RESEARCH

Prediction model of the hardness of waterlogged archaeological wood based on NIR spectroscopy

Tiantian Liu[1,2,3], Guanglan Xi[2,4*], Xiangna Han[1,2], Yafang Yin[1,2,3] and Liuyang Han[1,2*]

Abstract

The significance of waterlogged archaeological wood (WAW) lies in its profound informational value, encompassing historical, cultural, artistic, and scientific aspects of human civilization, and therefore need to be properly studied and preserved. In this study, the utilization of near-infrared (NIR) spectroscopy is employed as a predictive tool for assessing the hardness value of WAW. Given the submerged burial conditions, waterlogged wooden heritage frequently undergo substantial degradation in their physical and mechanical properties. The mechanical properties of waterlogged wooden heritage are essential for evaluating their state of preservation and devising appropriate conservation and restoration strategies. However, conventional methods for testing mechanical properties are limited by factors such as the availability of adequate sample size and quantity, adherence to the "principle of minimum intervention," and cost considerations. NIR spectroscopy is a non-destructive, rapid, sensitive, and low-cost analytical technique with great potential for application in this area. In this study, two large and significant ancient Chinese shipwrecks were investigated. One hundred ninety-seven samples were collected and analyzed using NIR spectroscopy and a portable C-type shore hardness testing method. A partial least squares (PLS) regression model was developed to predict the hardness of the WAW. The model was optimized and validated using different preprocessing methods and spectral ranges. The results indicate that the best models were obtained with first derivatives + multiple scattering corrections (MSC) and first derivatives + standard normal variate (SNV) preprocessing in the 1000–2100 nm spectral range, both with an R^2_c of 0.97, a root mean squared error of correction (RMSEC) of 2.39 and 2.40, and a standard error of correction (SEC) of 2.40 and 2.41. Furthermore, they exhibited an R^2_v of 0.89 and 0.87, a root mean squared error of cross-validation (RMSECV) of 4.43 and 4.67, a standard error of cross-validation (SECV) of 4.45 and 4.68, and RPD values of 3.02 and 2.88, respectively. A coefficient of determination of the established prediction model (R^2_p) of 0.89 with a relative standard deviation for prediction (RSD) of 6.9% < 10% was obtained using a sample from the prediction set to predict the established model inversely. These results demonstrate that NIR spectroscopy could enable a rapid, non-destructive, and accurate estimation of the hardness of WAW. Moreover, by carefully choosing appropriate preprocessing techniques and spectral ranges, the predictive capabilities and accuracy of the model can be further enhanced. This research also contributes to the development of a theoretical framework and a methodological approach for future studies in this field. Furthermore, the data obtained from this study are crucial for determining effective preservation strategies for waterlogged archaeological wood.

*Correspondence:
Guanglan Xi
xiguanglan@126.com
Liuyang Han
hanliuyang@ustb.edu.cn
Full list of author information is available at the end of the article

Liu *et al. Heritage Science*   (2023) 11:215

Page 2 of 13

Introduction

Wood, a natural polymeric organic material, plays a significant role in human activities and the development of civilization because it is environmentally friendly, resource-rich, and has biocompatible properties [1, 2]. Furthermore, wood was used as the main construction-material in ancient ships, which are of great conservation and research importance, as they are an important part of the historical and cultural heritage of mankind. However, the cell morphology and chemical structure of wood can be affected by biological hazards and natural degradation during long-term water burial [3], leading to the decomposition of the main components of wood cells such as hemicelluloses, cellulose, and lignin [4, 5]. Degradation increases the porosity and hygroscopicity of the wood tissue, and decreases its density, and ultimately causes a significant reduction in its mechanical properties [6], and finally causes a significant reduction in its mechanical properties [7]. Because the mechanical properties directly determine the service life of wood, they are an important index and essential data for evaluating the state of preservation of wooden heritage and formulating protection and restoration schemes [8]. However, due to WAW's scarcity, destructive research should be avoided when evaluating its mechanical properties [9].

Conventional mechanical testing of wood is typically performed using a universal mechanical testing machines; however, large sample sizes and destructive tests often consume large amounts of archaeological wood to obtain mechanical data for aforementionend conventional testing methods according to national standerd [10]. To reduce the consumption of relics during the evaluation of mechanical properties, researchers have experimented with various microdamage methods for mechanical testing, such as nanoindentation (NI) [1] and static thermomechanical analysis (TMA) techniques [11]. Although these methods reduce the number of cultural relic samples consumed during mechanical property evaluation, some waste is still involved in the testing and processing stages. Additionally, samples prepared according to specific test specifications are often difficult to reuse, and preventing damage to the integrity of wooden heritage proves challenging. To maximize the value of these valuable mechanical property data, researchers have attempted to combine mechanical property results with nondestructive assessment methods to construct mechanical property prediction models. This approach can enhance the speed and efficiency of mechanical property predictions in future studies while minimizing the number of wood relic materials consumed. Near-infrared spectroscopy (NIR) is a rapid and nondestructive technique for evaluating the properties of organic materials. This is reflected in the characteristic absorption of hydrogen-containing groups such as CH, OH, and NH in the near-infrared region [12, 13]. Since it was first proposed in the 1960s [14], NIR technology has been widely applied in various fields, including food, medicine, tobacco, petrochemicals, and agriculture [15, 16]. At present, this technology is widely used in the performance evaluation of sound wood [17, 18]. It can quickly and accurately predict the chemical composition [19, 20], wood species [21], density [22], moisture content [23–25], microfibril angle [26], and mechanical properties [27–30] of wood. However, using different wood species and analytical methods may lead to different prediction results [31]. In addition, the overlap of several spectral ranges can make it difficult to identify the differences [32], and the roughness of the sample surface and moisture can affect the reflection and absorption of the NIR spectra [33]. Researchers have used multiplication scatter correlation (MSC) and standard normal variate (SNV) to reduce particle size, surface scattering, and optical path variations in NIR spectra to address these issues [34, 35]. Furthermore, the effects of baseline drift or smoothing of background interference are eliminated by preprocessing methods, such as first and second derivatives, to improve the resolution of the raw spectra [36]. Although the noise and interference of the spectrum can be effectively reduced using preprocessing methods to improve the accuracy of the prediction model [37, 38], only a few studies on the NIR prediction model for WAW exist, mainly focusing on wood aging [39–42], fungal degradation [43–45] and heat treatment [46, 47]. Using NIR combined with chemometrics, Chen et al. successfully developed an orthogonal partial least squares discriminant analysis (OPLS-DA) model for archaeological hardwoods and softwoods and a predictive model for their degrees of degradation [32]. Yonenobu et al. found that a 1300-year-old archaeological wood contained less hemicellulose and cellulose and more lignin than sound wood [48]. According to Pecoraro et al., NIR spectroscopy was used to analyze decayed wood that was stored under waterlogged conditions for long periods of time [49]. The above studies confirmed NIR spectroscopy

Liu *et al. Heritage Science*   (2023) 11:215

Page 3 of 13

is reliable for detecting the relative chemical composition of decayed wood. However, due to the difficulties of obtaining archaeological wood and performing data analysis, this technique has yet to be widely applied to predict mechanical properties for wooden heritage [50].

The present study employed NIR spectroscopy and chemometric methods to develop a hardness prediction model for WAW, using one hundred ninety-seven samples from two significant ancient Chinese shipwrecks. Meanwhile, the effects of preprocessing methods and band selection on the NIR prediction model were examined.

Materials and methods
Wood samples
In this study, 30 wooden shell platings of the Chinese cruiser Chih Yuen shipwreck from the Qing Dynasty (1885–1894CE) [51] and Nanhai No.1 shipwreck from the Southern Song Dynasty (1127–1279CE) [52, 53] were selected as research subjects. All samples were identified as *Pinus* spp according to anatomical microscopic features [32, 52]. The shell platings were retrieved from the "Nanhai No. 1 Shipwreck Site"(located near Zhanjiang City, Guangdong Province, China) and the "Chih Yuen Shipwreck Site"(located near Dandong City, Liaoning Province, China) (Table 1), and were irregular blocks with relatively flat surfaces [53]. The samples in a waterlogged state were encapsulated in plastic bags and vacuumed. Finally, it was placed in a refrigerator freezer at 10 ℃.

Maximum water content and basic density
Maximum water content (MWC) and basic density (BD) were selected as two representative physical properties of the degree of degradation of waterlogged archaeological wood (WAW). There are five classes of the state of preservation (The class in the following text is equivalent to state of preservation): "less than 135% for the class 0-MWC; 135–225% for the class 1-MWC; 225–350% for the class 2-MWC; 350–500% for the class 3-MWC; more than 500% for the class 4-MWC" [54]. Shell platings Nos. 10 to 30 were tested for MWC and BD as follows: Four samples were taken from each shell plating for the MWC and the BD analysis, and their average value was used to characterise the degradation degree of each shell plating. They were tested in the same way as shell platings Nos. 1 to 9. The calculation methods are based on Eqs. (1) [55, 56] and (2) [57]:

$$BD = \frac{m_0}{V_{max}} \tag{1}$$

$$MWC = \frac{m_{max} - m_0}{m_0} \times 100\% \tag{2}$$

where m_0 is the constant weight of the sample at 102 ± 3℃ (g), V_{max} is the volume of the waterlogged sample measured by the water displacement method (cm^3), and m_{max} represents the mass of the waterlogged sample (g).

Hardness test
The Shore hardness tester is a commonly used tool for determining the hardness of rubber and plastic materials. In this study, the hardness of WAW is relatively lower than that of sound wood, with a surface that feels similar to that of sponge or rubber. Therefore, the C-type Shore hardness tester (Mitutoyo, Japan) was used to determine the surface hardness of the samples, based on the scope of application of this type of instrument. This type of Shore hardness tester consists of a hemispherical indenter with a diameter of 5 mm, and works by measuring the depth of the indenter pressed into the WAW, which is converted into a certain unit of hardness (HC). Note that the force applied should be just enough to bring the anvil and sample into complete contact and that the reading must be taken within 1 s after the anvil and sample have been fully pressed together. In order to pursue "the principle of minimal intervention" in heritage conservation, this approach was adopted without cutting the samples so that the samples could also be used for other studies. The tests were performed by fixing the waterlogged shell platings sample on the experimental bench and placing a flexible cotton grid on its surface (each grid cell was a 2×2 cm square), keeping the testing area defined. Five to eight hardness measurements were taken within each grid cell and averaged, resulting in a total of 197 grid cells on radial sections of 30 shell platings. Care was taken to avoid areas with obvious signs of decay, such as discoloration and the presence sulfur iron compounds. The samples remained waterlogged during the test.

NIR spectra acquisition
NIR spectral data were collected on the same grid cell surface area of the hardness test using a portable NIR spectrometer (TerraSpec4 Hi-Res, Malvern PANalytical, UK) with a spectral range of 350–2500 nm and a scan rate of 100 ms. The device has a resolution of 3 nm@700 and 6 nm@1400/2100 nm. During spectral acquisition, the ambient temperature was maintained at (25 ± 2 ℃) and the average relative humidity was around 20%. During the test, a wet towel was used to wipe away any running water from the sample surface, and then the NIR spectra were taken immediately after the hardness test. The NIR spectra acquisition time within each grid cell was limited to 2 min to ensure that the samples remained waterlogged throughout the NIR spectra acquisition process. Ten NIR spectra were averaged as one after collection at different locations within the same grid cell.

Liu *et al. Heritage Science* (2023) 11:215

Page 4 of 13

Table 1 List of waterlogged archaeological wood samples

Shell plating number	Number of samples	Source	Species	MWC (%)	BD (g/cm^3)
NO.1*	7	Chih Yuen		116%	0.55
NO.2*	5	Chih Yuen		117%	0.54
NO.3*	9	Chih Yuen		132%	0.51
NO.4*	39	Chih Yuen		135%	0.51
NO.5*	7	Chih Yuen		281%	0.29
NO.6*	3	Chih Yuen		156%	0.45
NO.7*	16	Chih Yuen		176%	0.43
NO.8*	12	Chih Yuen		181%	0.42
NO.9*	8	Chih Yuen		199%	0.38
NO.10	14	Chih Yuen			
NO.11	1	Chih Yuen			
NO.12	2	Chih Yuen			
NO.13	2	Chih Yuen			
NO.14	3	Chih Yuen			
NO.15	4	Chih Yuen	*Pinus* sp.		
NO.16	3	Chih Yuen			
NO.17	3	Chih Yuen			
NO.18	3	Chih Yuen			
NO.19	2	Chih Yuen			
NO.20	6	Nanhai No.1			
NO.21	5	Nanhai No.1			
NO.22	12	Nanhai No.1			
NO.23	4	Nanhai No.1			
NO.24	2	Nanhai No.1			
NO.25	11	Nanhai No.1			
NO.26	2	Nanhai No.1			
NO.27	4	Nanhai No.1			
NO.28	1	Nanhai No.1			
NO.29	1	Nanhai No.1			
NO.30	6	Nanhai No.1			

Chih Yuen refers to the Chinese cruiser Chih Yuen shipwreck; Nanhai No.1 refers to the Nanhai No.1 shipwreck. *These samples (samples Nos. 1 to 9) were also included in the Chen et al. study, and their MWC and BD have been tested [32]. Samples Nos. 10 to 30 are new samples of unknown state of preservation and independent of the Chen et al. study, and their MWC and will be tested in this study

MWC stands for maximum water content; BD represents basic density

Data analysis

Principal component analysis (PCA) and orthogonal partial least squares discriminant analysis (OPLS-DA) were employed for qualitative analysis, while PLS was used for quantitative analysis. The software Origin2021 (Origin Lab, USA) was utilized for principal component analysis, spectrum drawing, and significant difference analysis. The discriminant models were established using SIMCA-14.1 software (Umetrics, Umeå, Sweden). The collected near-infrared spectral data were preprocessed and quantitative models were established using the chemometrics software The Unscrambler X 10.4 (CAMO, Norway).

PCA and OPLS-DA

The purpose of performing PCA on spectral data is to visualize samples with varying degrees of degradation in multiple dimensions, and to better understand them through an OPLS-DA discriminant model. This allows for exploration of the correlation between the hardness

Liu *et al. Heritage Science*   (2023) 11:215

Page 5 of 13

and the degree of degradation, providing a basis for establishing a quantitative prediction model. The quality of the PCA model is evaluated using R^2X and Q^2. R^2X reflects the total variability of x and the degree of optimization of the model. Q^2 represents the cumulative contribution rate, describing the cumulative predictive ability of the model and the accuracy of the predictions. The closer R^2X and Q^2 are to 1, the better the model. The OPLS-DA model's goodness-of-fit and reliability were evaluated using R^2X_{cum} and R^2Y_{cum}, while the predictive ability was evaluated using Q^2_{cum}. To check for potential overfitting, a permutation test was performed 200 times. The outcomes are represented by the R^2 and Q^2 interception values on the Y-axis. For a valid model, the interception of R^2 should be below 0.30, and for Q^2, it should be below 0.05.

PLS regression analysis

In this study, the model is optimised by using a selection of preprocessing methods such as first derivative, second derivative, multivariate scattering correction and standard normal variate, combined with the different spectral ranges. The 197 WAW NIR spectroscopy data and hardness data were randomly divided into a calibration set and a prediction set in a ratio of 4:1, and the calibration set data were used to build a NIR prediction model for the hardness of the WAW, and the calibration model was validated using full cross validation. Finally, the model is externally tested using the prediction set data. The data were imported into Unscrambler X 10.4 (CAMO, Norway) for preprocessing and modeling. Based on the results of the established calibration set, validation set and prediction set models, the effects of different preprocessing methods and spectral ranges on the model quality were analyzed. The partial least squares (PLS) was used to model, as it can effectively solve the problem of the large amount of information in the NIR spectroscopy, eliminate the influence of external noise to a certain extent, improve data accuracy, and correlate the independent and dependent variable matrices to obtain the best model [56]. Whether the predictive performance and accuracy of the NIR prediction model established by the partial least squares method reach the standard of practical application requires a certain index to evaluate the model.

In this paper, coefficient of determination of calibration (R^2c), coefficient of determination of cross-validation (R^2v), coefficient of determination of prediction (R^2p), root mean square error of calibration (RMSEC), root mean square error of cross validation (RMSECV), root mean square error of prediction (RMSEP), standard error of calibration (SEC), standard error of cross validation (SECV), standard error of prediction (SEP),

ratio performance of deviation for cross validation (RPD) and relative standard of deviation for prediction (RSD) are used as the evaluation indexes of the model prediction effect. When the RPD value is greater than 2.5, the predictive performance of the model is satisfactory; if the RPD value can reach about 1.5, NIR can be used as a preliminary assessment tool [58, 59]. When the RSD value is less than 10%, the established model can be used for actual estimation [60].

Results and discussion

Hardness of waterlogged archaeological wood (WAW) and its variation in different degradation classes

As shown in Table 2, the MWC values of samples Nos. 10 to 30 in this study ranged from 218 to 631%, and the BD values ranged from 0.15 to 0.36 g/cm³. Then combining the data from Chen et al. study (samples Nos. 1 to 9) in Table 1 and classifying the state of preservation of samples Nos. 1 to 30 according to the method described in Sect. "Maximum water content and basic density", it can be seen that the state of preservation of all the samples ranges between 0 and 4. Furthermore, an analysis of the variability between the groups was performed on the hardness data for the different state of preservation. Figure 1 shows the graded box line-normal distribution plots and ANOVA results for the degrees of degradation of the WAW samples, with mean hardness values of 85.4 HC, 80.6 HC, 75.4 HC, 61.8 HC, and 53.1 HC respectively for classes 0, 1, 2, 3 and 4. Of the five classes of samples, there were significant differences in the hardness data for all grades except between classes 0 and 1, where there was no significant difference (their letters representing variability are all A, $p < 0.05$). The distribution of the hardness data for each class indicates that, except the Class 4, in which data was more concentrated, the remaining classes presented data points that were relatively scattered, suggesting a nonuniform degradation of the samples from Class 0 to Class 3. Overall, there was a correlation between the degree of degradation of the samples and their hardness; the hardness decreased as the degree of degradation increased. On the other hand, the sample data for each grade span a wide range, i.e., there are high standard deviations (STDV), which can be attributed to several factors related to the natural variability of the wood (earlywood/latewood, sapwood/heartwood, etc.). Even with the influence of factors unrelated to the degree of degradation, the results of the analysis of the hardness data proved their reliability in assessing degradation of WAW.

Figure 2 presents the raw spectra obtained by averaging the spectra for each class, visually illustrating the differences between the five classes of samples in different spectral ranges. The results are similar to those of
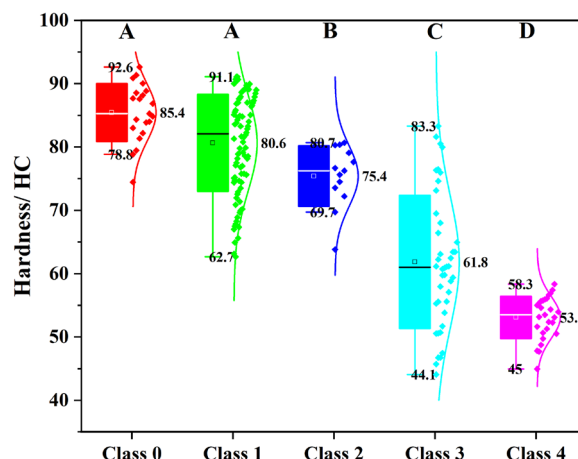
Liu *et al. Heritage Science*   (2023) 11:215

Page 6 of 13

Table 2 Results of MWC, BD and state of preservation of samples

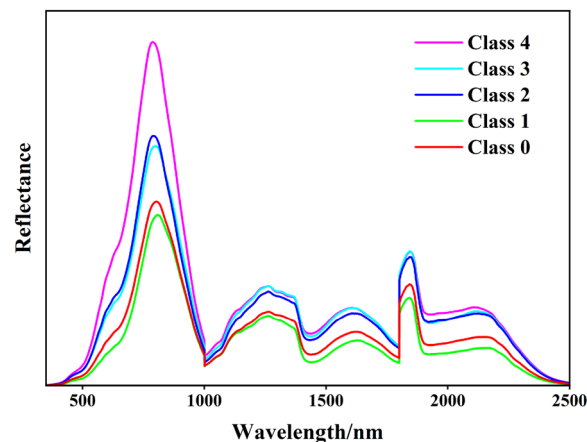| Shell plating number	MWC (%)	BD (g/cm³)	State of preservation
NO.1*	116	0.55	0
NO.2*	117	0.54	0
NO.3*	132	0.51	0
NO.4*	135	0.51	1
NO.5*	281	0.29	2
NO.6*	156	0.45	1
NO.7*	176	0.43	1
NO.8*	181	0.42	1
NO.9*	199	0.38	1
NO.10	218	0.36	1
NO.11	243	0.33	2
NO.12	252	0.32	2
NO.13	142	0.53	1
NO.14	328	0.27	2
NO.15	360	0.23	3
NO.16	373	0.23	3
NO.17	408	0.23	3
NO.18	416	0.21	3
NO.19	452	0.19	3
NO.20	463	0.19	3
NO.21	471	0.19	3
NO.22	473	0.19	3
NO.23	489	0.18	3
NO.24	526	0.17	4
NO.25	548	0.17	4
NO.26	552	0.16	4
NO.27	572	0.16	4
NO.28	575	0.16	4
NO.29	618	0.15	4
NO.30	631	0.15	4

The MWC and BD data for these samples with asterisks (*) are cited from Chen et al. study, which was tested in the same way as in this study

Fig. 1 WAW hardness grading box line-normal distribution scatter diagram; the same letter in Fig. 1 means that the parameters are not significantly different at the 95% confidence level

Fig. 2 Raw NIR spectra of different degrees of degradation of WAW

the variance analysis for hardness. The NIR spectra of classes 0 and 1, 2 and 3 have similar peak shapes, characteristic peak positions, and peak heights. Although the results of the hardness analyses showed four categories (0 + 1, 2, 3 and 4) and the raw spectra reflected only three categories (0 + 1, 2 + 3 and 4), these results still indicate that the NIR spectra of WAW with the same mechanical properties are similar [61]. On the other hand, NIR spectra are rich in information, but also complex, which necessitates further analyses by chemometrics.

Table 3 shows the results of the hardness data, with a range of 44.1–92.6 HC for the calibration set and 45.7–91.4 HC for the prediction set. The standard

Table 3 Statistics of the hardness of Calibration and Prediction

|	N	Max/HC	Min/HC	Mean/HC	Std Dev/HC
Calibration	158	92.6	44.1	73.2	13.46
Prediction	39	91.4	45.7	72.3	13.27

N: number of samples; Max: Maximum value of hardness; Min: Minimum value of hardness; Mean: Mean hardness; Std Dev: Standard deviation; HC is the unit of type C Shore hardness

deviations for the calibration and prediction sets were 13.5 HC and 13.3 HC, respectively. The data in both sets were relatively evenly distributed, and a wide range of hardness values was observed. This is beneficial for modeling and establishing a NIR prediction model that can accurately predict the hardness of a wide range of WAW samples.

Liu *et al. Heritage Science*   (2023) 11:215

Page 7 of 13

Preliminary classification of degree of degradation of waterlogged archaeological wood (PCA and OPLS-DA)

To investigate the effectiveness of NIR spectroscopy in predicting the hardness of WAW, raw spectral data from samples with different degrees of degradation were analyzed using an unsupervised PCA model. The results showed model fit parameters of $R^2X = 0.996$ and $Q^2 = 0.995$, with PC1, PC2, and PC3 explaining 89.3%, 6.2%, and 3.3% of the variance, respectively. The cumulative contribution reached 98.8%, covering essentially all the information in the samples. The PCA score plot (Fig. 3a) shows that the samples with different degrees of degradation were grouped into five clusters. However, there were overlaps between the clusters. Most Class 3 and 4 samples were located in different spaces of the PCA, and although there were some overlaps, they could still be distinguished from the other class samples. Furthermore, a supervised OPLS-DA discriminant model was used to analyze the spectral data. The model

with R^2X_{cum} of 0.59 and R^2Y_{cum} of 0.57 and Q^2 of 0.52, indicating the predictability of OPLS-DA is acceptable. The results showed that (Fig. 3b) Class 0, Class 1, and Class 2 samples still significantly overlapped. However, Class 3 and Class 4 samples could be distinguished from other grade samples, corresponding to the results of the hardness variability analysis. In addition, the applicability of the OPLS-DA model was tested using 200 random permutations (Fig. 3c). The intercepts of R^2 and Q^2 were 0.0098 and -0.118, less than 0.30 and 0.05, indicating that the model was not over-fitted and reliable for classification discrimination [62]. These results confirm that NIR spectroscopy can effectively perform a preliminary discriminant analysis of WAW with different degrees of degradation. Furthermore, this analysis serves as a foundation for quantitatively assessing the mechanical properties of waterlogged archaeological wood using NIR spectroscopy.

(a)

(b)

(c)
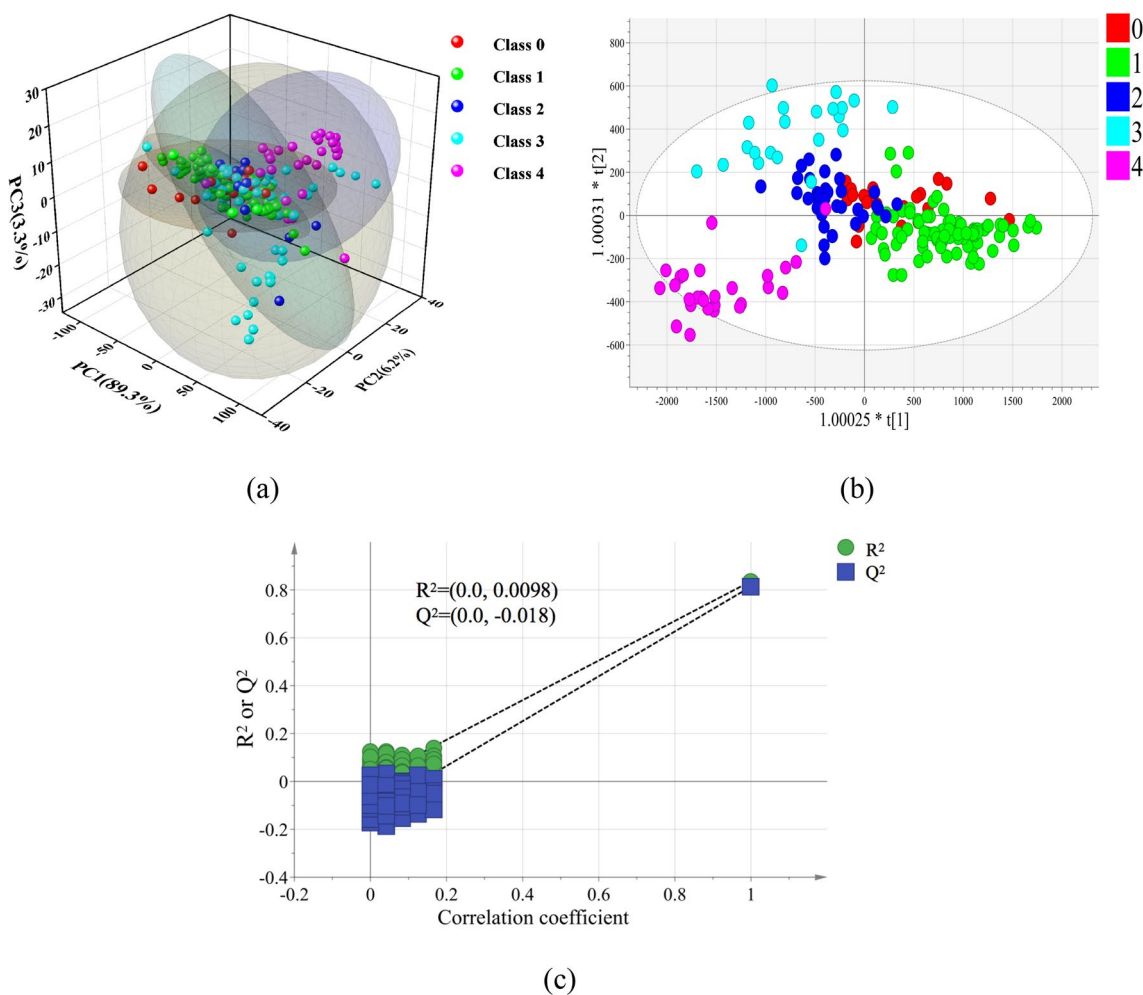
Fig. 3 PCA score plot (**a**), OPLS-DA score plot (**b**), and OPLS-DA model 200 permutation tests (**c**) of WAW with different degrees of degradation

Liu *et al. Heritage Science*   (2023) 11:215

Page 8 of 13

Quantification

Among the 158 calibration set samples used to establish prediction model of WAW hardness, the maximum hardness is 92.6 HC, the minimum is 44.1 HC, the average is 73.2 HC, and the standard deviation is 13.5 (Table 3). The effects of different preprocessing methods for spectra on the quality of the models are discussed for the raw spectral range (350–2500 nm). Data processed with MSC, SNV, first derivative, second derivative, and a combination of the two preprocessing methods were used to build predictive models. A control group of predictive models was built from the raw spectra (Table 4). Table 3 shows the model correlation coefficients established by the raw NIR spectra reached 0.74 and 0.70 for R^2c and R^2v, 6.83 and 7.39 for RMSEC and RMSECV, and 6.85 and 7.41 for SEC and SECV. Meanwhile, the RPD value of 1.82, which is larger than compareison value 1.5 in previous publications [58, 59], indicating that the model developed by raw NIR spectroscopy can preliminarily assess the hardness of WAW.

Only one preprocessing method was used, but applying either MSC or SNV did not lead to a significant improvement in the model's quality. In contrast, the performance and accuracy of the model improved significantly after derivative processing. The R^2c and R^2v after the first and second derivatives reached more than 0.90 and 0.80. The RPD values greater than 2.5, indicating that the derivative-processed prediction model has satisfactory predictive performance [63]. The model achieved the highest quality when the first derivative was combined with both MSC and SNV as the preprocessing methods. The R^2c and R^2v were as high as 0.95 and 0.88, and the RMSEC and SEC were reduced by approximately 50% compared to the raw spectra. The RMSECV and SECV were reduced by approximately 30% compared to the

raw spectra, while the RPD values reached 2.87 and 2.78. Hence, the prediction accuracy of the WAW hardness model based on the first derivative and MSC and SNV combined with the preprocessed NIR spectra was significantly improved.

Modeling using full-band spectral data results in a large amount of information computation, overlapping areas of information, and noise, ultimately affecting the predictive performance and accuracy of the model. In this study, the spectra were divided into three ranges (350–1000 nm, 1000–1800 nm, and 1000–2100 nm) according to the distribution of the spectral information of the wood characteristic compounds, and the effects of inaccessible preprocessing methods on the models generated based on the three spectral ranges were analyzed (Tables 5–7).

As shown in Table 5, modeling in the 350–1000 nm spectral range using the first derivative combined with MSC and SNV was optimal. The R^2c and R^2v of the model based on the first derivative + MSC preprocessing were 0.93 and 0.84, the RMSEC and RMSECV were 3.55 and 5.40, the SEC and SECV were 3.56 and 5.42, and the RPD value was 2.48. The R^2c and R^2v of the model based on the first derivative + SNV preprocessing were 0.94 and 0.85, the RMSEC and RMSECV were 3.33 and 5.30, the SEC and SECV were 3.34 and 5.32, and the RPD value was 2.53. Table 6 displays the outcomes of the optimal model established for the 1000–1800 nm spectral range, which employed the first derivative in combination with MSC and SNV as the preprocessing methods. The R^2c and R^2v of the model based on the first derivative + MSC preprocessing were 0.94 and 0.87, the RMSEC and RMSECV were 3.42 and 4.83, the SEC and SECV were 3.42 and 4.84, and the RPD value was 2.78. The R^2c and R^2v of the model based on the first derivative + SNV preprocessing were 0.96 and 0.88, the RMSEC and RMSECV

Table 4 Hardness model results of WAW based on raw NIR spectra

Mathematical treatment	Calibration set			Validation set				Factors
	R^2c	RMSEC	SEC	R^2v	RMSECV	SECV	RPD	
–	0.74	6.83	6.85	0.70	7.39	7.41	1.82	7
MSC	0.79	6.15	6.17	0.76	6.73	6.75	1.99	7
SNV	0.74	6.80	6.82	0.68	7.73	7.76	1.73	7
1st	0.91	4.07	4.08	0.85	5.19	5.21	2.58	7
2nd	0.95	3.08	3.09	0.84	5.35	5.37	2.51	7
1st + MSC	0.95	3.05	3.06	0.88	4.68	4.69	2.87	7
1st + SNV	0.95	3.11	3.12	0.88	4.82	4.84	2.78	7
2nd + MSC	0.93	3.46	3.47	0.85	5.23	5.24	2.57	5
2nd + SNV	0.93	3.46	3.47	0.86	5.15	5.16	2.61	5

Factors Opt. number of factors, *MSC* Multiple Scattering Correction, *SNV* Standard Normal Variate, *1st* First Deviation, *2nd* Second Derivative, *R^2c* Coefficient of determination of calibration, *RMSEC* Mean Squared Error for Calibration, *SEC* Standard Error of Calibration, *R^2v* Coefficient of determination of cross-validation, *RMSECV* Mean Squared Error for Cross-Validation, *SECV* Standard Error of Cross-Validation, *RPD* Ratio Performance of Deviation for Cross-Validation

Liu *et al. Heritage Science*   (2023) 11:215

Page 9 of 13

Table 5 Hardness model results for WAW based on NIR spectra in the 350–1000 nm range

Mathematical treatment	Calibration set			Validation set				Factors
	R^2c	RMSEC	SEC	R^2v	RMSECV	SECV	RPD	
–	0.68	7.57	7.59	0.62	8.39	8.42	1.60	7
MSC	0.70	7.42	7.44	0.62	8.32	8.35	1.61	7
SNV	0.68	7.66	7.69	0.59	8.61	8.64	1.56	7
1st	0.88	4.69	4.70	0.82	5.72	5.74	2.34	8
2nd	0.88	4.76	4.77	0.81	5.87	5.89	2.29	5
1st + MSC	0.93	3.55	3.56	0.84	5.40	5.42	2.48	9
1st + SNV	0.94	3.33	3.34	0.85	5.30	5.32	2.53	10
2nd + MSC	0.82	5.74	5.76	0.80	6.13	6.15	2.19	2
2nd + SNV	0.80	6.01	6.03	0.79	6.16	6.18	2.18	1

Factors Opt. number of factors, *MSC* Multiple Scattering Correction. *SNV* Standard Normal Variate. *1st* First Deviation, *2nd* Second Derivative, R^2c Coefficient of determination of calibration, *RMSEC* Mean Squared Error for Calibration, *SEC* Standard Error of Calibration, R^2v Coefficient of determination of cross-validation. *RMSECV* Mean Squared Error for Cross-Validation, *SECV* Standard Error of Cross-Validation, *RPD* Ratio Performance of Deviation for Cross-Validation

Table 6 Hardness model results for WAW based on NIR spectra in the 1000–1800 nm range

| Mathematical treatment	Calibration set			Validation set				Factors
	R^2c	RMSEC	SEC	R^2v	RMSECV	SECV	RPD	
–	0.91	4.09	4.10	0.89	4.45	4.47	3.01	8
MSC	0.91	4.11	4.12	0.89	4.54	4.55	2.96	7
SNV	0.91	4.00	4.02	0.89	4.56	4.58	2.94	8
1st	0.92	3.70	3.72	0.89	4.51	4.52	2.98	6
2nd	0.91	4.06	4.07	0.77	6.49	6.50	2.07	4
1st + MSC	0.94	3.42	3.43	0.87	4.83	4.84	2.78	5
1st + SNV	0.96	2.77	2.78	0.88	4.62	4.63	2.91	7
2nd + MSC	0.88	4.74	4.75	0.71	7.29	7.31	1.84	3
2nd + SNV	0.87	4.69	4.70	0.71	7.23	7.25	1.86	3

Factors Opt. number of factors, *MSC* Multiple Scattering Correction. SNV: Standard Normal Variate, *1st* First Deviation, *2nd* Second Derivative, R^2c Coefficient of determination of calibration, *RMSEC* Mean Squared Error for Calibration, *SEC* Standard Error of Calibration, R^2v Coefficient of determination of cross-validation, *RMSECV* Mean Squared Error for Cross-Validation, *SECV* Standard Error of Cross-Validation, *RPD* Ratio Performance of Deviation for Cross-Validation

Table 7 Hardness model results for WAW based on NIR spectra in the 1000–2100 nm range

Mathematical treatment	Calibration set			Validation set				Factors
	R^2c	RMSEC	SEC	R^2v	RMSECV	SECV	RPD	
–	0.93	3.55	3.56	0.89	4.42	4.44	3.03	11
MSC	0.94	3.38	3.39	0.90	4.31	4.33	3.11	10
SNV	0.94	3.16	3.17	0.88	4.59	4.60	2.93	12
1st	0.96	2.61	2.62	0.88	4.73	4.73	2.85	7
2nd	0.93	3.49	3.50	0.82	5.73	5.75	2.34	5
1st + MSC	0.97	2.39	2.40	0.89	4.43	4.45	3.02	7
1st + SNV	0.97	2.40	2.41	0.87	4.67	4.68	2.88	7
2nd + MSC	0.93	3.57	3.58	0.84	5.45	5.47	2.46	4
2nd + SNV	0.95	3.04	3.05	0.84	5.49	5.51	2.44	5

Factors Opt. number of factors, *MSC* Multiple Scattering Correction, *SNV* Standard Normal Variate, *1st* First Deviation, *2nd* Second Derivative, R^2c Coefficient of determination of calibration, *RMSEC* Mean Squared Error for Calibration, *SEC* Standard Error of Calibration, R^2v Coefficient of determination of cross-validation, *RMSECV* Mean Squared Error for Cross-Validation, *SECV* Standard Error of Cross-Validation, *RPD* Ratio Performance of Deviation for Cross-Validation

were 2.77 and 4.62, the SEC and SECV were 2.78 and 4.63, and the RPD value was 2.91. Table 7 presents the outcomes of the optimal model established for the 1000–2100 nm spectral range, which utilized the first derivative in combination with MSC and SNV as the preprocessing methods. The R^2c and R^2v of the model based on the first derivative + MSC preprocessing were 0.97 and 0.89, the RMSEC and RMSECV were 2.39 and 4.43, the SEC and SECV were 2.40 and 4.45, and the RPD value was 3.02. The R^2c and R^2v of the model based on the first derivative + SNV preprocessing were 0.97 and 0.87, the RMSEC and RMSECV were 2.40 and 4.67, the SEC and SECV were 2.41 and 4.68, and the RPD value was 2.88.

In summary, the optimal models developed for the four spectral ranges (350–2500 nm, 350–1000 nm, 1000–1800 nm, and 1000–2100 nm) were based on first derivative + MSC and first derivative + SNV preprocessing. For the WAW, the best quality model was developed for the 1000–2100 nm spectral range (Fig. 4a, b). Similarly, prior

research on the physical and mechanical properties of sound wood has reported that models developed using the 1000–2100 nm spectral range demonstrate improved predictive performance [64]. This may be because this range contains most of the characteristic information of WAW, and the spectral profiles in this range are clearer and less noisy (Fig. 5a, b).

In this study, a model for predicting the hardness of WAW based on NIR spectroscopy was developed and demonstrated higher prediction accuracy and stability than other similar studies. For example, Raul et al. employed NIR spectroscopy to forecast the hardness of charcoal post heat treatment at varying temperatures; the model R^2c at different temperatures was only 0.507, and R^2v was only 0.427 [65]. Tetsuya et al. developed a NIR spectroscopy prediction model simulating the compression modulus of archaeological wood with a coefficient of determination of 0.90 [47]. However, when the technique was applied to predict the bending modulus of

(a)   (b)
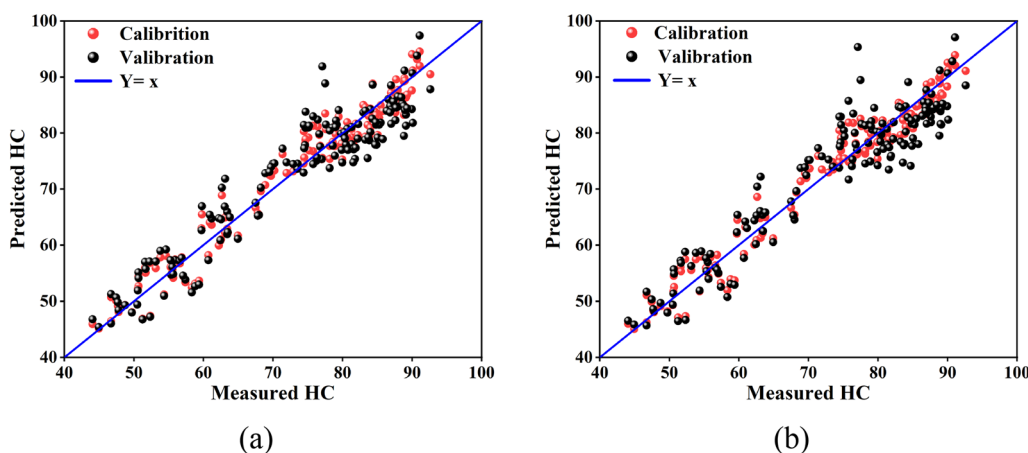
Fig. 4 Predication model for the 1000–2100 nm spectral range. Note: Diagonal straight lines (Y=x) indicates the perfect correlation trend

(a)   (b)
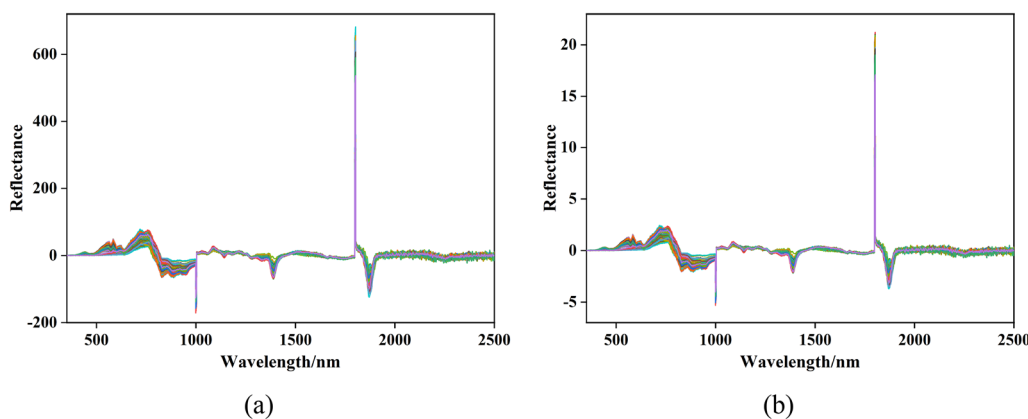
Fig. 5 NIR spectra of samples after first order derivative + MSC (**a**) and first-order derivative + SNV (**b**) treatments

archaeological wood, the model coefficient of determination was reduced to 0.82 [51]. Wang et al. integrated two preprocessing techniques to establish a NIR spectroscopy predictive model for the bending strength of *Catalpa bungei wood*, achieving optimal coefficient of determination of 0.843 and 0.846 for the bending strength and modulus of elasticity [66]. The lower coefficient of determination may be related to the larger coefficients of variation in compression modulus, flexural modulus, and flexural strength in the above referring studies. However, in this study, the large gradient in degradation and the uniform distribution of the hardness dataset between light and heavy degradation facilitated the mathematical modeling of the NIR spectroscopy prediction model.

In addition, this study optimized a NIR hardness prediction model for WAW based on different preprocessing methods and spectral ranges, resulting in a model with high predictive performance and accuracy. However, an external validation of the model is required to test its accuracy and the actual prediction results.

External validation

The model obtained for the first derivatives + MSC preprocessing in the 1000–2100 nm spectral range was selected for external testing and applied to 39 prediction set samples (39 samples were randomly selected from the total sample). The predictions were compared with the measured results (Fig. 6). Figure 6 shows that the model has a prediction set coefficient of determination R^2p of 0.89, an RMSEP of 5.04, and an RSD value of 6.9% < 10%, indicating that the model has high predictive performance and accuracy for initial application in practice [67]. However, the results of grading the degree of degradation based on the maximum water content

do not correspond to the results predicted by hardness, which may be related to the non-uniform degradation of the samples. In contrast, the hardness data corresponds to the NIR spectra (Fig. 6). For example, even though some samples belong to the same classification based on the maximum water content method and the error between the model-derived hardness predictions and the measured hardness values is small, the hardness results are distributed in different intervals. Hence, relying solely on maximum water content or basic density for grading the degradation of WAW is unreliable. The inclusion of mechanical properties like hardness becomes necessary.

Conclusions

This study developed a hardness prediction model for archaeological waterlogged wood (WAW) by combining NIR spectroscopy with chemometrics. It also examined the impact of preprocessing methods and spectral ranges. The optimal prediction models were established when the preprocessing method was either first-derivative + MSC or first-derivative + SNV for a spectral range of 1000–2100 nm. Both models achieved R^2c of 0.97, R^2v of 0.89 and 0.87, and RPD values of 3.02 and 2.88. The prediction model established through external validation achieved a correlation coefficient (R^2p) of 0.89, an RMSEP of 5.04, and an RSD value of 6.9% < 10%, indicating high predictive performance and accuracy. Furthermore, hardeness data and results from the inverse prediction model indicated that the degradation of WAW was non-uniform. The classification of degradation based on maximum water content did not precisely align with the hardness prediction results. Additionally, the hardness data exhibited a stronger correlation with NIR spectra compared to maximum water content. Therefore, it is important to include hardness as an indicator for assessing the state of preservation of WAW.

The hardness and NIR spectroscopy data of WAW obtained under waterlogged conditions provide a theoretical foundation for the rapid, nondestructive, and accurate evaluation of the mechanical properties of wooden heritage. In this study, a preliminary near-infrared prediction model for the hardness of WAW has been established, which can provide a new means for rapid, nondestructive and accurate assessment of the mechanical properties of WAW. At the same time, this study should be further improved, such as how to overcome the high STDV of the hardness data due to the natural variability of the wood (earlywood/latewood, sapwood/heartwood) and other factors, and how to improve the prediction accuracy of the model. In the future, this research will expand to include important mechanical properties like bending modulus, bending strength, and compressive strength. This will establish a theoretical
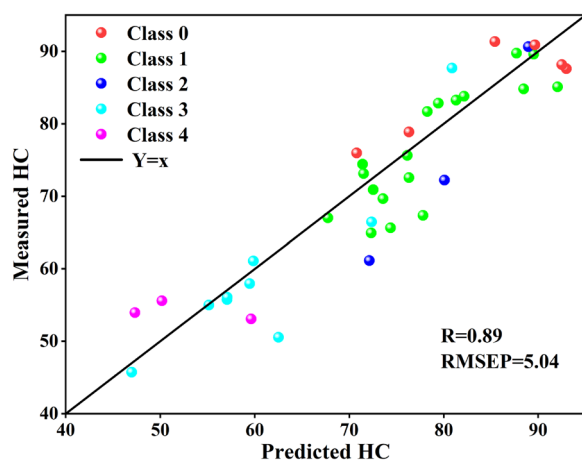
Fig. 6 Inverse prediction model. Note: Diagonal straight lines (Y=x) indicates the perfect correlation trend

Liu *et al. Heritage Science*   (2023) 11:215

Page 12 of 13

basis for predictive models to nondestructively, quickly, and accurately assess the mechanical properties of WAW. Additionally, it may provide important data for assessing nondestructively the state of preservation of WAW and thus assist in the development of an appropriate conservation strategy.

Declarations

Ethics approval and consent to participate
This manuscript does not involve the use of human participants, human data, human tissue, or animals. Therefore, ethics approval and consent are not applicable to this study.

Consent for publication
This manuscript does not contain any individual person's data in any form. Therefore, consent to publish is not required.

Competing interests
The authors declare that they have no competing interests.

Author details
[1]Key Laboratory of Archaeomaterials and Conservation, Ministry of Education, University of Science and Technology Beijing, Beijing 100083, China. [2]Institute for Cultural Heritage and History of Science & Technology, University of Science and Technology Beijing, Beijing 100083, China. [3]Research Institute of Wood Industry, Chinese Academy of Forestry, Beijing 100091, China. [4]National Center for Archaeology, Beijing 100013, China.

References
1. Han L, Han X, Tian X, Zhou H, Yin Y, Guo J. NIR spectroscopy and thermogravimetric analysis to assess the effect of three reinforcing agents on the micromechanical properties of archaeological wood from "Xiaobaijiao I" shipwreck. Spectr Spectr Anal. 2022;42:1529–34.
2. Andrey Borisovich S, et al. Fire hazard and fire resistance of wooden structures. Cham: Springer; 2023. p. 1–15.
3. Chen J, Huang X, Chen X, Chen Z. Corrosion type and conservation of archaeological waterlogged wood. Mater Rev. 2022;29(96–101):128.
4. Xia Y, Chen TY, Wen JL, Zhao YL, Qiu J, Sun RC. Multi-analysis of chemical transformations of lignin macromolecules from waterlogged archaeological wood. Int J Biol Macromol. 2018;109:407–16.
5. Han L, Wang K, Wang W, Guo J, Zhou H. Nanomechanical and topochemical changes in elm wood from ancient timber constructions in relation to natural aging. Materials. 2019;12:786.
6. Broda M, Hill CAS. Conservation of waterlogged wood—past. Present Future Perspectives For. 2021;12:1193.
7. Wu M, Qin Z, Han L, Han X, Chen K, Zhang Z, Yin Y. Review of materials and methods for dehydrating and reinforcing waterlogged wooden cultural relics. China Cult Herit. 2022;108:84–93.
8. Han L, Tian X, Keplinger T, Zhou H, Li R, Svedstrom K, Burgert I, Yin Y, Guo J. Even visually intact cell walls in waterlogged archaeological wood are chemically deteriorated and mechanically fragile: a case of a 170 Year-old shipwreck. Molecules. 2020;25:1113.
9. Han L, Han X, Liang G, Tian X, Ma F, Sun S, Yin Y, Xi G, Guo H. Even samples from the same waterlogged wood are hygroscopically and chemically different by simultaneous DVS and 2D COS-IR spectroscopy. Forests. 2023;14:15.
1. Fujimoto T, Kurata Y, Matsumoto K, Tsuchikawa S. Application of near infrared spectroscopy for estimating wood mechanical properties of small clear and full length lumber specimens. J Near Infrared Spectrosc. 2008;16:529–37.
11. Wu M, Han X, Qin Z, Zhang Z, Xi G, Han L. A quasi-nondestructive evaluation method for physical-mechanical properties of fragile archaeological wood with TMA: a case study of an 800-year-old shipwreck. Forests. 2022;13:38.
12. Manley M. Near-infrared spectroscopy and hyperspectral imaging: non-destructive analysis of biological materials. Chem Soc Rev. 2014;43:8200–14.
13. Hoffmeyer P, Pedersen JG. Evaluation of density and strength of Norway spruce wood by near infrared reflectance spectroscopy. Holz Roh Werkstoff. 1995;53:165–70.
14. Norris KH, Butler WL. Techniques for obtaining absorption spectra on intact biological samples. Ire Trans Biomed Electron. 1961;8:153–7.
15. Sandak A, Sandak J, Zborowska M, Prądzyński W. Near infrared spectroscopy as a tool for archaeological wood characterization. J Archaeol Sci. 2010;37:2093–101.
16. Chen H, Zhou M, Wu Q, Liu Y, Li X. Wang, Near-infrared spectral consistency test model for rapid identification of aminotrimadol tablets. Strait Pharm J. 2022;34:18–21.
17. Yang Z, Li K, Zhang M, Xin D, Zhang J. Rapid determination of chemical composition and classification of bamboo fractions using visible-near infrared spectroscopy coupled with multivariate data analysis. Biotechnol Biofuels. 2016;9:35.
18. Assis Loureiro B, Arriel TG, Guedes Ramalho FM, Hein PRG, Trugilho PF. NIR-based models for estimating selected physical and chemical wood properties from fast-growing plantations. iFor Biogeosci For. 2022;15:372–80.
19. Alves AMM, Simões RFS, Santos CA, Potts BM, Rodrigues J, Schwanninger M. Determination of *Eucalyptus globulus* wood extractives content by near infrared-based partial least squares regression models: comparison between extraction procedures. J Near Infrared Spectrosc. 2012;20:275–85.
20. da Silva AR, Monteiro Pastore TC, Batista Braga JW, Davrieux F, Arakaki Okino EY, Rauber Coradin VT, Alves Camargos JA, do Prado AGS. Assessment of total phenols and extractives of mahogany wood by near infrared spectroscopy (NIRS). Holzforschung. 2013;67:1–8.
21. Abe H, Watanabe K, Ishikawa A, Noshiro S, Fujii T, Iwasa M, Kaneko H, Wada H. Non-destructive method for species identification of historical wooden statues using NIR spectroscopy: analysis using wood samples of the internationally authorized collection. Wood Preservation. 2015;41:162–70.
22. Watanabe K, Kobayashi I, Kuroda N, Harada M, Noshiro S. Predicting oven-dry density of Sugi (*Cryptomeria japonica*) using near infrared (NIR) spectroscopy and its effect on performance of wood moisture meter. J Wood Sci. 2012;58:383–90.
23. Kobori H, Inagaki T, Fujimoto T, Okura T, Tsuchikawa S. Fast online NIR technique to predict MOE and moisture content of sawn lumber. Holzforschung. 2015;69:329–35.

Liu *et al. Heritage Science*   (2023) 11:215

Page 13 of 13

24. Koumbi-Mounanga T, Groves K, Leblon B, Zhou G, Cooper PA. Estimation of moisture content of trembling aspen (*Populus tremuloides* Michx.) strands by near infrared spectroscopy (NIRS). Eur J Wood Wood Prod. 2014;73:43–50.

25. Hans G, Leblon B, Cooper P, La Rocque A, Nader J. Determination of moisture content and basic specific gravity of *Populus tremuloides* (Michx.) and *Populus balsamifera* (L.) logs using a portable near-infrared spectrometer. Wood Mater Sci Eng. 2014;10:3–16.

26. Giroud G, Defo M, Bégin J, Ung C-H. Application of near-infrared spectroscopy to determine the juvenile-mature wood transition in black spruce. For Prod J. 2015;65:129–38.

27. Mancini M, Leoni E, Nocetti M, Urbinati C, Duca D, Brunetti M, Toscano G. Near infrared spectroscopy for assessing mechanical properties of *Castanea sativa* wood samples. J Agric Eng. 2019;50:191–7.

28. Schimleck LR, Matos JLM, Trianoski R, Prata JG. Comparison of methods for estimating mechanical properties of wood by NIR spectroscopy. J Spectrosc. 2018;2018:1–10.

29. Yu L, Liang Y, Zhang Y, Cao J. Mechanical properties of wood materials using near-infrared spectroscopy based on correlation local embedding and partial least-squares. J For Res. 2019;31:1053–60.

30. Schimleck LR, Jones PD, Peter GF, Daniels RF, Clark A. Success in using NIR spectroscopy to estimate wood properties of *Pinus taeda* radial strips not due to autocorrelation. NIR News. 2017;17:13–4.

31. Li Y. Visible-near infrared spectroscopy for wood species identification and density prediction. Harbin: Northeast Forestry University; 2019.

32. Chen J, Liu S, Yin L, Cao H, Xi G, Zhang Z, Liu J, Luo R, Han L, Yin Y, Guo J. Non-destructive preservation state estimation of waterlogged archaeological wooden artifacts. Spectrochim Acta A Mol Biomol Spectrosc. 2023;285:121840.

33. Schimleck LR, Jones PD, Clark A III, Daniels RF, Peter GF. Near infrared spectroscopy for the nondestructive estimation of clear wood properties of *Pinus taeda* L. from the southern United States. For Prod J. 2005;55:21–8.

34. Dhanoa MS, Lister SJ, Sanderson R, Barnes RJ. The link between multiplicative scatter correction (MSC) and standard normal variate (SNV). Transform NIR Spectra. 1994;2:43–7.

35. Chu XL, Yuan HF, Lu WZ. Progress and application of spectral data pretreatment and wavelength selection methods in NIR analytical technique. Prog Chem. 2004;16:528–42.

36. Kan X. Research on non-destructive testing of wood moisture content by near-infrared spectroscopy at different temperatures. Harbin: Northeast Forestry University; 2022.

37. Isaksson T, Næs T. The effect of multiplicative scatter correction (MSC) and linearity improvement in NIR spectroscopy. Appl Spectrosc. 1988;42:1273–84.

38. Woo Y-A, Terazawa Y, Chen JY, Iyo C, Terada F, Kawano S. Development of a new measurement unit (MilkSpec-1) for rapid determination of fat, lactose, and protein in raw milk using near-infrared transmittance. Spectroscopy. 2002;56:599–604.

39. Pizzo B, Pecoraro E, Macchioni N. A new method to quantitatively evaluate the chemical composition of waterlogged wood by means of attenuated total reflectance fourier transform infrared (ATR FT-IR) measurements carried out on wet material. Appl Spectrosc. 2013;67:553–62.

40. Sandak A, Sandak J, Babiński L, Pauliny D, Riggio M. Spectral analysis of changes to pine and oak wood natural polymers after short-term waterlogging. Polym Degrad Stab. 2014;99:68–79.

41. Tsuchikawa S, Yonenobu H, Siesler HW. Near-infrared spectroscopic observation of the ageing process in archaeological wood using a deuterium exchange method. Analyst. 2005;130:379–84.

42. Sandak A, Rozanska A, Sandak J, Riggio M. Near infrared spectroscopic studies on coatings of 19th century wooden parquets from manor houses in South-Eastern Poland. J Cult Herit. 2015;16:508–17.

43. Fackler K, Schwanninger M. How spectroscopy and microspectroscopy of degraded wood contribute to understand fungal wood decay. Appl Microbiol Biotechnol. 2012;96:587–99.

44. Fackler K, Schwanninger M. Polysaccharide degradation and lignin modification during brown rot of spruce wood: a polarised fourier transform near infrared study. J Near Infrared Spectrosc. 2010;18:403–16.

45. Fackler K, Schwanninger M. Accessibility of hydroxyl groups of brown-rot degraded spruce wood to heavy water. J Near Infrared Spectrosc. 2011;19:359–68.

46. Inagaki T, Mitsui K, Tsuchikawa S. Near-infrared spectroscopic investigation of the hydrothermal degradation mechanism of wood as an analogue of archaeological objects. Part I Softwood Appl Spectrosc. 2008;62:1209–15.

47. Inagaki T, Matsuo M, Tsuchikawa S. NIR spectral–kinetic analysis for thermally degraded Sugi (*Cryptomeria japonica*) wood. Appl Phys A. 2016;122:208.

48. Yonenobu H, Tsuchikawa S. Near-infrared spectroscopic comparison of antique and modern wood. Appl Spectrosc. 2003;57:1451–3.

49. Pecoraro E, Pizzo B, Alves A, Macchioni N, Rodrigues JC. Measuring the chemical composition of waterlogged decayed wood by near-infrared spectroscopy. Microchem J. 2015;122:176–88.

50. Inagaki T, Yonenobu H, Asanuma Y, Tsuchikawa S. Determination of physical and chemical properties and degradation of archeological Japanese cypress wood from the Tohyamago area using near-infrared spectroscopy. J Wood Sci. 2018;64:347–55.

51. Zhou C. Liaoning Dandong Chih Yuen shipwreck site investigation, study on natural and cultural. Heritage. 2020;5:12–20.

52. Li R, Guo J, Macchioni N, Pizzo B, Xi G, Tian X, Chen J, Sun J, Jiang X, Cao J, Zhang Z, Yin Y. Characterisation of waterlogged archaeological wood from Nanhai No. 1 shipwreck by multidisciplinary diagnostic methods. J Cult Herit. 2022;56:25–35.

53. Zhang H, Shen D, Zhang Z, Ma Q. Characterization of degradation and iron deposits of the wood of Nanhai I shipwreck. Herit Sci. 2022;10(1):1–3.

54. Macchioni N, Capretti C, Sozzi L, Pizzo B. Grading the decay of waterlogged archaeological wood according to anatomical characterisation. The case of the Fiavé site (N-E Italy). Int Biodeterior Biodegradation. 2013;84:54–64.

55. Yin S, Li C, Meng Y, Wang C, Zhao J, Li Y. Near infrared spectral estimation and model optimization of Tilia tuan based on different pretreatments. J Central South Univ For Technol. 2020;40:171–80.

56. Kollmann FP, Cote WA. Principles of wood science and technology I: solid wood. New York: Springer; 1968.

57. Babiński L, Izdebska-Mucha D, Waliszewska B. Evaluation of the state of preservation of waterlogged archaeological wood based on its physical properties: basic density vs. wood substance density. J Archaeol Sci. 2014;46:372–83.

58. Jones PD, Schimleck LR, Peter GF, Daniels RF, Clark Iii A. Nondestructive estimation of *Pinus taeda* L. wood properties for samples from a wide range of sites in Georgia. Can J For Res. 2005;35:85–92.

59. Schimleck LR, Doran JC, Rimbawanto A. Near infrared spectroscopy for cost effective screening of foliar oil characteristics in a *Melaleuca cajuputi* breeding population. J Agric Food Chem. 2003;51:2433–7.

60. Xu H. NIR-based modeling wood density of major hardwood species in northeastern region. Harbin: Northeast Forestry University; 2015.

61. Jia R. Study on physical and mechanical properties and their correlation with microstructure of different poplar clones. Beijing: Chinese Academy of Forestry; 2021.

62. He X, Yangming H, Górska-Horczyczak E, Wierzbicka A, Jeleń HH. Rapid analysis of Baijiu volatile compounds fingerprint for their aroma and regional origin authenticity assessment. Food Chem. 2021;337:128002.

63. Williams PC, Sobering DC. Comparison of commercial near infrared transmittance and reflectance instruments for analysis of whole grains and seeds. J Near Infrared Spectrosc. 1993;1:25–32.

64. Zhang H, Li Y, Zhang H, Zhang Y, Li P. Moisture content prediction of Korean pine wood based on near infrared spectroscopy at varied wave lengths. J North-East For Univ. 2011;39:83–5.

65. de Abreu Neto R, Ramalho FMG, Costa LR, Hein PRG. Estimating hardness and density of wood and charcoal by near-infrared spectroscopy. Wood Sci Technol. 2020;55(1):215–30.

66. Wang R, Shi L, Wang Y. Rapid prediction of bending resistance properties of *Liriodendron tulipifera* wood using NIR spectroscopy, spectroscopy and spectral. Analysis. 2023;43:557–62.

67. Qi JL, Yang GP, Wang ZK, Niu ZY. Determination of organotin compounds in timber and wood products by GC-MS. Chin J Anal Chem. 2013;41:119–22.

Publisher's Note