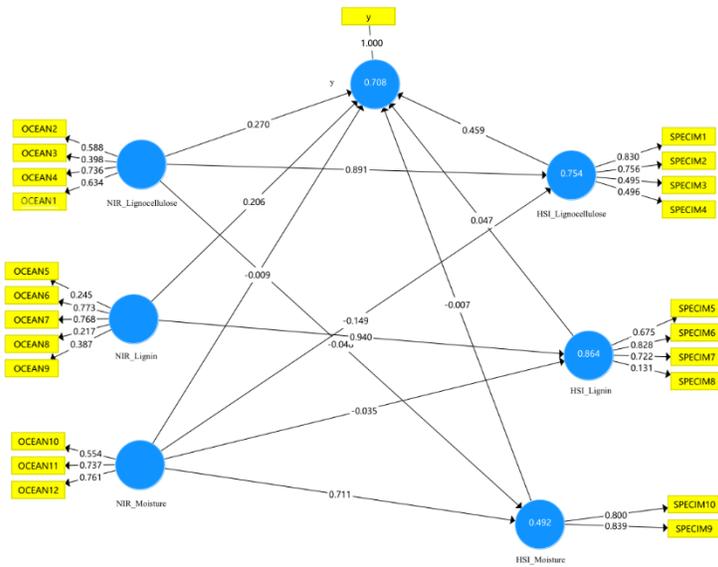


Fig. 13. Structural Equation path before optimization

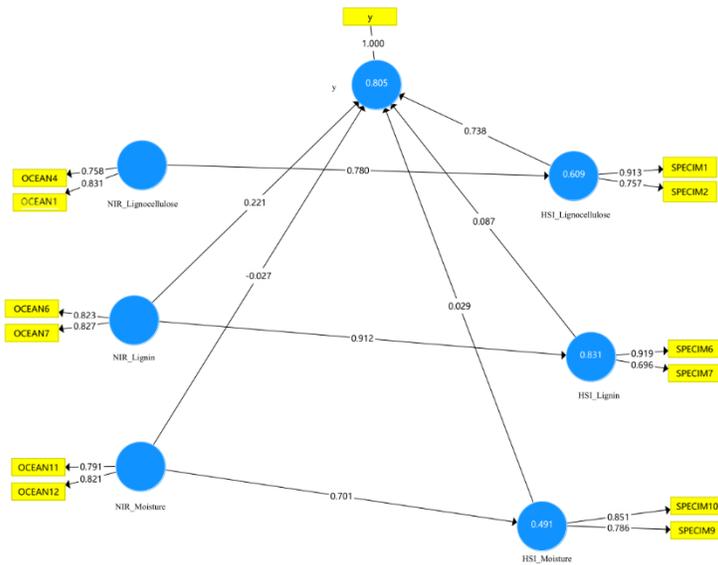


Fig. 14. Structural Equation path after optimization

The unoptimized fully connected structural equation is depicted in Fig. 13. Most of the composite reliability values in this structural equation model less than 0.8 and an AVE index less than 0.5. The structural equation model was poor quality. Spectral bands with poor SEM contribution values should be deleted, and the structural equation path should be made simpler. The optimized structural equation model is shown in Fig. 14. At this point, all the composite reliability values ranged from 0.774 to 0.825, indicating good internal consistency reliability; the AVEs ranged from 0.632 to 0.703, which were all within the recommended range as presented in Table 1. Therefore, the entire latent variables fulfilled the threshold value and met the standard recommended for convergent validity.

Table 1. Results of Structural Equation Index: Unoptimized vs Optimized

	Unoptimized		Optimized	
	Composite Reliability	AVE	Composite Reliability	AVE
Latent Variable 1	0.685	0.362	0.774	0.632
Latent Variable 2	0.747	0.438	0.825	0.703
Latent Variable 3	0.617	0.289	0.81	0.681
Latent Variable 4	0.728	0.476	0.787	0.649
Latent Variable 5	0.705	0.42	0.795	0.664
Latent Variable 6	0.803	0.672	0.803	0.671

The degree of correlation between the formative items and the indicators is known as multicollinearity. reflection indicators are frequently seen as interchangeable, which makes high correlations expected. However, multicollinearity issues are brought on by significant correlations between formative indicators. More frequently, VIF is written as:

$$VIF = \frac{1}{1 - R^2} \quad (42)$$

A multiple regression model is used to calculate formative construct scores, where the formative indicators are the independent variables and the construct scores are the dependent variable. The variance inflation factor (VIF), which is present in most statistical software, is investigated to ascertain whether multicollinearity is a problem. The VIF values of each factor are shown in Table 2, and multicollinearity is unlikely to be concerned if the VIF is 3.0 or less.

Table 2. VIF value of Each Factor

Factor	VIF
OCEAN11	1.098
OCEAN12	1.098
OCEAN2	1.042
OCEAN4	1.117
OCEAN6	1.151
OCEAN7	1.151
SPECIM1	1.22
SPECIM10	1.134
SPECIM2	1.22
SPECIM6	1.145
SPECIM7	1.145
SPECIM9	1.134
y	1
OCEAN1	1.077

Result of Wood Tensile Strength Transfer Model

Following SNV-SG pre-processing, transferable models based on the PLS-SEM framework from NIR to online HSI were constructed by employing several types of calibration transfer method for the prediction of the Solid Wood Panels Tensile Strength, and CCA and SST method were compared. The NIR-HSI spectra of the 9 calibration samples were subjected to each standardization procedure for spectrum correction. The prediction accuracy in terms of R^2 and RMSE was used to assess the effectiveness of the calibration transfer.

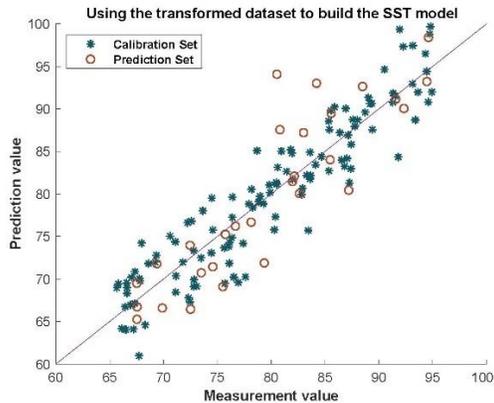


Fig. 15. Result of tensile strength testing using the SST calibration transfer model

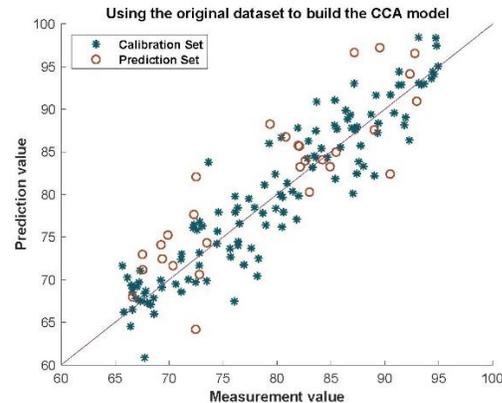


Fig. 16. Result of tensile strength testing using the CCA calibration transfer model

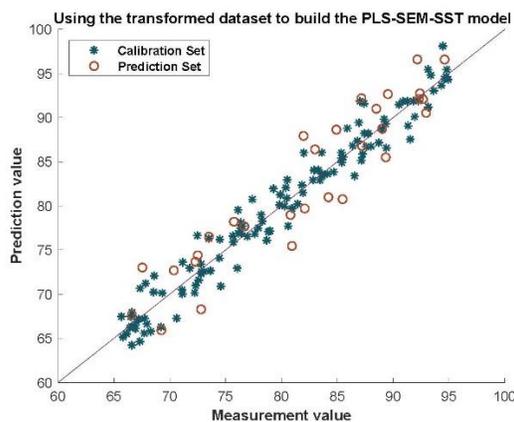


Fig. 17. Result of tensile strength testing using the EFEA-SST calibration transfer model

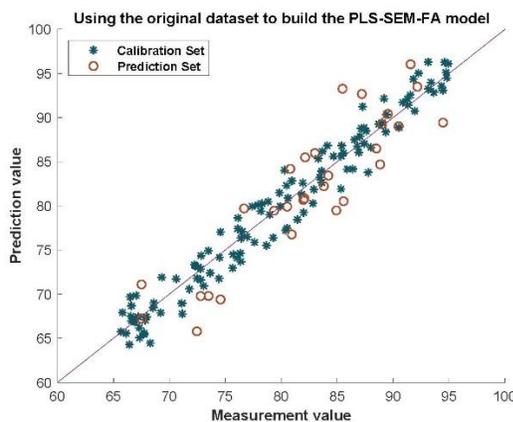


Fig. 18. Result of tensile strength testing using the EFEA-FA calibration transfer model

Figures 15 through 18 and Table 3 depict the results. Different types of calibration transfer methods have a major impact on performance. For example, by optimizing the CCA methods with the PLS-SEM framework, the correlation coefficient was enhanced from 0.792 to 0.813. The correlation coefficient was increased from 0.765 to 0.865 by optimizing the SST approaches with this framework. When compared to the results of the EFEA-FA approach of validation samples ($R_p=0.813$ and $RMSEP=16.646$, the indicator of EFEA-SST standardization ($R_p=0.865$ and $RMSEP=11.309$) had the best Flexural Strength prediction accuracy.

Table 3. Calibration Transfer Model Prediction Results of Timber Tensile Strength

	Training Set		Prediction Set	
	R_c	RMSEC	R_p	RMSEP
SST	0.859	12.809	0.765	21.790
CCA	0.865	11.404	0.792	24.810
EFEA-SST	0.958	3.337	0.865	11.309
EFEA-FA	0.960	3.316	0.813	16.646

Result of Wood Flexural Strength Transfer Model

Similar to how the tensile model is built, the spectral transfer model for wood bending is built. After SNV-SG pre-processing, transferable models based on the PLS-SEM framework from NIR to online HSI were built using a variety of calibration transfer methods to forecast flexural strength of Solid Wood Panels, and the CCA and SST methods were contrasted.

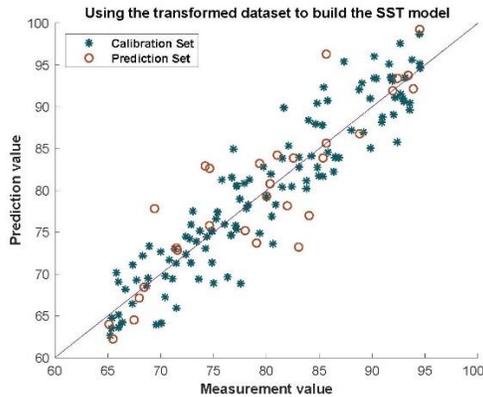


Fig. 19. Result of flexural strength testing using the SST calibration transfer model

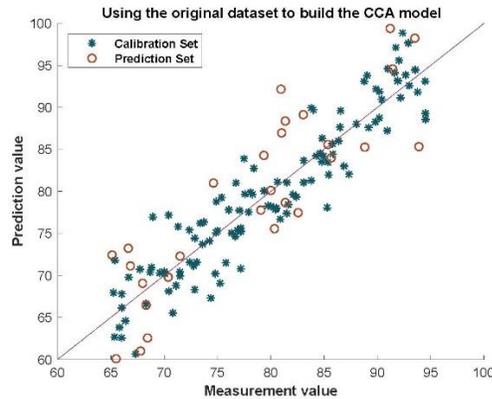


Fig. 20. Result of flexural strength testing using the CCA calibration transfer model

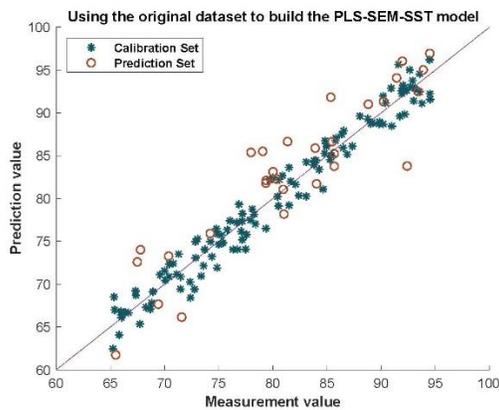


Fig. 21. Result of flexural strength testing using the EFEA-SST calibration transfer model

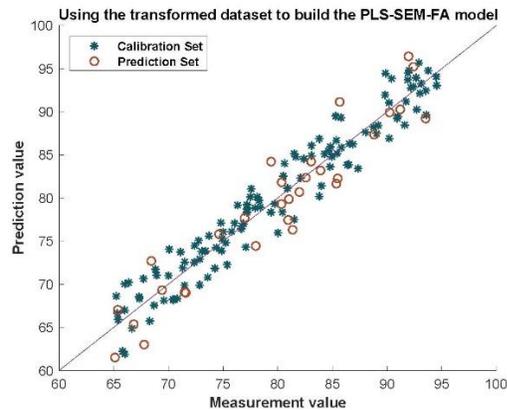


Fig. 22. Result of flexural strength testing using the EFEA-FA calibration transfer model

The outcomes are displayed in Figs. 19 through 22 and Table 4. The performance of various calibration transfer methods varies significantly. In particular, the correlation coefficient of the CCA approaches was improved from 0.795 to 0.912 with the application of the PLS-SEM framework. The SST techniques' correlation coefficient was improved using this framework, going from 0.778 to 0.843. When compared to the results obtained

using the EFEA-SST technique of validation samples ($R_p=0.843$ and $RMSEP =15.178$), the EFEA-FA standardization approach produced an equivalent Flexural Strength prediction accuracy of $R_p =0.912$ and $RMSEP =10.653$.

Table 4. Calibration Transfer Model Prediction Results of Timber Tensile Strength

	Training Set		Prediction Set	
	R_c	RMSEC	R_p	RMSEP
SST	0.869	11.607	0.778	22.024
CCA	0.872	11.985	0.795	25.847
EFEA-SST	0.964	2.792	0.843	15.178
EFEA-FA	0.939	4.752	0.912	10.653

CONCLUSIONS

1. The current NIRS calibrations model predicting flexural and tensile strength in solid wood panels was successfully transferred using all four transfer methods (EFEA-FA, EFEA-SST, CCA, and SST). The innovative method built on the PLS-SEM framework transfer demonstrated the best calibration transfer performance.
2. The best calibration transfer performance was found when the EFEA-FA procedure was applied predicting flexural strength. Note that the RMSEP value after EFEA-FA standardization was significantly lower (10.653) than the error obtained in the original calibration ($RMSEP=25.847$). Additionally, the R^2 values increased from 0.795 to 0.912.
3. In the same way, the application of EFEA-SST algorithm allowed to obtain the best prediction results for the Tensile Strength property. The RMSEP value obtained (Table 3) indicated that the prediction of Tensile Strength was greatly improved using EFEA-SST transferred rather than untransferred models. The prediction performed gave a RMSEP of 11.309, and a R^2 of 0.865.

ACKNOWLEDGMENTS

This research was supported of the Fundamental Research Funds for the China State Forestry Administration “948” projects (2015-4-52), Heilongjiang Natural Science Foundation (C2017005).

REFERENCES CITED

- Chen, Y., Wu, H. L., Wang, T., Liu, B. B., Ding, Y. J., and Yu, R. Q. (2022). “Piecewise direct standardization assisted with second-order calibration methods to solve signal instability in high-performance liquid chromatography-diode array detection systems,” *Journal of Chromatography A* 1667, article 462851. DOI: 10.1016/j.chroma.2022.462851

- Dash, G., and Paul, J. (2021). "CB-SEM vs PLS-SEM methods for research in social sciences and technology forecasting," *Technological Forecasting and Social Change* 173, article 121092. DOI: 10.1016/j.techfore.2021.121092
- Du, W., Chen, Z. P., Zhong, L. J., Wang, S. X., Yu, R. Q., Nordon, A., Littlejohn, D., and Holden, M. (2011). "Maintaining the predictive abilities of multivariate calibration models by spectral space transformation," *Analytica Chimica Acta* 690(1), 64-70. DOI: 10.1016/j.aca.2011.02.014
- Fan, W., Liang, Y., Yuan, D., and Wang, J. (2008). "Calibration model transfer for near-infrared spectra based on canonical correlation analysis," *Analytica Chimica Acta* 623(1), 22-29. DOI: 10.1016/j.aca.2008.05.072
- Fujimoto, T., Kobori, H., and Tsuchikawa, S. (2012). "Prediction of wood density independently of moisture conditions using near infrared spectroscopy," *Journal of Near Infrared Spectroscopy* 20(3), 353-359. DOI: 10.1255/jnirs.994
- GB/T 1928-2009 (2009). "General requirements for physical and mechanical tests of wood," The state forestry administration of the People's Republic of China, Beijing, China.
- GB/T 1938.1-2009 (2009). "Method of testing in tensile strength parallel to grain of wood," Anhui Agricultural University, Anhui, China.
- GB 1936.1-2009 (2009). "Method of testing in bending strength of wood," Chinese Academy of Forestry, Beijing, China.
- Hair, J. F., Howard, M. C., and Nitzl, C. (2020). "Assessing measurement model quality in PLS-SEM using confirmatory composite analysis," *Journal of Business Research* 109, 101-110. DOI: 10.1016/j.jbusres.2019.11.069
- Hair, J. F., Hult, G. T. M., Ringle, C. M., Sarstedt, M., Danks, N. P., and Ray, S. (2021). *Partial Least Squares Structural Equation Modeling (PLS-SEM) Using R*, Springer, Berlin. DOI: 10.1007/978-3-030-80519-7
- Henseler, J., Dijkstra, T. K., Sarstedt, M., Ringle, C. M., Diamantopoulos, A., Straub, D. W., Ketchen, D. J., Hair, J. F., Hult, G. T. M., and Calantone, R. J. (2014). "Common beliefs and reality about PLS: Comments on Rönkkö and Evermann (2013)," *Organizational Research Methods* 17(2), 182-209. DOI: 10.1177/1094428114526928
- Horvath, L., Peszlen, I., Peralta, P., and Kelley, S. (2011). "Use of transmittance near-infrared spectroscopy to predict the mechanical properties of 1-and 2-year-old transgenic aspen," *Wood Science and Technology* 45(2), 303-314. DOI: 10.1007/s00226-010-0330-x
- Jonckere, J. D., and Rosseel, Y. (2022). "Using bounded estimation to avoid nonconvergence in small sample structural equation modeling," *Structural Equation Modeling: A Multidisciplinary Journal* 29(3), 412-427. DOI: 10.1080/10705511.2021.1982716
- Li, L., Huang, W., Wang, Z., Liu, S., He, X., and Fan, S. (2022). "Calibration transfer between developed portable Vis/NIR devices for detection of soluble solids contents in apple," *Postharvest Biology and Technology* 183, article 111720. DOI: 10.1016/j.postharvbio.2021.111720
- Li, X., Li, Z., Yang, X., and He, Y. (2021). "Boosting the generalization ability of Vis-NIR-spectroscopy-based regression models through dimension reduction and transfer learning," *Computers and Electronics in Agriculture* 186, article 106157. DOI: 10.1016/j.compag.2021.106157
- Lima, M. D. R., Guedes Ramalho, F. M., Trugilho, P. F., Bufalino, L., Francisco Dias Júnior, A., de Paula Protásio, T., and Hein, P. R. G. (2022). "Classifying waste wood

- from Amazonian species by near-infrared spectroscopy (NIRS) to improve charcoal production,” *Renewable Energy* 193, 584-594. DOI: 10.1016/j.renene.2022.05.048
- Mendoza, F. A., Cichy, K. A., Sprague, C., Goffnett, A., Lu, R., and Kelly, J. D. (2018). “Prediction of canned black bean texture (*Phaseolus vulgaris* L.) from intact dry seeds using visible/near infrared spectroscopy and hyperspectral imaging data,” *Journal of the Science of Food and Agriculture* 98 (1), 283-290. DOI: 10.1002/jsfa.8469
- Nakawajana, N., Lerdwattanakitti, P., Saechua, W., Posom, J., Khwantri, S., and Wongpichet, S. (2021). “A low-cost system for moisture content detection of bagasse upon a conveyor belt with multispectral image and various machine learning methods,” *Processes* 9(5). DOI: 10.3390/pr9050777
- Qiao, L., Mu, Y., Lu, B., and Tang, X. (2021). “Calibration maintenance application of near-infrared spectrometric model in food analysis,” *Food Reviews International* 2021, 1-17. DOI: 10.1080/87559129.2021.1935999.
- Rehman, T. U., Zhang, L., Ma, D., and Jin, J. (2022). “Common latent space exploration for calibration transfer across hyperspectral imaging-based phenotyping systems,” *Remote Sensing* 14(2). DOI: 10.3390/rs14020319
- Salguero-Chaparro, L., Palagos, B., Peña-Rodríguez, F., and Roger, J. M. (2013). “Calibration transfer of intact olive NIR spectra between a pre-dispersive instrument and a portable spectrometer,” *Computers and Electronics in Agriculture* 96, 202-208. DOI: 10.1016/j.compag.2013.05.007
- Sarstedt, M., and Cheah, J. H. (2019). “Partial least squares structural equation modeling using SmartPLS: A software review,” *Journal of Marketing Analytics* 7(3), 196-202. DOI: 10.1057/s41270-019-00058-3
- Sarstedt, M., Hair, J. F., Pick, M., Liengaard, B. D., Radomir, L., and Ringle, C. M. (2022). “Progress in partial least squares structural equation modeling use in marketing research in the last decade,” *Psychology & Marketing* 39(5), 1035-1064. DOI: 10.1002/mar.21640
- Schuberth, F., Henseler, J., and Dijkstra, T. K. (2018). “Confirmatory composite analysis,” *Frontiers in Psychology* 9. DOI: 10.3389/fpsyg.2018.02541.
- Shan, P., Zhao, Y., Wang, Q., Ying, Y., and Peng, S. (2020). “Principal component analysis or kernel principal component analysis based joint spectral subspace method for calibration transfer,” *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* 227, article 117653. DOI: 10.1016/j.saa.2019.117653
- Smith, L. K., Nixon, R. S., Sudweeks, R. R., and Larsen R. A. (2022). “Elementary teacher characteristics, experiences, and science subject matter knowledge: Understanding the relationships through structural equation modeling,” *Teaching and Teacher Education* 113, article 103661. DOI: 10.1016/j.tate.2022.103661
- Sun, X., Zhang, P., Shen, Z., Yuan, H., Song, C., Han, X., Xiu, Y., Lv, Z., and Hu, A. (2021). “Investigation on spectral standardization among multi-channel of an on-line near-infrared spectrometer,” *Vibrational Spectroscopy* 113, article 103206. DOI: 10.1016/j.vibspec.2021.103206
- Tenenhaus, M., and Hanafi, M. (2010). “A bridge between PLS path modeling and multi-block data analysis,” In: *Handbook of Least Partial Squares*, V. E. Vinzi, W. W. Chin, J. Henseler, and H. Wang (eds.), Springer, Berlin, pp. 99-123. DOI: 10.1007/978-3-540-32827-8_5

- Tian, J., Chen, X., Liang, Z., Qi, W., Zheng, X., Lu, D., and Chen, B. (2022). "Application of NIR spectral standardization based on principal component score evaluation in wheat flour crude protein model sharing," *Journal of Food Quality* 2022, 9009756. DOI: 10.1155/2022/9009756
- Tuncer, F. D. (2022). "Comparing modelling performance of chemometric methods for wood discrimination by near infrared spectroscopy," *Wood Material Science & Engineering* 2022, 1-12. DOI: 10.1080/17480272.2022.2039960
- Tunny, S. S., H Amanah, H. Z., Faqeerzada, M. A., Wakholi, C., Kim, M. S., Baek, I., and Cho, B. K. (2022). "Multispectral wavebands selection for the detection of potential foreign materials in fresh-cut vegetables," *Sensors* 22(5). DOI: 10.3390/s22051775
- Vidal, C., and Pasquini, C. (2021). "A Comprehensive and fast microplastics identification based on near-infrared hyperspectral imaging (HSI-NIR) and chemometrics," *Environmental Pollution* 285, article 117251. DOI: 10.1016/j.envpol.2021.117251
- Wang, Y., Ren, Z., Li, M., Lu, C., Deng, W. W., Zhang, Z., and Ning, J. (2023). "From lab to factory: A calibration transfer strategy from HSI to online NIR optimized for quality control of green tea fixation," *Journal of Food Engineering* 339, article 111284. DOI: 10.1016/j.jfoodeng.2022.111284
- Watanabe, K., Yamashita, K., and Noshiro, S. (2012). "Non-destructive evaluation of surface longitudinal growth strain on sugi (*Cryptomeria japonica*) green logs using near-infrared spectroscopy," *Journal of Wood Science* 58 (3), 267-272. DOI: 10.1007/s10086-011-1238-2
- Yakubu, H. G., Kovacs, Z., Toth, T., and Bazar, G. (2022). "The recent advances of near-infrared spectroscopy in dairy production—A review," *Critical Reviews in Food Science and Nutrition* 62 (3), 810-831. DOI: 10.1080/10408398.2020.1829540
- Yap, X. Y., Chia, K. S., and Suarin, N. A. S. (2022). "Adaptive artificial neural network in near infrared spectroscopy for standard-free calibration transfer," *Chemometrics and Intelligent Laboratory Systems* 230, article 104674. DOI: 10.1016/j.chemolab.2022.104674
- Zheng, K., Zhang, X., Iqbal, J., Fan, W., Wu, T., Du, Y., and Liang, Y. (2014). "Calibration transfer of near-infrared spectra for extraction of informative components from spectra with canonical correlation analysis," *Journal of Chemometrics* 28(10), 773-784. DOI: 10.1002/cem.2637

Article submitted: December 12, 2022; Peer review completed: February 11, 2023;
Revised version received: February 22, 2023; Accepted: February 23, 2023; Published:
April 6, 2023.

DOI: 10.15376/biores.18.2.3620-3641