Near-infrared spectroscopy as a potential method for identification of anatomically similar Japanese diploxylons

Yoshiki Horikawa*, Suyako (Mizuno) Tazuru, Junji Sugiyama

Research Institute for Sustainable Humanosphere, Kyoto University, Kyoto, Japan

*Tel: +81-774-38-3634

Fax: +81-774-38-3635

E-mail: yhorikawa@rish.kyoto-u.ac.jp

Keywords: Discriminant analysis; NIR spectroscopy; Japanese diploxylons; Wood identification; Aging wood

Abstract

A reliable technique for distinguishing anatomically similar diploxylons, *Pinus densiflora* and *P. thunbergii*, was designed by employing near-infrared (NIR) spectroscopy in combination with multivariate analysis. In total, 24 wood blocks, with half of them being of *P. densiflora* and the rest of *P. thunbergii*, were selected from the collections of the Kyoto University xylarium and scrutinized to build an acceptable model for discriminating between the two species. The prediction model was constructed only from heartwood, and the best performance was obtained for wavenumbers of 7300–4000 cm^{-1} in the second derivative spectra. To apply this model to actual materials obtained from historical wooden buildings, 12 aging wood samples were analyzed and compared by microscopic identification. Unexpectedly, the spectral differences between the species were smaller than those caused by aging, and the prediction error was approximately 50%. The spectra of the aging samples were quite distinct in the specific region characteristic of absorbed water (5220 cm^{-1}); this was demonstrated clearly by principal component analysis. Therefore, for the proposed model to be suitable for use in practical applications, further investigations of aging wood samples and the corresponding spectroscopic data are necessary in order to understand the effects of aging on the spectral data.

Introduction

Pinus densiflora and *P. thunbergii* are varieties of pine trees that are very popular in Japan. The former is known as akamatsu and mematsu, and the latter as kuromatsu and omatsu. Both are planted widely in Japan for timber production and as ornamental trees and are a characteristic feature of classical Japanese gardens. *P. densiflora* is commonly seen growing on the low mountains and hillsides, while *P. thunbergii* is native to the coastal areas.

Anatomically, the two species are nearly identical in terms of the resin canal, which is surrounded by thin-walled epithelial cells and window-type cross-field pitting and exhibits a distinct transition from earlywood to latewood. The key difference between the two species was reported to be the degree of dentate thickening of the ray tracheids (Fig. 1). However, this difference is rather subjective and can be misleading, particularly in old samples, whose cell walls have nearly deteriorated. Consequently, in many previous studies, these pine wood species have been identified simply as diploxylons and their particular species has been left undecided. Therefore, an alternative method that could allow for the identification of the particular species without requiring special experimentation would be highly desirable.

In this regard, near-infrared (NIR) spectroscopy, which is known as a rapid, accurate and reproducible analysis technique, is an attractive choice. NIR spectroscopy is also suitable for assessing wood materials because the bands attributable to the vibrations of the chemical bonds involved in the formation of the cell wall allow for the direct and indirect estimation of

the chemical and physical properties of the materials. When combined with multivariate analysis, NIR spectroscopy can be used to distinguish between different wood species. Schimleck et al. demonstrated that principal component analysis (PCA) could be used to distinguish between pine and eucalyptus and also to differentiate between samples of the same eucalyptus species grown at different sites [1]. Soft independent modeling of class analogy has also been used to classify wood samples, including red and white oak [2] and larch species [3]. Regression analysis, especially partial least square (PLS) regression, is a powerful method of accurately estimating the chemical compositions of wood samples [4] and of determining their enzymatic hydrolysis [5] and decay resistance [6], in addition to their physical properties, such as fiber length [7-9], cellulose microfibril angle [10], and stiffness [11]. The method of distinguishing species coupled with regression analysis, called partial least squares-discriminant analysis (PLS-DA), has been used as a tool for differentiating true mahogany from three other similar species [12, 13]. Sandberg and Sterley [14] could successfully distinguish between heartwood and sapwood samples of Norway spruce using the PLS algorithm. Watanabe et al. [15] could differentiate between aging and degraded samples of softwood, such as *Chamaecyparis obtusa*, *Torreya nucifera*, and *C. pisifera* using PLS.

In this study, we first describe a simple technique that uses NIR spectroscopy in combination with PLS-DA for distinguishing *P. densiflora* from *P. thunbergii*, which were

classified at the xylarium of Kyoto University. The discriminant model was initially examined using complete sets of the wood samples. Later, the sapwood and heartwood samples were analyzed separately. Next, we demonstrate the applicability of the proposed method in determining the species of aging samples of wood, and discuss the factors that influence the precision of discrimination.

Materials and methods

Sampling

Wood blocks of *P. densiflora* designated as KYOw00029, 00225, 08058, 08059, 09268, 13942, 19360, and 19361, and those of *P. thunbergii* designated as KYOw00030, 00520, 05509, 05639, 08071, 10321, 11386, 13913, and 19176 by the xylarium at the Research Institute for Sustainable Humanosphere, Kyoto University (http://database.rish.kyoto-u.ac.jp/cgi-bin/bmi/en/namazu.cgi) were used for establishing the discriminant model. The wood samples in these blocks were collected from all the sapwood and heartwood zones. However, in the case of the wood blocks of KYOw00029, 05639, 08071, 09268, and 13913, only sapwood was collected. On the other hand, only heartwood was taken from KYOw00520, 19176, 19360, and 19361. Three parts were collected randomly from each wood block after NIR spectral analysis. Finally, wood samples from Chion-In temple in Kyoto, Japan and designated as KYO_ID_5165, 5166, 5168, 5170, 5173, 5175, 5185, 5187, 5189, 5192, 5197, and 5252 [16] were used to test applicability of the

proposed method. Blocks were collected from each of these aging samples.

Optical microscopy

In the case of the wood samples obtained from the xylarium for the construction of the calibration model, radial sections approximately 30 μm in thickness were cut using a sliding microtome and were stained with safranin. In the case of the wood samples from Chion-in, the corresponding sections were obtained by hand sectioning and were not stained. The sections were observed using a light microscope (Olympus BX51) equipped with a digital camera (Olympus DP73).

NIR spectroscopy

Each wood block after air-drying was milled with a rough file to produce a powder sample. Then, a tablet was prepared by collecting approximately 0.04 g of the powder, which was hand pressed following a previously published protocol [17]. The NIR spectrum was obtained using a PerkinElmer Spectrum 100N system for wavenumbers of 10000–4000 cm^{-1} at a spectral resolution of 16 cm^{-1}; 32 scans were made for each sample. The prepared tablet was placed directly on the NIR integrating sphere diffuse reflectance accessory (PerkinElmer), which had a triglycine sulfate detector. Both faces of each tablet were scanned. The absorbance spectrum was recorded by normalizing the single-beam spectrum against the background spectrum using a Teflon-based material (Spectralon; LabSphere, North Sutton, NH). The original spectrum was treated using the Savitzky–Golay second derivative [18]

using 9 points and a fifth-order polynomial for the smoothing before the multivariate analysis.

Multivariate analysis

PLS-DA and PCA were performed using a commercial software (Unscrambler v.9.8; CAMO Software, Inc., Woodbridge, NJ). Calibration and prediction samples from the 144 spectra (72 each for the sapwood and heartwood samples) were randomly selected as the ratio at 2 to1, that is, 96 were used for calibration, and 48 were used for the prediction set. Of the spectra used for calibration, 48 belonged to *P. densiflora* and 48 belonged to *P. thunbergii*. In the case of the spectra used for prediction, 24 belonged to *P. densiflora* and 24 belonged to *P. thunbergii*, as shown in Table 1. For the development of a discriminant model for use in the multivariate analysis, we assigned *P. densiflora* a class value of +1 and *P. thunbergii* a class value of -1 in the calibration set. The PLS factors were determined by cross validation; a single sample was kept out of the model, and its characteristics were predicted by constructing a model without the sample. Excessively high numbers may result in overfitting; therefore, the number of PLS factors was kept at fewer than 11. The coefficient of determination for calibration (R_c^2) and the root mean square error of calibration (RMSEC) were used to assess the calibration performance. The models developed were evaluated by using the coefficient of determination of prediction (R_p^2) and the root mean square error of prediction (RMSEP). The percentage of correct prediction was determined as the proportion

of the number of species discriminated correctly compared to the total number of samples from prediction set. PLS-DA to distinguish between sapwood and heartwood was also performed using the same procedure.

PCA was performed on the basis of the second derivative spectra of all the wood samples for wavenumbers of 7300–4000 cm^{-1}. The PC loading was obtained from the model built for score plots.

Results and Discussion

Discriminant model for determining the type of wood present

Fig. 2a and b show original and second derivative spectra from heart and sapwood samples of *P. densiflora* and *P. thunbergii*. From the band at 5220 cm^{-1} assigned to absorbed water in second derivative NIR spectra, sapwood samples seemed higher moisture contents than those of heartwood. However, it was difficult to identify whether *P. densiflora* or *P. thunbergii* from spectra because spectral pattern including the bands at 5980 and 5800 specific to lignin and hemicellulose respectively, were almost same between these species. Therefore, we applied multivariate analysis and Table 2 shows the statistical summary of the discriminant models obtained on the basis of the original spectra and the second derivative spectra. To generate a better model, the regions of the NIR spectra corresponding to the wavenumbers of 10000–4000 cm^{-1} were separated into four distinct ranges on the basis of the properties of the molecular vibrations. In the first range (10000–7300 cm^{-1}), the second or

152 third overtones were involved, although less information was obtained from the wood
153 samples. The second range (7300–6050 cm^{-1}) mainly corresponded to OH overtone vibrations.
154 The third range (6050–5500 cm^{-1}) corresponded to the CH vibrations and the vibrations from
155 the aromatic framework, while in the fourth range (5500–4000 cm^{-1}), several combinatorial
156 vibrations were present.

157 For the samples containing both sapwood and heartwood, the discriminant models
158 shown in Table 2a were constructed on the basis of the NIR spectra without subjecting the
159 spectra to any spectral pretreatment. All the models were unreliable because the R_p^2 values
160 were less than 0.60. Next, we obtained the second derivative spectra and created the
161 discriminant models shown in Table 2b. Secondary differentiation can extract information
162 hidden in the original spectra. Thus, researchers have often applied this algorithm to construct
163 regression models. This spectral pretreatment decreased the number of factors relatively;
164 however, the models obtained were not markedly better. Fig. 3a shows a histogram
165 corresponding to the discriminant model based on the second derivative spectra for 7300–
166 4000 cm^{-1}. In this region, a few samples of both *P. densiflora* and *P. thunbergii* had class
167 values of approximately 0, which indicated that this model could not be used for
168 distinguishing between the two species.

169 Sapwood could, therefore, be distinguished from heartwood, and discriminant
170 models could be built, as summarized in Tables 2c and d. However, as was the case with the

dataset corresponding to the samples containing both sapwood and heartwood, all the models showed poor performances, as the R_p^2 values were lower than 0.75. Fig. 3b shows a histogram based on the second derivative spectra for 7300–4000 cm^{-1}; for this region, the RMSEP value was 0.54 and the R_p^2 value was 0.71, with some of the prediction samples from *P. thunbergii* having class values of 0 and similar to those of *P. densiflora*.

In the case of heartwood, even though a large number of factors were required, the calibration performance was comparatively better (Table 2e). However, the models obtained using the vibrations over 10000–7300 cm^{-1} and 7300–6050 cm^{-1} were less reliable; this was particularly true in the latter case, where the major bands were assigned to cellulose [19-21]. This suggested that the cellulose contents were indistinct between *P. densiflora* and *P. thunbergii* as well as their crystalline properties.

The discriminant models obtained using the second derivative spectra are shown in Table 2f. The models exhibited better performances as the R_p^2 values corresponding to a few of the NIR spectral regions were higher than 0.85. The best performance was obtained for 7300–4000 cm^{-1}; this region showed an RMSEP value of 0.37, R_p^2 value of 0.86 and 100 % accuracy of identification. As shown in Fig. 3c, the prediction model based on the corresponding region allowed us to classify all the samples from *P. densiflora* as having positive values, while the samples for classifying *P. thunbergii* were placed in another group and had negative values.

NIR spectroscopy is sensitive to the functional groups and is thus influenced by the chemical and structural features of the cell walls of the trees being investigated. Therefore, the difficulties encountered in classification using sapwood samples indicated that the chemical natures of *P. densiflora* and *P. thunbergii* were essentially indistinct. However, the fact that using heartwood samples yielded better results suggested that the heartwood components of the two species might be slightly different.

Applicability in investigating aging wood used in traditional buildings

The applicability of the regression model developed was tested by reexamining actual wood materials used in traditional wooden buildings built in the medieval period. Chion-In temple in Kyoto is well known and is the main temple of Jōdo Shū ("The Pure Land School"). The wood materials used in the Shūedō (i.e., the Assembly Hall) had been studied during 2005–2010, and 25 wood samples had been classified as being of diploxylons [16]. Twelve specimens were used in the present study. The enlarged radial sections of the ray tracheids of these samples are shown in Fig. 4. Of these 12 samples, only three were anatomically identified as *P. thunbergii* on the basis of the degree of dentate thickening in the ray tracheids. The PLS-DA models built up on the basis of heartwood were used in the identification of these materials. The percentage of coincidence with the anatomical identification results is listed in Table 3. In contrast to the prediction set samples shown in Tables 2e and f, the discriminant models failed to predict the species perfectly. One possible

reason behind this failure seems to be the fact that the wood used in Chion-in was sapwood, while the calibration models used for identification were created using heartwood. However, sapwood is usually not used as a building material, in order to minimize deterioration and maintain the structural strength. Given this background, we investigated these aging wood samples further using NIR spectroscopy in combination with multivariate analysis, as mentioned in the next segment.

Effects of aging

To understand the reason for the failure in prediction of Chion-in materials, PCA was carried out in the wavenumber range 7300–4000 cm^{-1} of the second derivative spectra (Fig. 5a). The score plots showed that some of the wood samples from Chion-In temple localized on the left side and far from those belonging to *P. densiflora* and *P. thunbergii*. It is known that noncrystalline polysaccharides such as hemicellulose decrease in quantity in aging samples of *C. obtusa*, whereas the crystalline cellulose region is not affected [22]. Furthermore, Yokoyama et al. reported that the equilibrium moisture content in *C. obtusa* decreases after aging [23]. Therefore, it seems that aging under dry conditions degraded the hemicelluloses, which are the adsorption sites for water in wood materials, resulting in a decrease in the equilibrium moisture content. In this regard, the wood samples from Chion-in temple were different in that there was no statistical difference between modern and aging woods in hemicellulose contents, given the presence of the band at approximately 5800 cm^{-1};

this band is specific to furanose/pyranose, which form from hemicellulose [24] and exhibited a value of almost 0 in the PC1 loading (Fig. 5b). In addition, the amount of absorbed water in the Chion-in samples was higher, as a positive band was noticed at approximately 5220 cm^{-1} and was assignable to the combinational vibration of water; this was clearly visible in the PC1 loading. Moreover, the lignin content of the Chion-in samples was lower, as a band was noticed at 5970 cm^{-1}; this band is characteristic of aromatic skeletal vibrations [24] and exhibited negative values during PC1 loading. These interpretation was supported by the comparison with the second derivative spectra between modern and aging wood (Fig. 5c). Therefore, the lignin in the Chion-in samples seemed to be modified to a greater degree than was the hemicellulose, which resulted in a decrease in the hydrophobicity, as this increased the amount of absorbed water. These features were not observed in the spectra of the unaging wood samples. Hence, the samples from Chion-in could not be classified accurately. In order to be able to employ the proposed classification method for identifying historical and archeological wood samples, we have to consider the effects of aging on the characteristics of the samples, including on the quantity of absorbed water and the chemical components such as lignin and polysaccharides, whose chemical structure can be changed by oxidative and/or enzymatic reactions. Therefore, further investigations need to be performed to determine the optimal conditions for measurements as well as suitable data treatments to account for the spectral variations caused by aging, in order to be able to distinguish between *P. densiflora*

247 and *P. thunbergii* on the basis of the differences in their spectra.

248 **Conclusions**

249 When using unaging heartwood samples, we were able to identify *P. densiflora* and *P.*
250 *thunbergii* by employing NIR spectroscopy in combination with multivariate analysis.
251 However, when aging wood samples were used, the proposed method was ineffective in
252 distinguishing between the two species. Thus, the method is not suitable for classifying wood
253 samples from historical and archeological buildings. However, further research is underway
254 to find the spectral features between these microscopically similar species more significant
255 than those caused by aging.
256

257 **Acknowledgments**

258 The study was supported in parts by Grants-in-Aid for Scientific Research (Grant Numbers
259 25252033, 22300309, and 24780169) from the Japan Society for the Promotion of Science
260 (JSPS). The authors thank Ms. Izumi Kanai and Mr. Akio Adachi for their technical support.

261 **References**

262 [1] Schimleck LR, Evans R, Ilic J, Matheson AC (2002) Estimation of wood stiffness of
263 increment cores by near-infrared spectroscopy. Can J For Res 32: 129–135

264 [2] Adedipe OE, Dawson-Andoh B, Slahor J, Osborn L (2008) Classification of red oak
265 (*Quercus rubra*) and white oak (*Quercus alba*) wood using a near infrared spectrometer and
266 soft independent modelling of class analogies. J Near Infrared Spectrosc 16: 49–57

267 [3] Gierlinger N, Schwanninger M, Wimmer R (2004) Characteristics and classification
268 of Fourier-transform near infrared spectra of the heartwood of different larch species (*Larix*
269 sp.). J Near Infrared Spectrosc 12: 113–119
270 [4] Jones PD, Schimleck LR, Peter GF, Daniels RF, Clark A (2006) Nondestructive
271 estimation of wood chemical composition of sections of radial wood strips by diffuse
272 reflectance near infrared spectroscopy. Wood Sci Technol 40: 709–720
273 [5] Horikawa Y, Imai T, Takada R, Watanabe T, Takabe K, Kobayashi Y, Sugiyama J
274 (2012) Chemometric analysis with near-infrared spectroscopy for chemically pretreated
275 Erianthus toward efficient bioethanol production. Appl Biochem Biotechnol 166: 711–721
276 [6] Leinonen A, Harju AM, Venalainen M, Saranpaa P, Laakso T (2008) FT-NIR
277 spectroscopy in predicting the decay resistance related characteristics of solid Scots pine
278 (*Pinus sylvestris L.*) heartwood. Holzforschung 62: 284–288
279 [7] Hauksson JB, Bergqvist G, Bergsten U, Sjostrom M, Edlund U (2001) Prediction of
280 basic wood properties for Norway spruce. Interpretation of near infrared spectroscopy data
281 using partial least squares regression. Wood Sci Technol 35: 475–485
282 [8] Inagaki T, Schwanninger M, Kato R, Kurata Y, Thanapase W, Puthson P,
283 Tsuchikawa S (2010) *Eucalyptus camaldulensis* density and fiber length estimated by
284 near-infrared spectroscopy. Wood Sci Technol 46: 143–155
285 [9] Schimleck LR, Jones RD, Peter GF, Daniels RF, Clark A (2004) Nondestructive

286 estimation of tracheid length from sections of radial wood strips by near infrared
287 spectroscopy. Holzforschung 58: 375–381
288 [10] Schimleck LR, Evans R, Jones PD, Daniels RF, Peter GF, Clark A (2005)
289 Estimation of microfibril angle and stiffness by near infrared spectroscopy using sample sets
290 having limited wood density variation. IAWA J 26: 175–187
291 [11] Schimleck LR, Michell AJ, Vinden P (1996) Eucalypt wood classification by NIR
292 spectroscopy and principal components analysis. Appita J 49: 319–324
293 [12] Braga JWB, Pastore TCM, Coradin VTR, Camargos JAA, da Silva AR (2011) The
294 use of near infrared spectroscopy to identify solid wood specimens of *Swietenia Macrophylla*
295 (Cites Appendix II). IAWA J 32: 285–296
296 [13] Pastore TCM, Braga JWB, Coradin VTR, Magalhaes WLE, Okino EYA, Camargos
297 JAA, de Muniz GIB, Bressan OA, Davrieux F (2011) Near infrared spectroscopy (NIRS) as a
298 potential tool for monitoring trade of similar woods: Discrimination of true mahogany, cedar,
299 andiroba, and curupixa. Holzforschung 65: 73–80
300 [14] Sandberg K, Sterley M (2009) Separating Norway spruce heartwood and sapwood in
301 dried condition with near-infrared spectroscopy and multivariate data analysis. Eur J For Res
302 128: 475–481
303 [15] Watanabe K, Abe H, Kataoka Y, Nodhito S (2011) Species separation of aging and
304 degraded solid wood using near infrared spectroscopy. Jpn J Histor Bot 19: 117–124

[16] Mizuno S, Sugiyama J (2011) Wood identification of building components of Syue-do, Chion-in temple designated as nationally important cultural property. J Soc Architect Hits Japan 56: 124–135

[17] Horikawa Y, Imai T, Takada R, Watanabe T, Takabe K, Kobayashi Y, Sugiyama J (2011) Near-infrared chemometric approach to exhaustive analysis of rice straw pretreated for bioethanol conversion. Appl Biochem Biotechnol 164: 194–203

[18] Savitzky A, Golay MJE (1964) Smoothing and differentiation of data by simplified least squares procedures. Anal Chem 36: 1627–1639

[19] Tsuchikawa S, Siesler HW (2003a) Near-infrared spectroscopic monitoring of the diffusion process of deuterium-labeled molecules in wood. Part I: Softwood. Appl Spectrosc 57: 667–674

[20] Tsuchikawa S, Siesler HW (2003b) Near-infrared spectroscopic monitoring of the diffusion process of deuterium-labeled molecules in wood. Part II: Hardwood. Appl Spectrosc. 57: 675–681

[21] Watanabe A, Morita S, Ozaki Y (2006) A study on water adsorption onto microcrystalline cellulose by near-infrared spectroscopy with two-dimensional correlation spectroscopy and principal component analysis. Appl Spectrosc 60: 1054–1061

[22] Tsuchikawa S, Yonenobu H, Siesler HW (2005) Near-infrared spectroscopic observation of the ageing process in archaeological wood using a deuterium exchange

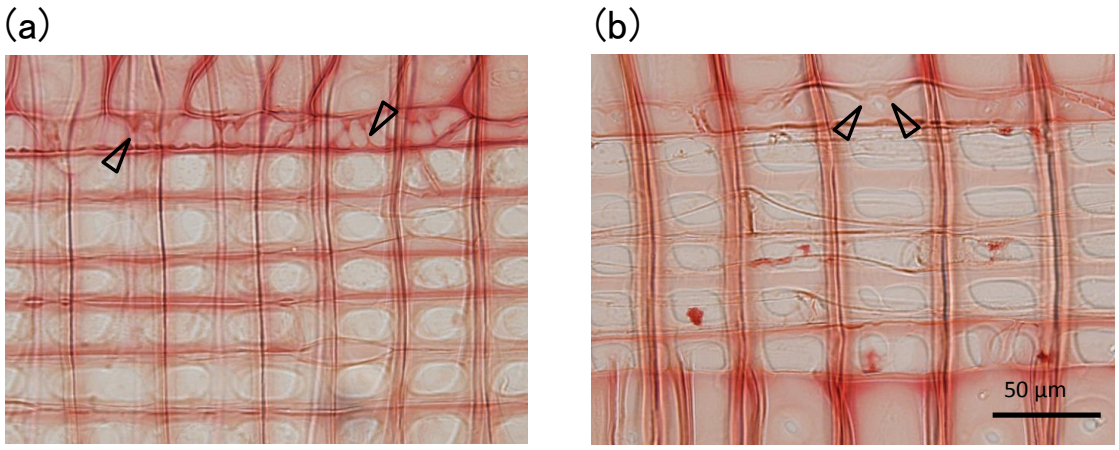
method. Analyst 130: 379–384

[23] Yokoyama M, Gril J, Matsuo M, Yano H, Sugiyama J, Clair B, Kubodera S, Mitsutani T, Sakamoto M, Ozaki H, Imamura M, Kawai S (2009) Mechanical characteristics of aged Hinoki wood from Japanese historical buildings. C R Physique 10: 601–611

[24] Siesler HW, Ozaki Y, Kawata S, Heise HW Near-infrared Spectroscopy, Principal Instruments, Applications; Willey-VCH Verlag GmbH, Weinhem, Germany, 2002

331 Figures

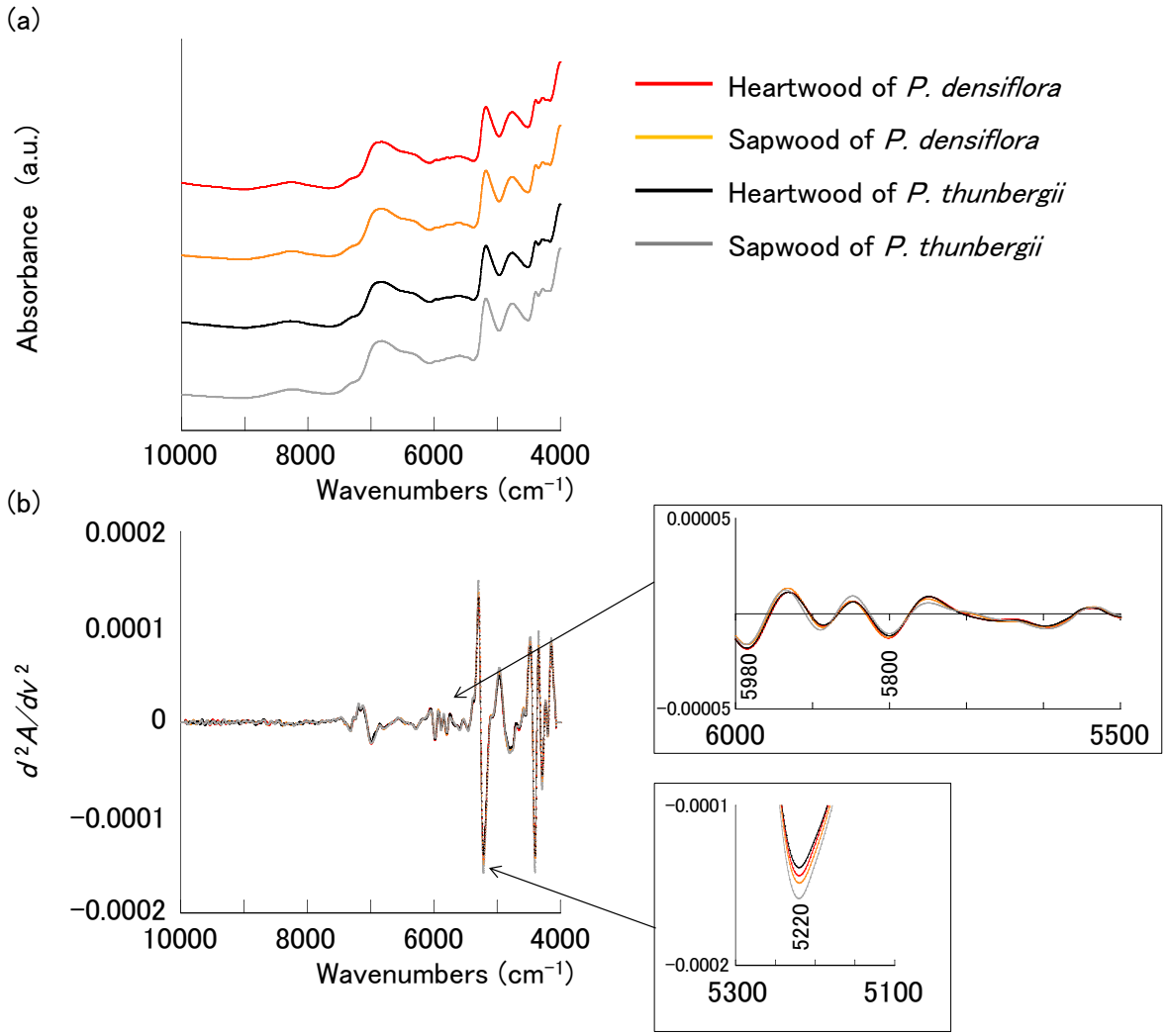
332



333

334 Fig. 1

335



336

337

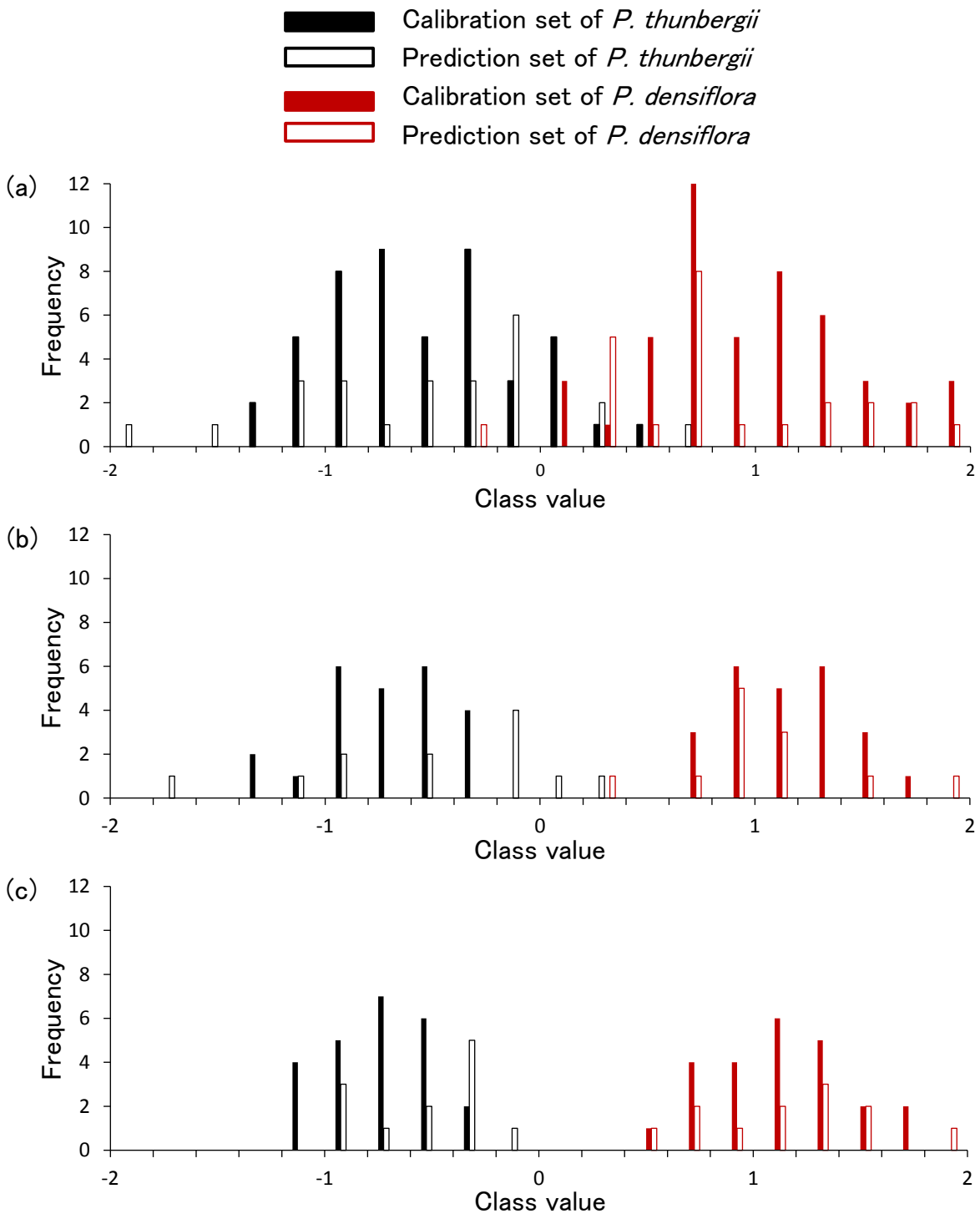
338 Fig. 2

339

340

341

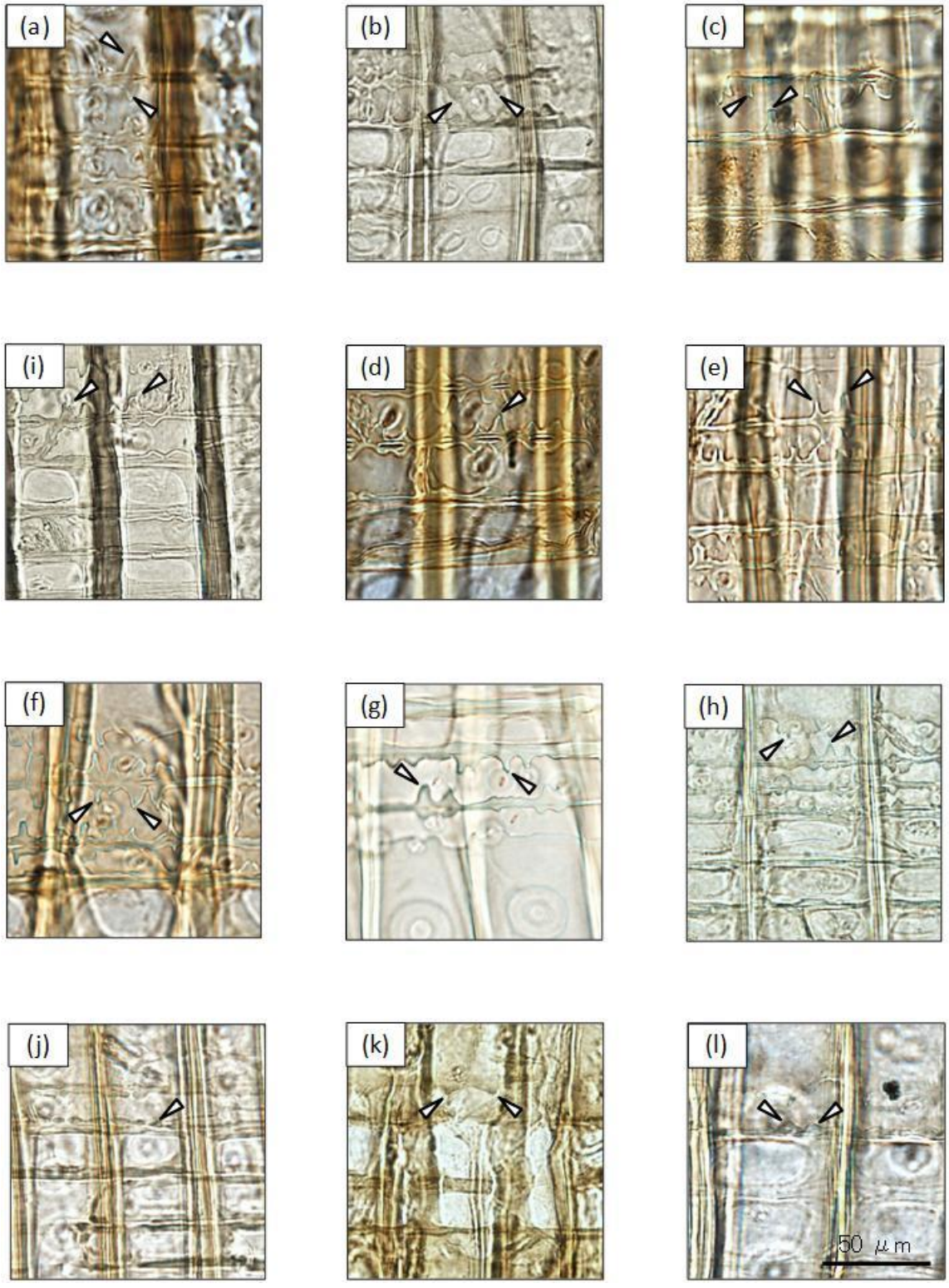
342



343
344

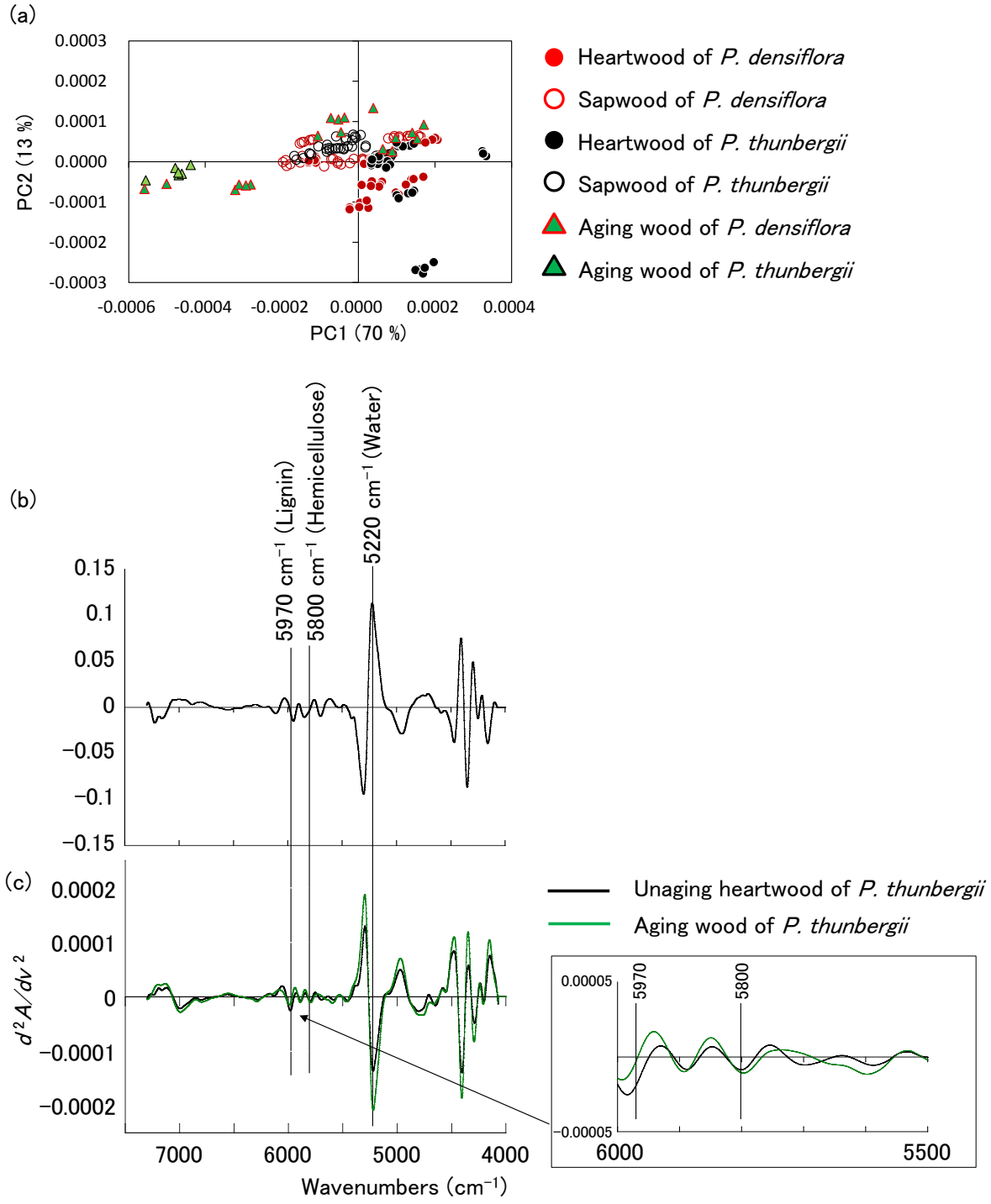
345 Fig. 3

346



347

348 Fig. 4



349

350 Fig. 5

351

352 **Figure legends**

353 Fig. 1. Optical micrographs of the standard radial section. (a) *P*. *densiflora* shows dentate
354 thickening within the ray tracheid, while (b) these features are smooth in *P. thunbergii*. The
355 arrow heads indicate dentate thickening.

356

357 Fig. 2. (a) Original NIR spectra of heartwood and sapwood from *P. densiflora* and *P*.
358 *thunbergii* designated as KYOw13942 and 00030, which were included in calibration set. (b)
359 Second derivative spectra obtained from the 4 spectra in (a).

360

361 Fig. 3. Histograms of the class values computed by PLS-DA on the basis of the second
362 derivative spectra for wavenumbers of 7300–4000 cm^{-1} and obtained from (a) a mixture of
363 sapwood and heartwood, (b) sapwood, and (c) heartwood.

364

365 Fig. 4. Optical micrographs of the radial sections acquired from wood samples from Chion-In
366 temple. The images in (a)–(i) correspond to KYO_ID_5165, 5166, 5168, 5170, 5173, 5175,
367 5185, 5189, and 5252, respectively, which were identified as being of *P. densiflora*. The
368 images in (j)–(l), which were acquired from KYO_ID_5187, 5192, and 5197, respectively,
369 were identified as being of *P. thunbergii*. The arrow heads indicate dentate thickening.

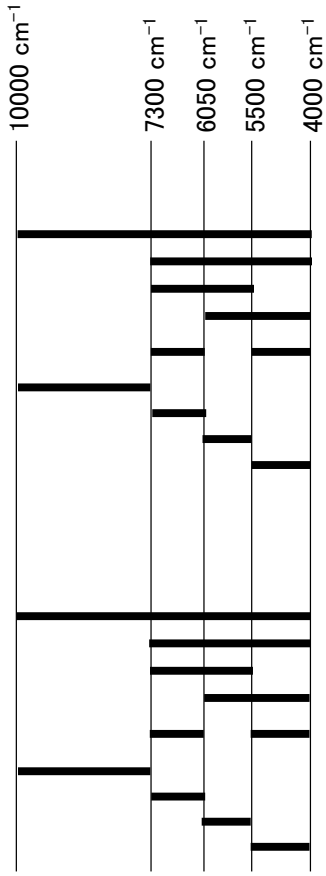
370

371 Fig. 5. (a) The principal components analysis (PCA) scores plotted on the first and second
372 principal components on the basis of the second derivative NIR spectra in the 7300–4000
373 cm^{-1} region.
374 (b) The spectrum obtained from PC1 loading in PCA. The bands at 5970, 5800, and 5220
375 cm^{-1} are assigned to lignin, hemicellulose, and the absorbed water, respectively.
376 (c) Second derivative spectra obtained from unaging heartwood and aging wood of *P.*
377 *thunbergii* designated as KYOw19176 and KYO_ID_5197.
378
379

Table 1. The number of NIR spectra of the wood samples used for calibration and prediction.

		Calibration set	Prediction set	Total
P. densiflora	Sapwood	24	12	36
	Heartwood	24	12	36
P. thunbergii	Sapwood	24	12	36
	Heartwood	24	12	36
Total		96	48	144

Table 2. Statistical summary of the discriminant models based on the calibration and prediction sets obtained from a mixture of sapwood and heartwood samples (a and b), and individual sapwood (c and d) and heartwood (e and f) samples. The discriminant models were obtained by using the original spectra (a, c and e) and the second derivative spectra (b, d and f). A schematic illustration is shown on the left to indicate each spectral region.

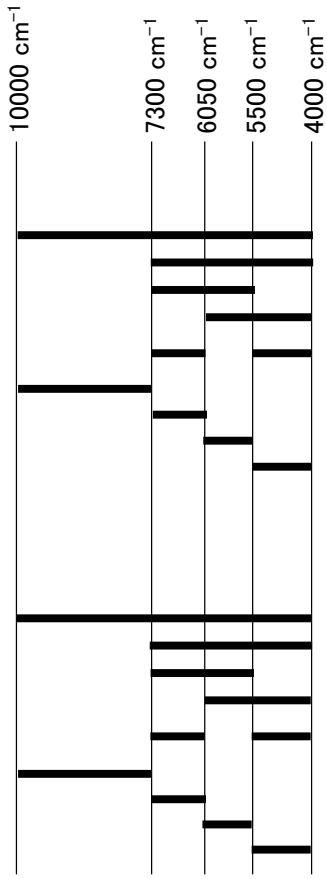


(a)

Spectral region (cm^{-1})	Factors	Calibration set R_c^2	Calibration set RMSEC	Predicition set R_p^2	Predicition set RMSEP	Correct prediction (%)
10000 – 4000	8	0.52	0.69	0.49	0.71	93.8
7300 – 4000	10	0.66	0.58	0.58	0.65	93.8
7300 – 5500	10	0.54	0.68	0.45	0.74	85.4
6050 – 4000	10	0.65	0.59	0.57	0.66	89.6
7300 – 6050 5500 – 4000	10	0.65	0.59	0.53	0.68	91.7
10000 – 7300	9	0.68	0.56	0.32	0.82	79.2
7300 – 6050	7	0.32	0.82	0.25	0.87	77.1
6050 – 5500	9	0.53	0.68	0.48	0.72	83.3
5500 – 4000	9	0.61	0.62	0.56	0.66	89.6

(b)

Spectral region (cm^{-1})	Factors	Calibration set R_c^2	Calibration set RMSEC	Predicition set R_p^2	Predicition set RMSEP	Correct prediction (%)
10000 – 4000	7	0.72	0.53	0.54	0.68	89.6
7300 – 4000	9	0.75	0.50	0.56	0.66	91.7
7300 – 5500	10	0.69	0.56	0.56	0.66	93.8
6050 – 4000	8	0.72	0.53	0.54	0.68	87.5
7300 – 6050 5500 – 4000	8	0.71	0.54	0.55	0.67	89.6
10000 – 7300	2	0.42	0.76	0.23	0.88	68.8
7300 – 6050	2	0.33	0.82	0.28	0.85	72.9
6050 – 5500	10	0.56	0.66	0.53	0.69	81.3
5500 – 4000	8	0.69	0.56	0.52	0.69	85.4

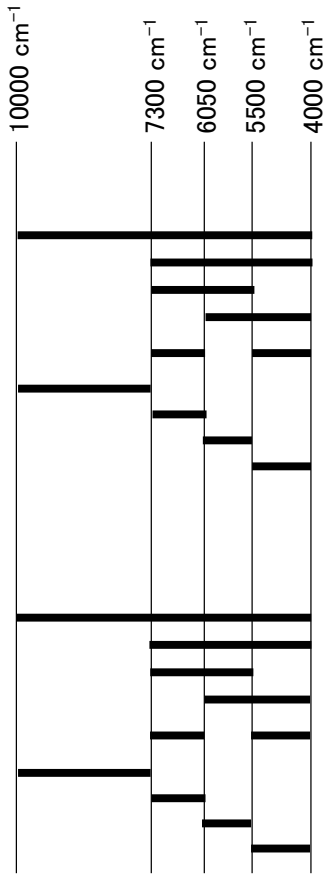


(c)

Spectral region (cm^{-1})	Factors	Calibration set R_c^2	Calibration set RMSEC	Predicition set R_p^2	Predicition set RMSEP	Correct prediction (%)
10000 – 4000	10	0.91	0.30	0.74	0.51	91.7
7300 – 4000	9	0.90	0.31	0.69	0.56	87.5
7300 – 5500	5	0.35	0.80	0.30	0.84	75.0
6050 – 4000	10	0.90	0.32	0.64	0.60	87.5
7300 – 6050 5500 – 4000	10	0.87	0.36	0.65	0.59	87.5
10000 – 7300	7	0.73	0.52	0.21	0.89	83.3
7300 – 6050	6	0.33	0.82	0.01	1.07	54.2
6050 – 5500	9	0.77	0.48	0.39	0.78	79.2
5500 – 4000	10	0.86	0.38	0.60	0.63	87.5

(d)

Spectral region (cm^{-1})	Factors	Calibration set R_c^2	Calibration set RMSEC	Predicition set R_p^2	Predicition set RMSEP	Correct prediction (%)
10000 – 4000	5	0.79	0.46	0.59	0.64	87.5
7300 – 4000	8	0.91	0.30	0.71	0.54	95.8
7300 – 5500	4	0.72	0.53	0.67	0.58	87.5
6050 – 4000	8	0.88	0.35	0.66	0.58	95.8
7300 – 6050 5500 – 4000	6	0.74	0.51	0.64	0.60	87.5
10000 – 7300	1	0.19	0.90	0.02	0.99	54.2
7300 – 6050	5	0.64	0.60	0.33	0.82	75.0
6050 – 5500	5	0.65	0.59	0.59	0.64	91.7
5500 – 4000	6	0.72	0.53	0.61	0.62	87.5



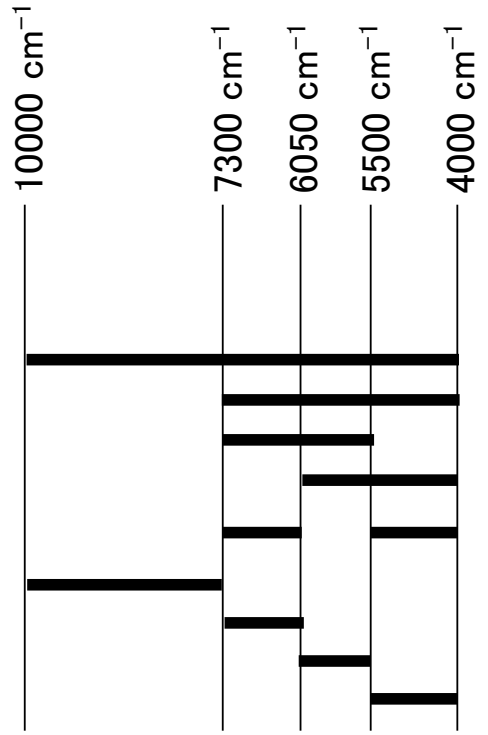
(e)

Spectral region		Calibration set		Predicition set		Correct
(cm^{-1})	Factors	R_c^2	RMSEC	R_p^2	RMSEP	prediction (%)
10000 – 4000	9	0.92	0.28	0.80	0.45	100
7300 – 4000	9	0.93	0.27	0.75	0.50	97.9
7300 – 5500	7	0.81	0.44	0.77	0.48	100
6050 – 4000	9	0.93	0.27	0.81	0.43	100
7300 – 6050 5500 – 4000	9	0.91	0.30	0.78	0.47	100
10000 – 7300	8	0.81	0.44	0.56	0.67	93.8
7300 – 6050	10	0.86	0.37	0.65	0.59	95.8
6050 – 5500	9	0.91	0.30	0.84	0.40	100
5500 – 4000	10	0.95	0.22	0.83	0.42	100

(f)

Spectral region		Calibration set		Predicition set		Correct
(cm^{-1})	Factors	R_c^2	RMSEC	R_p^2	RMSEP	prediction (%)
10000 – 4000	8	0.97	0.18	0.83	0.41	100
7300 – 4000	7	0.92	0.28	0.86	0.37	100
7300 – 5500	8	0.91	0.30	0.73	0.52	95.8
6050 – 4000	7	0.92	0.29	0.86	0.38	100
7300 – 6050 5500 – 4000	7	0.92	0.28	0.85	0.39	100
10000 – 7300	7	0.98	0.12	0.46	0.74	75.0
7300 – 6050	5	0.82	0.43	0.48	0.72	87.5
6050 – 5500	10	0.90	0.31	0.81	0.44	100
5500 – 4000	7	0.91	0.30	0.85	0.39	100

Table 3. Prediction accuracies of the wooden materials used in Chion-In temple as functions of the spectral pretreatment and spectral range. The predictions were made by employing the discriminant models based on the original and second derivative spectra of the heartwood sample for wavenumbers of 7300–4000 cm^{-1}. A schematic illustration is provided on the left to show each spectral region.



Spectral region (cm^{-1})	Correct prediction (%)	
	Original	2nd derivative
10000 – 4000	29.2	45.8
7300 – 4000	25.0	41.7
7300 – 5500	41.7	45.8
6050 – 4000	20.8	41.7
7300 – 6050 5500 – 4000	20.8	45.8
10000 – 7300	25.0	41.7
7300 – 6050	62.5	50.0
6050 – 5500	37.5	45.8
5500 – 4000	20.8	45.8