Material Evaluation of Cellulosic Materials by Terahertz Time-domain Spectroscopy

(テラヘルツ時間領域分光法によるセルロース系材料の材質評価)

A Dissertation

Submitted to the Graduate School of Bioagricultural Sciences

Nagoya University

In Partial Fulfillment of the Requirements for the Degree

of

Doctor of Agricultural Sciences

By

WANG Han

(王晗)

Lab. System Engineering for Biology

Department of Forest and Environmental Resources Sciences

Graduate School of Bioagricultural Sciences

NAGOYA UNIVERSITY

July 2022

Contents

1. Introduction

Since the second industrial revolution, the use of fossil fuels has made industries very efficient, but as by-products of efficient industries, global environmental problems such as global warming and climate anomalies have now become major challenges for human survival. As the first of six positive climate actions proposed by the UN Secretary-General, decarbonization of all aspects of the industry is necessary to combat and contain climate change and its impacts. In order to achieve this goal, it is important to actively use environmentally friendly materials such as cellulose materials, including wood, bamboo, straw, and its derivatives materials. Cellulosic materials as mentioned above are degradable, recyclable, and have the characteristics of low energy consumption in the production process, however, wood and bamboo, the native cellulosic materials are difficult to ensure homogeneity in production, due to the large differences in individual properties. Therefore, developing a new method that can easily, quickly and non-destructively evaluate materials is necessary.

Some parameters are notable, such as the density and moisture content (MC) of wood, bamboo, the crystalline structures, and the relative content of the cellulose crystalline in a polycrystalline cellulose material, etc., which influence the stiffness and physicochemical properties of the cellulosic materials. The most important of these is the crystalline information of cellulose, including lattice structures, cellulose crystallinity, and if possible, the amount of crystalline cellulose in the materials. Because the other parameters are largely affected by the crystalline state. For example, as the amount of cellulose crystallinity increases, the density of the material increases, and so does the MOE , while the hygroscopic performance decreases because the regular arrangement of crystals makes it difficult for water molecules to enter the interior (Sugino et al. 2007).

Several techniques have been used to determine these parameters. The density, cellulose crystallinity and MC of wood have been evaluated by X-ray (Lindgren 1991; Bhuiyan et al. 2000; Keunecke et al. 2012), Fourier-transform infrared spectroscopy (FT-IR), near-infrared (NIR) spectroscopy in many studies (Fujimoto et al. 2007; Inagaki et al. 2009; Popescu et al. 2011; Ma et al. 2017). And for the cellulose, which is the most abundant and widely distributed polysaccharide in nature, plays a role in supporting the structural framework of cell walls (of plants, algae, and bacteria). For investigating the cellulose relative content and crystalline structure, the typical

examples are using cross-polarization/magic angle spinning (CP/MAS) and carbon 13 nuclear magnetic resonance (13C NMR) are used to determine the crystalline form of cellulose (Atalla and VanderHart 1984; VanderHart and Atalla 1984; Horii et al. 1987), X-ray diffraction (XRD) which is the most widely used in cellulose crystalline study, to determine the crystalline form and the crystallinity index (Wada et al. 2001; Nakamura et al. 2004). FT-IR, NIR spectroscopy have also been reported in the study of cellulose materials to observe the response of chemical bonds (Langkilde and Svantesson 1995; Schenzel and Fischer 2001; Inagaki et al. 2009).

Terahertz radiation lies in the far-infrared part of the electromagnetic spectrum, between the microwave and the infrared radiation, with the frequency range from 0.3 to 3 THz designated by The International Telecommunication Union (ITU), where the frequency range corresponding to the wavelength from 0.1 to 1 mm. The upper boundary varies from the criteria, another always be used is up to 10 THz with a wavelength of 0.01 mm. Limited by the effective THz source and sensitive detectors, this band was also called the "THz gap". With the development of a series of new technologies and new materials in the 1980s, especially the development of ultra-short pulse laser, the generation and modulation of broadband stable pulsed THz sources have become possible, and terahertz time-domain spectroscopy (THz-TDS) has been developed (Williams 2006).

There are some advantages of the THz range (Baxter and Guglietta 2011). Firstly, many non-conductive dry materials (wood, ceramic, paper, etc.) are transparent at THz frequency (Koch et al. 1998; Reid and Fedosejevs 2006), therefore, the transmission imaging of many materials with the sub-millimeter spatial resolution is possible, which is much clear than the image obtained from microwave radiation. Secondly, the characteristic absorption can be used to identify substances. There are already cases that used THz in the security field to detect explosives, weapons, and drugs (Kawase et al. 2003; Federici et al. 2005). And most important, THz is very sensitive to the intermolecular effects; not only the hydrogen-bonding and Van der Waals forces that can be easily detected but also the phonons in crystal lattices directly respond to this region (Mantsch and Naumann 2010; Baxter and Guglietta 2011). There are many reports about using THz to detect the biomolecules and the molecular dynamics, especially for the water molecular, DNA, and proteins (Fischer et al. 2002; Walther et al. 2002; Ebbinghaus et al. 2007; Born et al. 2009). And investigating the pharmaceutical polymorphs structures and the crystallinity, which influence the drug solubility

and dissolution rates (Pickwell and Wallace 2006; Zeitler et al. 2007; Takeuchi et al. 2012). Furthermore, THz radiation is non-ionized, which is safe for human bodies. As the advantages listed above, THz radiation showed potential in the evaluation of cellulosic materials. However, the research on this is still at a relatively preliminary stage.

Therefore, this study aims to effectively utilize the advantages of THz radiation in response to crystalline materials and attempts to construct an evaluation system for cellulosic materials based on THz-TDS. To achieve the goal, five research topics were done as shown below:

1. By building an MC-dependent dielectric function of water in the THz region, the accuracy of the prediction model of density and MC of wood was significantly improved compared with the previous model.

2. Then, to predict the microfibril angle (MFA) which influences the stiffness of wood materials, a multiple linear regression model was built with parameters of the absorption and refractive index spectra in the THz region, although the prediction accuracy is limited, it still showed the possibility of predicting MFA in this region.

3. As mentioned above, to improve the prediction model of MFA, it is necessary to investigate the absorption model of cellulose, which is the basic unit of the microfibrils in the THz region. Cellulose I_α and I_β were investigated by THz-TDS, and the correlation between the relative content with absorption intensity was confirmed.

4. After this, the amount of cellulose crystalline was determined by THz-TDS with wood, microcrystalline (MCC) cellulose and pseudo-wood (MCC mixed with lignin), instead of the crystalline index calculated from XRD, which is a relative parameter.

5. Following that, the construction change of the native cellulose by a chemical treatment was investigated by THz-TDS.

In this thesis, after an introduction to the principle of THz radiation used in this study, the five topics above will be introduced and discussed separately.

2. THz radiation

As shown in Fig.2-1, THz lies in the far-infrared region, where THz spectroscopy can determine the optical properties of materials in this region as a function of frequency. This information, including absorption coefficients and refractive indices, can be used to gain insight into material properties for various applications as mentioned above. In this chapter, the common methods of THz generation and detection, and the principle of THz-TDS will be introduced.

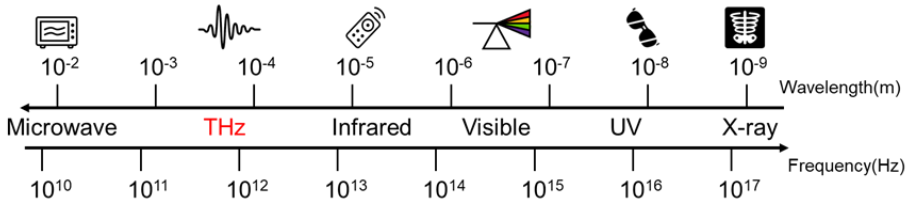

Fig. 2-1 Frequency and wavelength of electromagnetic radiation

2.1 THz generation

Most broadband pulsed radiation sources are currently produced by the excitation of semiconductor materials by ultrashort laser pulses. Photoconductive dipole antenna and optical rectification are the two most common methods to generate THz radiation.

2.1.1 Photoconductive dipole antenna

Using a photoconductive material (such as GaAs and ZnTe, and some organic crystals) as a radiation antenna, the semiconductor material is irradiated with an ultrashort pulse laser having photon energy larger than the forbidden bandwidth of the semiconductor, and an electron-hole pair is excited; the exciting free carrier is instantaneously accelerated by an external bias electric field. A transient current with a rapidly increasing current intensity is generated, and the stored electrostatic potential energy is released in the form of an electromagnetic pulse and propagates through the antenna to free space. When the pulse width of the excitation light pulse is on the femtosecond scale (10^{-15} sec), the pulse width radiated will be on the order of picoseconds (10^{-12} sec), that is, a THz electromagnetic wave (Mourou et al. 1981; Ferguson and Zhang 2002; Baxter and Guglietta 2011).

2.1.2 Optical rectification

Two beams will mix when they are in a nonlinear medium like a nonlinear crystal lattice, resulting in sum-frequency generation and difference-frequency generation, in the emergent radiation, there is not only the radiation that has the same frequency as the incident radiation but also new frequency

radiation (for example, sum-frequency radiation). Moreover, when a high-intensity monochromatic laser propagates in a nonlinear medium, it excites a constant (non-time-varying) electrode field through the difference-frequency oscillation effect inside the medium. A constant electrode field does not radiate electromagnetic waves but creates a DC electric field inside the medium. This phenomenon is called the optical rectification effect, which is the inverse of the electro-optical effect. While an ultrashort pulse laser is an incident on the nonlinear medium, a low-frequency electromagnetic pulse will radiate. When the pulse width of the incident laser is on the order of sub-picoseconds, the upper-frequency limit of the electromagnetic pulse radiated is on the order of THz since the upper-frequency limit of the radiated electromagnetic wave is related to the pulse width of the incident laser (Rice et al. 1994; Nahata et al. 1996; Baxter and Guglietta 2011).

Since optical rectification is to couple the incident beam power from the optical frequency to the THz band, this coupling efficiency is lower compared to the photoconductive antenna, on the other hand, the advantage of this method is that it can output a wider THz radiation bandwidth.

2.2 THz detection

At present, to detect THz radiation, the most commonly used methods are photoconductive detection and electro-optical sampling. Both methods record the time-domain waveform of the THz-radiated electric field signal and obtain the amplitude and phase-frequency distribution by Fourier transform (Baxter and Guglietta 2011).

2.2.1 Photoconductive detection

The photoconductive sample is based on the inverse process of the photoconductive emission mechanism, the sampling pulse excites free carriers in the photoconductive medium, and the THz electric field acts as a deflection electric field to cause the carrier movement to produce a current proportional to the THz pulse electric field. By changing the time delay of the detector beam and the THz pulse, the variation of the entire THz electric field with time can be obtained (Shen et al. 2004).

2.2.2 Electro-optical sampling

Electro-optical sampling is a detection technique based on linear electro-optical effects. The probe light and the THz radiation are simultaneously applied to an electro-optic crystal, the angle between the polarization direction and the propagation direction, the optical axis direction of the

electro-optical crystal, and appropriately select the modulation polarizer and compensator. Transient birefringence is caused in the crystal by the THz pulse, which affects the propagation of the laser sampling pulse in the crystal, that is, the polarization state of the probe light is modulated by the THz electric field. The entire THz electromagnetic radiation waveform information including amplitude and phase can be obtained by electro-optic sampling (Nahata et al. 1996).

2.3 Principle of THz-TDS

THz-TDS is a relatively new spectroscopy method of obtaining spectral information of both the amplitude and the phase by directly measuring the time domain waveform of the electric field and Fourier transforming it. Since THz-TDS can not only detect amplitude but also phase, the permittivity and the complex refractive index including the information of refractive index and absorption coefficient can be calculated from the time delay and intensity ratio between the reference and the sample spectrum. The spectrum contains physical and chemical information, and the complex refractive index and permittivity are related to the physical properties, which can be used for material characterization, and layer inspection (Baxter and Guglietta 2011).

Fig. 2-2 shows the basic configuration of the THz time-domain spectrometer. The pulse wave emitted from the femtosecond laser is split into pump light and probe light by a beam splitter. THz pulse is generated by irradiating the pump light in the THz emitter. The emitted THz pulse wave passes through the sample and reaches the detector. And the probe light passes through the time delay stage and reaches the THz detector. The detector is configured to detect the THz wave only when the probe light is received. A movable mirror is placed in the time delay stage, and its movement changes the optical path length. As a result, the timing at which the probe light enters the detector changes gradually so that the electric field strength in the time domain can be measured.

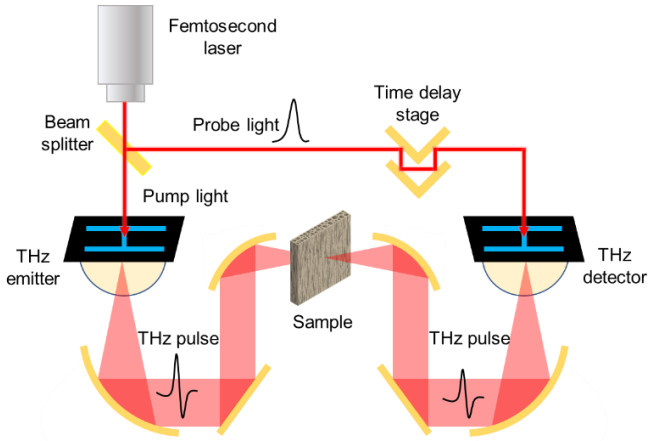

Fig. 2-2 Geometry of a THz time-domain spectroscopy

In this thesis, two different systems are used: T-Ray 4000 Time-Domain THz System (Picometrix LLC) in the chapter 3, and Tera Prospector (Nippo Precision Co., Ltd., with Frequency range: 0.1 to 4.0 THz Resolution: 6 GHz or less, S/N ratio: 2500:1 or more for time-domain signal, dynamic range: 5 digits or more) that using a Femtosecond laser as the light source (AS-20STD, IMRA America, Inc., with center wavelength: 780nm±5nm, plus width: 100 fs or less, power: 20mW or more) in the chapters 4 to 7. Both are photoconductive generation and detection types.

2.4 Analysis of THz data in transmission model

The time-domain signal measured by THz-TDS is shown in Fig.2-3 (a), all the THz time-domain signals were firstly Fourier transformed in the frequency domain to obtain the power spectrum and the phase spectrum are as shown in Fig. 2-3 (b) and (c), respectively. The data were analyzed by MATLAB to solve for the refractive index (n) and absorption coefficient (α) in the frequency domain.

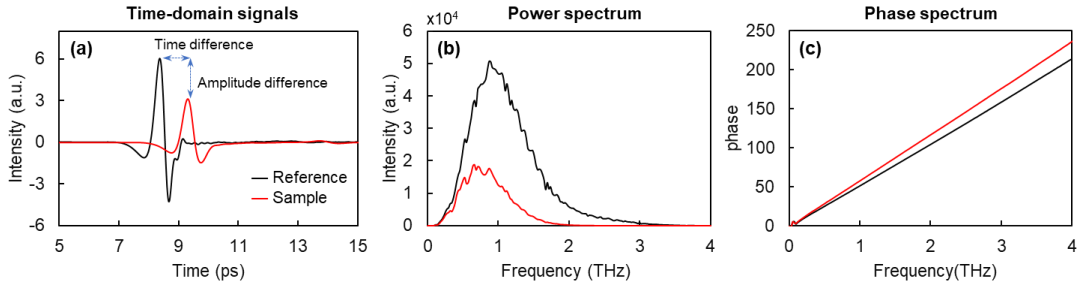

Fig. 2-3 (a) Time-domain signals of black: reference, red: sample, (b) the power spectrum and (c) the phase spectrum.

The complex-valued THz fields of the reference THz signal (E_{ref}) and THz signal transmitted through the sample (E_{sample}) can be expressed as:

$$E_{ref}(v) = E_{ref}e^{-i(\omega t+\varphi_{ref})} \quad (2-1)$$

$$E_{sample}(v) = E_{ref}t_{rs}t_{sr}e^{-i(\omega t+\varphi_{sample})} \quad (2-2)$$

The ratio of E_{ref} and E_{sample} is:

$$\frac{E_{sample}(v)}{E_{ref}(v)} = t_{rs}t_{sr}e^{-\frac{i2\pi L}{c}(n_{sample}-n_{ref})} \quad (2-3)$$

where *v* is the frequency, *c* is the speed of light in vacuum, *L* is the thickness of the sample, and t_{rs} and t_{sr} are the Fresnel transmission coefficients of the sample-to-air and air-to-sample interfaces, respectively. The Fresnel transmission coefficient is defined as:

$$t_{ij} = -\frac{2n_i}{n_i + n_j} \quad (2-4)$$

n_i is the index of refraction for substance *i*, where the refractive index of air (for the reference signals) is 1, therefore, the complex refractive index of the sample is defined as a sum of the refractive index (n_{sample}) and extinction coefficient (k_{sample}):

$$\tilde{n}_{sample} = n_{sample} + ik_{sample} \quad (2-5)$$

At last, the refractive index (n) and absorption coefficient (α) of the sample can be solved by combing these formulars together:

$$n = -\frac{\varphi c}{2\pi v L} + 1 \quad (2-6)$$

$$\alpha = -\frac{2}{L}\ln\left[R\frac{(n+1)^2}{4n}\right] \quad (2-7)$$

where φ is the phase difference between the reference and measured samples ($\varphi_{ref} - \varphi_{sample}$), *R* is the ratio of the amplitude in the frequency domain of the measured samples to the reference (Reid and Fedosejevs 2006). The samples used in this study were all ‘‘optically thick samples,’’ therefore the Fabry–Pérot effect, which is the backward and forward reflections in the sample can be ignored (Duvillaret et al. 1996).

3. An accurate prediction of moisture content by applying a dielectric function of water in THz region

3.1 Introduction

The changing of environmental humidity in a wood factory is the main reason cause MC changing in wood, wood absorbs and desorbs water periodically from the surrounding environment, which in turn leads to dimensional changes, and may even cause some cracking and warping, which need to be avoided. Therefore, determining the moisture content (MC) of wood material accurately is essential, not only to predict the physical properties of wood but also can be used to provide guidelines for humidity management, minimize seasonal fluctuations, and retain the integrity of the wood.

The behavior of water within wood is not simple, it changes as the MC changes. This is reasonable, MC changing leads to a change in the interaction of cell walls and water molecules through hydrogen bonding. There are two types of water in wood material, “free” water which exists in wood as liquid or vapor in the lumen, and “bound” water which exists in the amorphous region on the surface crystallization region of cellulose, hemicellulose, and cellulose by hydrogen bonding. The fiber saturation point (FSP) is defined as the MC at the transition point from the saturated state to the unsaturated state, that is, where lumens have no liquid water around an MC of 25–30% (Tiemann 1906). Therefore, the understanding of the behavior of the bound water is necessary for precise control of the drying process of wood products.

In previous work, using THz-TDS density and MC of wood were simultaneously determined (Inagaki et al. 2014a), However, the MC tends to be systematically underestimated in the prediction. In this study, a species-independent calibration is performed to determine the MC-dependence of the dielectric function and used to optimize the simultaneous prediction of density and MC below the FSP.

3.2 Method and materials

3.2.1 Sample preparation

Aspen (*Populus tremuloides*, hardwood), birch (*Betula spp*, hardwood), western hemlock (*Tsuga heterophylla*, softwood), and maple (*Acer saccharum*, hardwood) four species of the wood samples were cut from the radial plane with an approximate 5 mm (T) thickness and 30 mm (L) x 30 mm

(R) squares, where the wood grain is parallel to the edge, 15 polished samples for each species.

In the first step, all the samples were dried in a convection oven at 103 ± 2 °C for 48 hours to determine the oven-dry mass and weighed (±0.01 g) before and after the THz measurements to ensure that the samples did not absorb moisture from the atmosphere. The dimensions (length, width and thickness) of all samples were measured using a caliper (±0.01 mm) prior to THz measurement. After completing the THz measurement of the oven-dry wood, the samples were placed in a desiccator at different relative humidities (RH) to obtain different MC samples, and the RH was gradually controlled using a saturated salt solution. The salts used were: anhydrous calcium sulfate ($CaSO_4$, RH = 0%), potassium hydroxide (KOH, RH = 8%), potassium fluoride (KF, RH = 30%), sodium chromate ($Na_2Cr_2O_7 \cdot 2H_2O$, RH = 54%), ammonium chloride (NH_4Cl, RH = 77%) and potassium nitrate (KNO_3, RH = 94%) obtained from SIGMA-ALDRICH. 99% RH was obtained using a desiccator with water only. The samples were placed in the desiccator for at least 48 hours to ensure they reached an equilibrium MC. After equilibrium was reached, as determined by no change in mass, samples were taken out of the desiccators and their mass and dimensions were recorded. THz measurements were taken immediately for two THz field polarizations (parallel and perpendicular) relative to the visible grain of the wood as shown in Fig. 3-1, the THz measurements were very fast (10s for each measurement). And the masses of the samples were also recorded after the THz measurement to ensure that the samples did not absorb or desorb water. Samples were measured oven-dry, followed by 11% RH, and progressively higher RH's. Densities and MC were calculated from physical dimensions, the average mass of the samples and the mass of the oven-dry wood samples. The above process was repeated for each sample from low to high MC. Experiments and wood conditioning were carried out at room temperature (21 ± 1 °C).

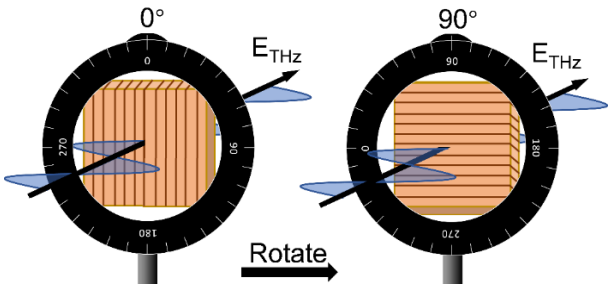

Fig. 3-1 Polarization of the THz beam and the manual rotation stage to switch the directions of sample. 0° and 90° are corresponding to polarizations perpendicular and parallel to the visible grain orientation

3.2.2 THz measurement

Polarized THz transmission spectroscopy was performed using a modified Picometrix T-Ray 4000 THz spectrometer. This system produces THz measurements in an 80 ps window at a rate of 1,000 waveforms per second. The bandwidth extends from approximately 0.01-2.00 THz. The emitter and detector are fiber-coupled photoconductive switches, generating linearly polarized THz radiation, where the polarization direction is defined by the orientation of the photoconductive antenna. The emitter and receiver photoconductive antennas are arranged with their linear polarization axis in the horizontal direction in the lab frame of reference. For each species of wood studied, the time-domain waveforms of the transmitted THz radiation were recorded at 0° and 90° orientations of the grain with respect to the THz polarization, Wood samples are mounted on a manual rotation stage to switch the directions with an accurate rotation (±0.25°), when the sample was rotated to 0°/90°, the grain of wood is perpendicular/parallel to the THz polarization as shown in Fig. 3-1. Each measurement being an average of 10,000 waveforms. A reference THz signal without a wood sample in place was obtained prior to each transmission measurement. The sample sizes were chosen to be larger than the THz beam diameter to avoid diffraction effects. The THz spectroscopy was performed in transmission, with a THz beam diameter (1/e electric field) of approximately 30 mm. The complex refractive index of the samples was determined from a transmitted and reference THz signals as described in chapter 2 (Duvillaret et al. 1996; Reid and Fedosejevs 2006). The dielectric function $\tilde{\varepsilon}$ is the relevant parameter in applying the effective medium theories (EMTs) described in this work can be determined from the refractive index n that calculated from Eq. (2-6):

$$\tilde{\varepsilon} = \varepsilon' + i\varepsilon'' = (n + i\kappa)^2 \qquad (3-1)$$

and the conversion between refractive index and dielectric constant is using the follow equations:

$$\varepsilon' = n^2 - \kappa^2 \qquad (3-2)$$

$$\varepsilon'' = 2n\kappa \qquad (3-3)$$

Where κ is the extinction coefficient can be calculated from the absorption coefficient α calculated with Eq. (2-7) and the wavelength λ:

$$\kappa = \frac{\alpha\lambda}{4\pi} \qquad (3-4)$$

3.3 Model construction

In this study, we followed the previous model (Inagaki et al. 2014a), to correct the systematic under underestimated in the predictions of MC and density, an MC-dependent dielectric function of water combing an effective medium theory (EMT) model was constructed. The EMT model calculates the properties of composite materials based on the properties of each component (Kirkpatrick 1971). In simple terms, wood can be thought of as a composite of cell walls, air, and water, therefore, when the dielectric function and the volume fraction of the three components are determined, the density and MC can be calculated from the measured dielectric function of wood. The scheme of this EMT model and the fitting process are shown in Fig. 3-2, and it will describe in detail in this section.

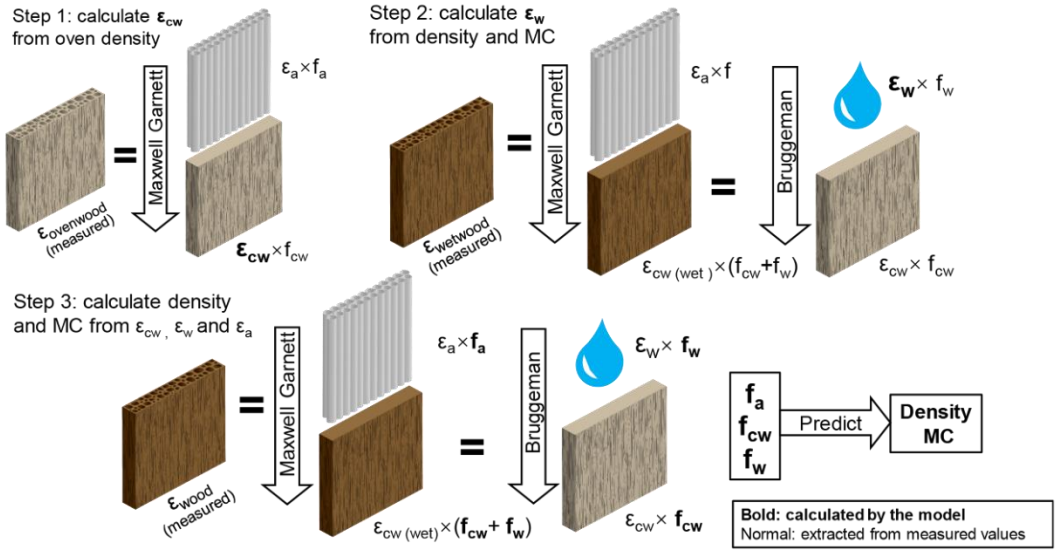

Fig. 3-2 Principle of the model. Where ε_{cw} can be determined by finding the best-fit of the measured dielectric function of oven dry wood sample, same as the first step, the ε_w can be determined as a MC-dependent function by finding the best-fit of the measured dielectric function of wet wood sample (discussed in detail in the results). After the calculation of ε_{cw} and ε_w, the volume fractions (air, cell wall, water) can be calculated to predict density and MC for wood sample simultaneously

3.3.1 Determining the dielectric function of cell wall

The first step, the dielectric function of the cell wall was determined. The geometry of oven-dry wood can be considered as being composed of cylinders of air in a background matrix consisting of the cell wall material. Using this concept, optical properties can be extracted from THz measurements using effective medium theory (Inagaki et al. 2014b), based on the composite

cylinder assemblage (CCA) formalism (Hashin 1983). In this formalism, the effective medium properties are determined by inserting an infinite cylinder with dielectric function ε_1 (in wood materials that is air) into a concentric cylindrical shell with dielectric function ε_2 (in wood materials that is cell wall substance). This geometry has a unique axial direction, and therefore depends on the polarization of the probing THz field.

For the case where the field is polarized perpendicular to the cylinder axis, an upper bound of the effective dielectric function in the CCA scheme is given by the Maxwell-Garnett (MG) formula:

$$\varepsilon_{eff}^{\perp,max} = \varepsilon_2 + \frac{f_1}{\frac{1}{\varepsilon_1 - \varepsilon_2} + \frac{f_2}{2\varepsilon_2}} \qquad (3-5)$$

and when the field is polarized along the cylinder axis, where the electric field is always parallel to the dielectric interface, it follows the simple mixing rule (Jepsen et al. 2007):

$$\varepsilon_{eff}^{\parallel} = f_1\varepsilon_1 + f_2\varepsilon_2 \qquad (3-6)$$

where ε_1 and f_1 corresponds to the dielectric function and the volume fraction of the infinite cylinder, with ε_2 and f_2 the dielectric function and the volume fraction of the cylindrical shell.

The dielectric function of cell wall material was calculated from the measured values of the dielectric function of oven-dry wood samples averaged using frequency a range of 0.1-0.2 THz. For the oven-dry wood, by using the MG-EMT model (Eq.(3-5) and Eq. (3-6)) for different polarizations, the dielectric function of cell wall material, which is the only free parameter can be simply calculated, since the volume fractions of air and cell wall can be calculated from density. The density of the cell wall substance is approximately constant, close to 1.4 g cm^{-3} (Mwaikambo and Ansell 2001), and the dielectric function of air is 1. Specifically, the dielectric function of cell wall material was determined by finding the best fit of the real and imaginary parts of it.

The best fit was defined as when the fitted dielectric function of cell wall material combined with the dielectric function of air can descript the measured value best, in other words, when the error between the predicted and the measured dielectric functions is minimized, and the error was determined by the sum of squared differences:

$$Err = \sum_i (\varepsilon_{pred} - \varepsilon_{meas})^2 \qquad (3-7)$$

Until now, the dielectric function of the cell wall of two polarizations can be extracted from the oven-dry wood THz spectra by using MG-EMT model.

3.3.2 The dielectric function of water: the Double Debye model and an MC-dependent dielectric function of water

For further prediction of the MC particularly below the fiber saturation point (FSP, no free water in the cell cavity and intercellular space) of wood materials, the dielectric function of water in wood is necessary. As the water in wet wood is randomly distributed in its cell wall material, the effective medium properties of wet wood can be determined by combining the dielectric function of water with the cell wall material which has been determined in the first step. With the dielectric function of water, the prediction of moisture content becomes feasible.

In the previous study (Inagaki et al. 2014a), the double Debye model appropriate for frequencies around 1 THz (Jepsen et al. 2007) was used to obtain the dielectric properties of the water in the cell wall:

$$\varepsilon(\omega) = \frac{\varepsilon_s - \varepsilon_1}{1 - i\omega\tau_1} + \frac{\varepsilon_1 - \varepsilon_\infty}{1 - i\omega\tau_2} + \varepsilon_\infty \qquad (3-8)$$

where the parameters for Eq. (3-8) were taken as presented in the report of Jepsen et al., ε_s = 78.36, ε_1 = 5.16, τ_1 = 7.89 ps, ε_∞ = 3.49, and τ_2 = 0.181 ps.

As mentioned above, the wet wood can be considered as water randomly distribute in cell wall material, thus the Bruggeman EMT model seems to be suitable to obtain an effective medium approximation for the hydrated cell wall, since the Bruggeman model has used to model the water within paper which is a cellulose fiber network very well (Reid and Fedosejevs 2006). The Bruggeman EMT is given by:

$$\varepsilon_{eff} = \frac{1}{4}\left(\beta + \sqrt{\beta^2 + 8\varepsilon_1\varepsilon_2}\right) \qquad (3-9)$$

$$\beta = (3f_1 - 1)\varepsilon_1 + (3f_2 - 1)\varepsilon_2 \qquad (3-10)$$

In this process, the same as the dielectric function of the cell wall, the dielectric function of wet cell wall material was calculated from the measured values of the dielectric function of wet wood samples averaged using a frequency range of 0.1-0.2 THz. The dielectric functions of water and cell wall material can be combined to represent the dielectric function of the wet cell wall by using the Bruggeman EMT model (Eq. (3-9) and Eq. (3-10)), with the dielectric function of cell wall material determined in the first step, which are 3.20 + 0.17i and 3.27 + 0.59i for parallel and perpendicular polarizations with respect to the THz filed, respectively.

However, the dielectric function of water would vary with MC, especially for MCs below FSP.

Since when the wood sample at a low-MC level, the hydrogen bond between water molecules and cellulose and hemicellulose can be very strong, and as the MC increases, more water is bound to the state close to the liquid water until it reaches the FSP. Thus, by considering the actual dielectric function of the water in the wood, which is a function depending on the MC, rather than the application of the previously used double Debye model, which represents the application of bulk liquid water, a more accurate model can be expected. Therefore, to solve the problems in the previous model that caused by assuming the dielectric function of water using the double Debye model, and improve the accuracy of the previous model, it is expected that the MC-dependent dielectric function of water should be applied. As that functional dependence is currently unknown from a first-principles perspective, the idea of finding the best-fit like the determination the dielectric function of cell wall was followed, which means the dielectric function of water would be assumed by a fitting model of water. The best fit defined in Eq. (3-7) of the dielectric function of water was determined based on the real and imaginary parts by minimizing the error between the predicted and measured dielectric functions using nonlinear least-squares curve fitting by the model using the measured parameters (the volume fractions of components from density and MC). With the MC-dependent dielectric function of water, the dielectric function of the wet cell wall can be expressed with Eq. (3-9)

3.3.3 Combination of the dielectric functions of wood components

At the last step, the effective medium properties of wood materials can be described by combining the frequency-independent value obtained for the dielectric function of the cell wall material, and the dielectric function of water (the Double Debye model and an MC-dependent one). However, to predict the density and MC of wood simultaneously, achieving the volume fractions f_a and f_w is indispensable, similar to the calculation of the model described above, the nonlinear least-squares curve fitting also used to give the most suitable volume fractions to the predicted dielectric function of each part (air, water, cell wall), since the dielectric functions of air, water and cell wall are all known, the only free parameters for the dielectric function model are the volume fractions of the components, by inserting the volume fractions that calculated from the nonlinear least-squares curve fitting into the effective medium approximation, so that the overall prediction of the dielectric function is closest to the measured value.

With the volume fractions f_a and f_w, density can be converted by the formula:

$$\rho_{wetwood} = \rho_{air} \times f_a + \rho_{ovencw} \times f_{ovencw} + \rho_w \times f_w \quad (3-11)$$

where $\rho_{wetwood}$ is the density of wet wood, ρ_{air} is the density of air at 20 ℃ which 0.0012 g cm^{-3}, ρ_{ovencw} is the average density of the oven-dry cell wall material, which is assumed to be 1.4 g cm^{-3} (Mwaikambo and Ansell 2001), ρ_w is the density of water which is 1 g cm^{-3}, and. f_{ovencw} is the volume fraction of the oven-dry cell wall material (f_{ovencw} = 1- (f_a + f_w)).

And MC were calculated by:

$$\rho_{ovenwood} = \rho_{air} \times (f_a + f_w) + \rho_{ovencw} \times f_{ovencw} \quad (3-12)$$

$$MC = \frac{(\rho_{wetwood} - \rho_{ovenwood})}{\rho_{ovenwood}} \times 100(\%) \quad (3-13)$$

where the $\rho_{ovenwood}$ is the density of the oven-dry wood, MC is the moisture content of wood materials.

The prediction ability and stability of the model are evaluated by two parameters: determination coefficient and root-mean-square error (RMSE):

$$R^2 = 1 - \frac{\sum(y - y_{pred})^2}{\sum(y - \bar{y})^2} \quad (3-14)$$

$$RMSE = \sqrt{\frac{\sum(y - y_{pred})^2}{n-2}} \quad (3-15)$$

where, y is the measured value, y_{pred} is predicted value that calculated from THz, $\bar{y}$ is the average of measured value and n is the number of samples.

3.4 Result and discussion

3.4.1 Simultaneous prediction of wood density and MC using the Double Debye model as the dielectric function of water

The previous study did the prediction by minimizing the error of measured and calculated dielectric functions of wood, and the dielectric function of water used the Double Debye model, which represents the bulk liquid water. The prediction results of density and MC are shown in Fig. 3-3 and Fig. 3-4 respectively.

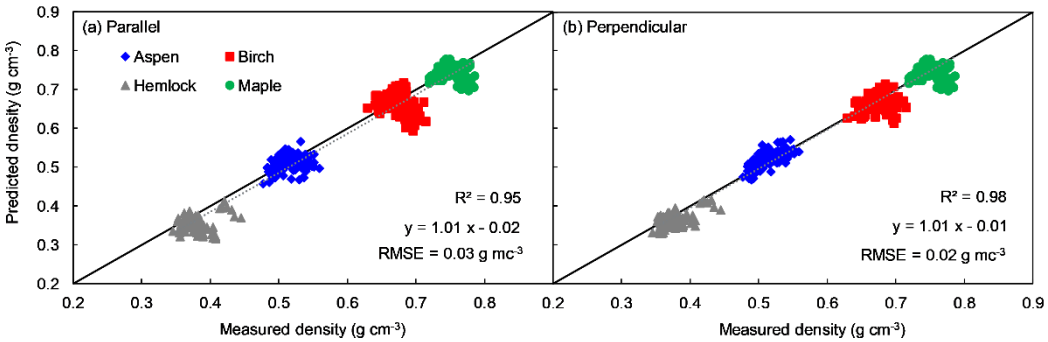

Fig. 3-3 Relationship between measured and predicted density based on the previous model. R^2, determination coefficient; RMSE, root mean square error. (a) Parallel and (b) perpendicular refer to the polarization of the THz filed with respect to the grain of the wood samples

Fig. 3-3 shows the relationship between the measured and predicted densities, the determination coefficient values R^2 of 0.95 and 0.98, and RMSE values of 0.03 g cm^{-3} and 0.02 g cm^{-3} for parallel and perpendicular polarizations, respectively. The model provides good prediction at relatively low MCs for both polarizations, and consistently underestimates the real density at higher MCs, where the data points that lie furthest from the curve appear at relatively high MC for all specie. The result is reasonable since wood samples contain more water as the MC increases, which leads to more significant deviations as the improper dielectric function is being used as argued above. Another contributing factor is that the S/N ratio is lowest when the sample density and MC are the highest.

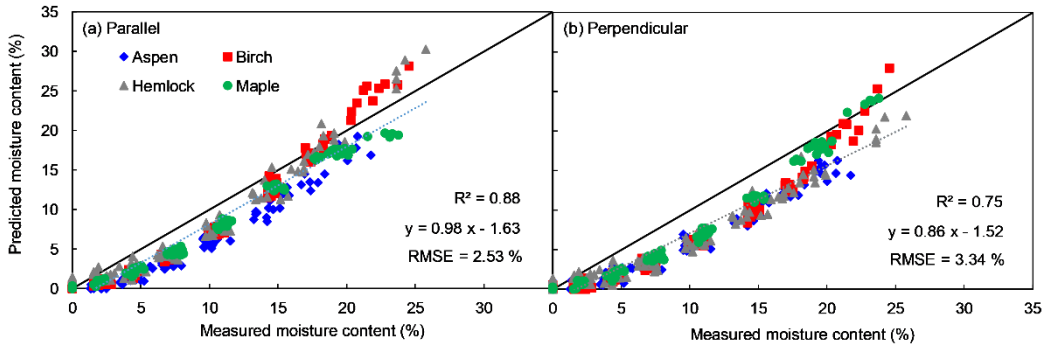

Fig. 3-4 Relationship between measured and predicted moisture content based on the previous model. R^2, determination coefficient; RMSE, root mean square error. (a) Parallel and (b) perpendicular refer to the polarization of the THz filed with respect to the grain of the wood samples

Fig. 3-4 shows the relationship between measured and predicted MCs, which gives determination coefficient values of R^2 0.88 and 0.75, with RMSE values of 2.53 % and 3.34 % for parallel and perpendicular polarizations, respectively. The predicted MCs tend to be underestimated at low MC, and gradually approach the measured value as MC increases. This can again be attributed to the behavior of the dielectric function of the water within the wood. Specifically, due to the behavior of

the dielectric function of water where it is expected that free water absorbs more strongly than bound water, the use of a highly absorbing free water dielectric function at a low MCs will result in a systematically underestimated value of the MC. And as MC increases, the behavior of water molecules progressively approaches that of free water as argued in the methods section, and the predicted MCs will approach the actual values more closely.

3.4.2 Relation between dielectric function and oven-dry density and MC

As shown in Fig. 3-5, the measured frequency-resolved real part (a) and imaginary part (b) of the dielectric function that was obtained from THz transmission spectroscopy of oven-dry wood, the real part of the dielectric function is linear with respect to density, whereas the imaginary part of the dielectric function increases with increasing density. And there is a relationship between the real and imaginary parts of the measured dielectric functions of the sample averaged over the frequency range of 0.1 to 0.2 THz and the oven-dried density for 4 species as shown in plots (c) and (d), where the THz beam was polarized perpendicular to the visible grain of the wood samples. As shown in Fig. 3-5 a strong correlation between oven-dry density and the dielectric function is observed that is consistent with the other research (Tanaka et al. 2014).

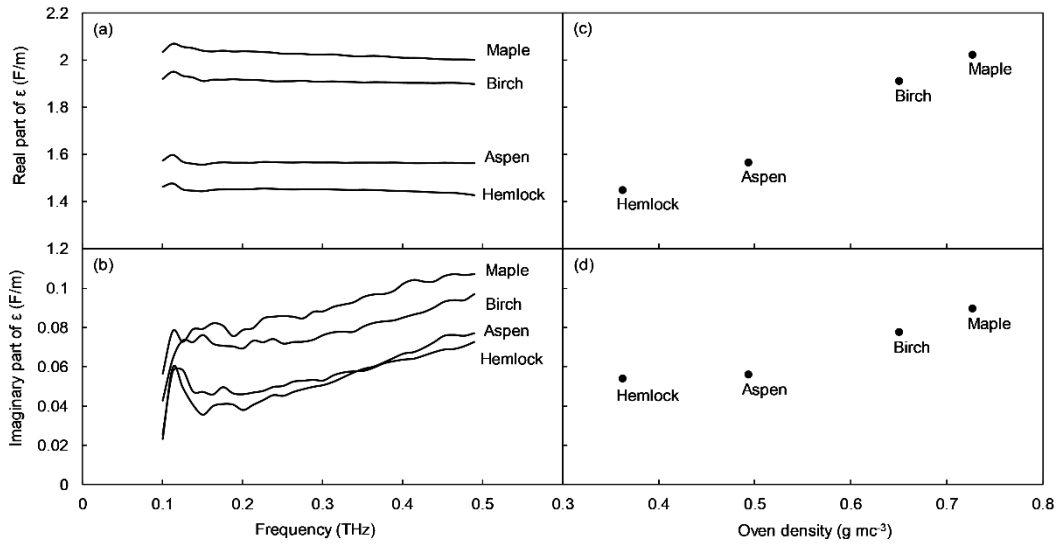

Fig. 3-5 Frequency-resolved real part (a) and imaginary part (b) of the measured dielectric function of oven-dry wood samples. Relationship between oven-dry density and the real part (c) and imaginary part (d) of the dielectric function averaged over the frequency range of 0.1–0.2 Tz is shown on the right where the THz beam was polarized perpendicular to the grain of the wood samples in these measurements

Fig. 3-6 shows a clear linear relation and a quadratic relation between MC and the real and imaginary parts of the dielectric function for the 4 species of wood investigated in this study, where

the THz beam was polarized perpendicular to the visible grain of the wood samples. The real and imaginary parts of the dielectric function show obvious correlations with MC for all 4 species.

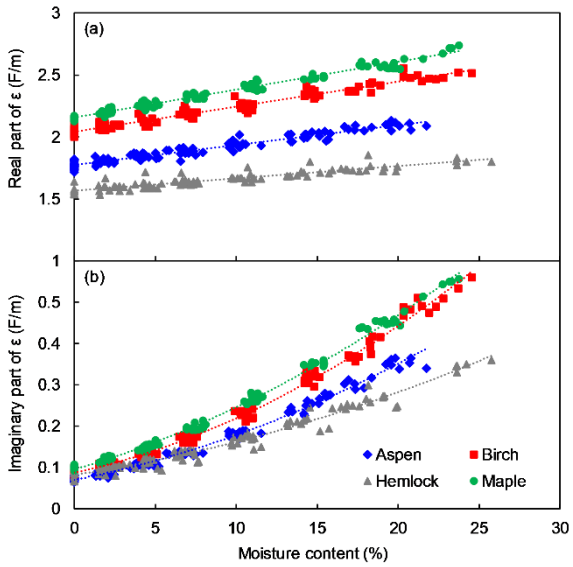

Fig. 3-6 Real and imaginary parts averaged over the frequency range of 0.1-0.2 THz of the measured dielectric function for 4 species of wood as it depends on MC, with the THz beam polarized perpendicular to the grain of the wood samples.

The same as that in Fig.3-5, a more specific relationship is highlighted in Fig. 3-7, plots (a) and (b) show the measured frequency-resolved real part and imaginary part of the dielectric function that was obtained from THz transmission spectroscopy birch wood with different MCs, respectively. Plots (c) and (d) show a strong correlation between the real and imaginary parts of the dielectric function averaged over the frequency range of 0.1 to 0.2 THz and MCs for birch wood. The real part of the dielectric function is relatively linear with respect to MC, whereas the imaginary part of the dielectric function increases quadratically with MC. It is interesting to note this correlation, which again implies that the imaginary part of the dielectric function of bound water is a function of MC as argued in the introduction section.

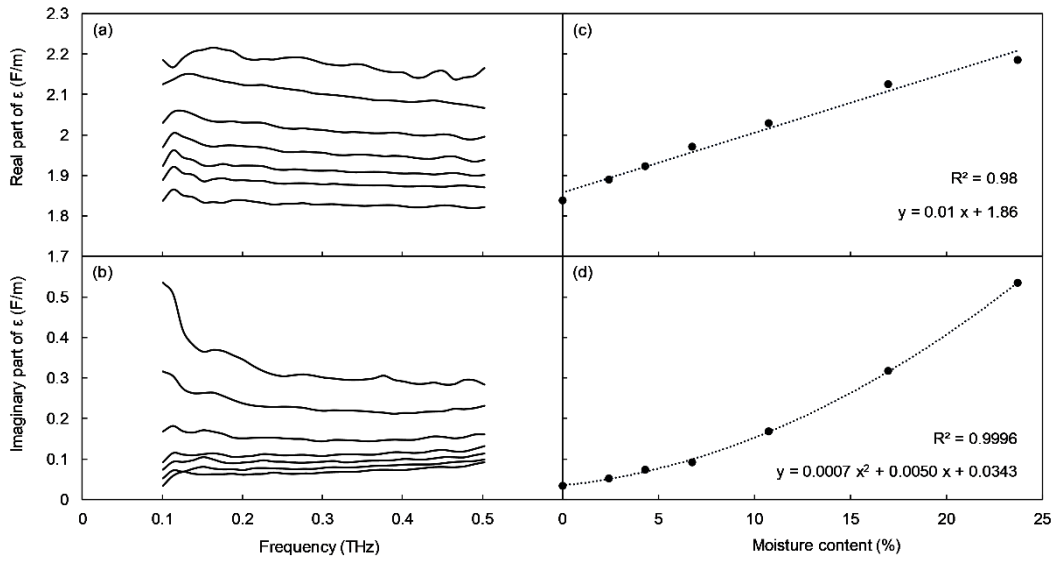

Fig. 3-7 Frequency-resolved real part (a) and imaginary part (b) of the measured dielectric function of birch at various MCs. Relationship between MC and the real part (c) and imaginary part (d) of the dielectric function averaged over the frequency range of 0.1-0.2 THz the with THz beam polarized perpendicular to the grain of the wood samples

3.4.3 Calculation of MC-dependent dielectric function of water

As mentioned above, we estimated that the dielectric function of water is MC-dependent, because as the MC in the wood increases, the molecular dynamics of water tend to shift from bound water to free water. Fig. 3-8 shows the relationship between the volume fraction of water and the best fit real and imaginary parts of the dielectric function of water averaged over the frequency range of 0.1-0.2 THz, with the THz radiation polarized parallel and perpendicular to the visible grain of the wood samples. The determination value, R^2, and significant difference value P are also shown in Fig. 3-8, the real and imaginary parts of the dielectric function for both polarizations show extremely small P values, implying the correlations are reliable. It can be seen that for the real part of the dielectric function, as the volume fraction of water increases, the dielectric function has a decreasing trend, while for the imaginary part of the dielectric function, the opposite trend is shown. Notably, the dielectric function of water is roughly species-independent, as reflected by the scatter around the fitted line, and thus can be used for essentially any wood sample, just like the dielectric function of cell wall materials. Only volume fractions of water > 0.02 are used here because using a very small volume fraction value in extracting the dielectric function of water can lead to large errors, since a very small volume fraction means almost no water in the sample, the information of water will be non-extractable. This is also why the dispersion of the data increases as the volume fraction

of water decreases, as shown in Fig. 3-8, hemlock is the only species with an uneven distribution of the dielectric function around the fitted line, the real part of the dielectric function of water is lower than that of other species, while the imaginary part is higher, suggesting that the behavior of water in hemlock may be different from other species, which may cause by that hemlock was the only softwood in our sample.

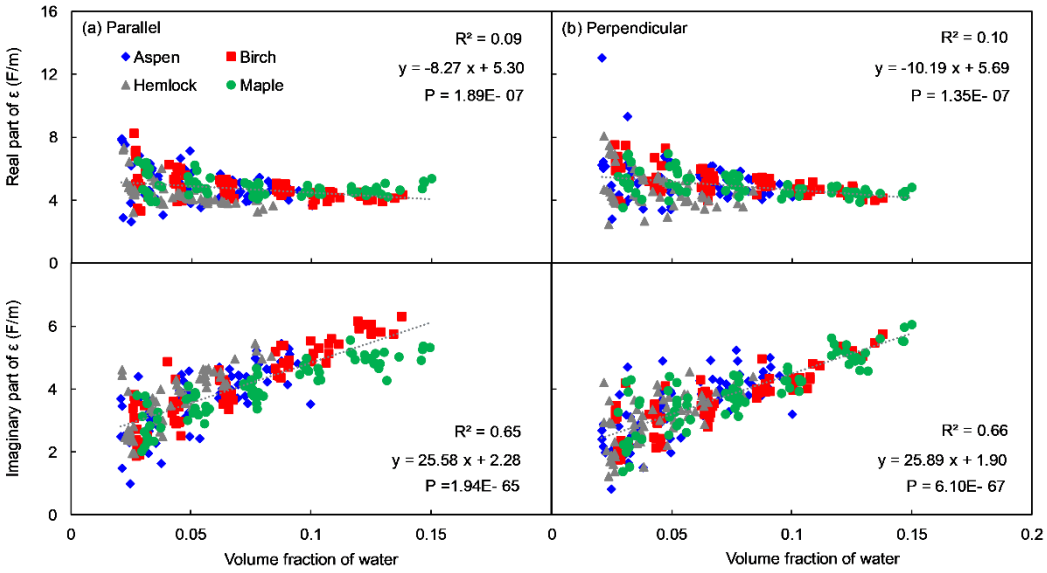

Fig. 3-8 Relationship between the volume fraction of water and real and imaginary part of dielectric function averaged over the frequency range of 0.1-0.2 THz. R^2, determination coefficient; P, significant difference. (a) Parallel and (b) perpendicular refer to the polarization of the THz radiation with respect to grain of the wood samples

As outlined in the methods section, the best-fit dielectric function of water was determined based on the dielectric functions of air and cell wall material, with independently determined volume fractions from physical measurements (density and MC), that gave the best fits to the experimental data. With the correlation between the best fit of real and imaginary parts of the dielectric function of water and the volume fraction of water, the dielectric function of water can be written as linear regression for the real and imaginary parts of the dielectric function separately, as a function of the volume fraction of water for the two different polarizations:

$$\varepsilon_w^{\parallel} = -8.3 \times f_w + 5.3 + (26 \times f_w + 2.3)i \qquad (3-16)$$

$$\varepsilon_w^{\perp} = -10 \times f_w + 5.7 + (26 \times f_w + 1.9)i \qquad (3-17)$$

3.4.4 Improvement in simultaneous prediction of wood density and MC

The prediction process is the same as that in 3.4.1, only the dielectric function of water was replaced from the Double Debye model by the MC-dependent dielectric function of water Eq. (3-

16) and Eq. (3-17). The only remaining free parameters are the volume fraction of water and air (f_w and f_a), as the volume fraction of cell wall material can be calculated from these as: (1- ($f_w + f_a$)). To extract f_w and f_a, which gives the predictions for density and MC from the measured THz spectra by using Eq. (3-11) and Eq. (3-13).

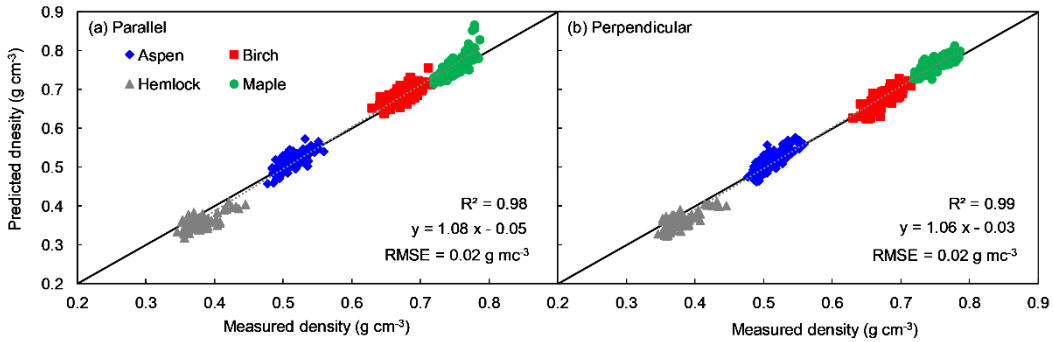

Fig. 3-9 Relationship between measured and predicted density based on the new MG-EMT model. R^2, determination coefficient; RMSE, root mean square error. (a) Parallel and (b) perpendicular refer to the polarization of the THz filed with respect to the grain of the wood samples

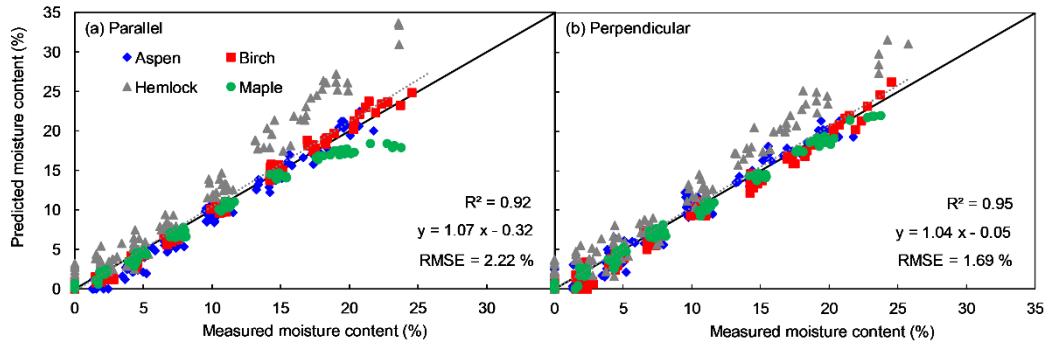

Fig. 3-10 Relationship between measured and predicted value of the moisture content based on the use of a MC-dependent dielectric function of water. R^2, determination coefficient; RMSE, root mean square error. (a) Parallel and (b) perpendicular refer to the polarization of the THz filed with respect to the grain of the wood samples

Fig. 3-9 shows the relationship between the measured and predicted densities for both polarizations using the model with the new MC-dependent dielectric function of water. The new model provided excellent predictions, with coefficients of determination R^2 of 0.98 and 0.99, and RMSE values of 0.02 g cm^{-3} and 0.02 g cm^{-3} for parallel and perpendicular polarizations, respectively. The prediction improves the accuracy compared to the previous results in Fig. 3-3, and successfully reduces the bias of the predicted values in the high MC range.

Fig. 3-10 shows the relationship between the measured and predicted MCs by using the MC-dependent dielectric function of water. For both parallel and perpendicular polarizations, a

significant improvement in the prediction accuracy of MCs is observed, with determination coefficient values, R^2, of 0.92 and 0.95 and RMSE values of 2.22 % and 1.69 %. Compared to the previous results, the accuracy of prediction has improved by a factor of 2, and the problem of systematically underestimating MC as shown in Fig. 3-4 has been resolved. It can therefore be concluded that the systematic deviation in MC prediction present in Fig. 3-4 can be attributed to the use of the double Debye model of bulk liquid water for the dielectric function of all water within the wood. However, the predictions for hemlock are systematically under-predicted for density and over-predicted for MC, which is certainly related to the deviation of MC-dependent fit of the dielectric function of water as shown in Fig. 3-8, and the possible reason was argued in advance, where water in hemlock behaves differently from other species, it is probably caused by the different structure and chemical components since hemlock is the only softwood species in our model.

This is the first attempt to speculate on the dielectric function of bound water in wood in the THz region, which may also help us to understand the behavior of bound water in wood and the process of wood drying at the molecular dynamics level.

3.5 Conclusions

The results of the previous prediction suffered from systematic errors that were attributed to using of a Double Debye model which is a dielectric function for bulk water, however, when water in wood at low MC undergoes hydrogen bonding within the cellulose network and behaves as bound water. In this study, the dielectric function of water in wood materials is considered to vary with the MC of the wood. The MC-dependent dielectric function is determined by doing the calibration, and the MC-dependent dielectric function of water within the wood is species-independent. Inserting this dielectric function of water in the model resulted in a significant improvement in the accuracy of the model for both density and MC predictions and rectified the systematic errors.

4. Prediction of the properties of cedar by using THz-TDS

4.1 Introduction

Japanese cedar (*Cryptomeria japonica*) is a large evergreen tree, that has the characteristics of a large diameter, rapid growth, and a straight stem, that can be used as wood materials effectively. It is widely used in Japan and has important economic value. In chapter 3, we have demonstrated the possibility of using THz-TDS to predict the density and MC of wood material with good accuracy. Thus, we further focused not only on density but also on MOE, cellulose crystallinity, and microfibril angle (MFA) of cedar by using THz-TDS. MFA which is the angle between the longitude direction and the microfibrils in S2 layer of wood cell wall, since the S2 layer contributes about 85% or more of wood cell wall thickness, MFA is strongly related to the physical properties of wood. However, the methods of measuring of MFA are complex, generally ways are iodide staining, polarization microscopy, and X-ray diffraction (Barnett and Bonham 2004). The physical and mechanical properties of wood are strongly dependent on these parameters and have been studied by some techniques such as bending test (for MOE), and X-ray diffraction (XRD, for MFA and cellulose crystallinity). As the strong birefringence showed by wood material in the THz region birefringence (Todoruk et al. 2012), which is the form birefringence and the intrinsic birefringence, caused by the cell wall alignment for the former, and cellulose crystalline zone for the latter, respectively. Which may be used to analyze the cellulose crystallinity and MFA. In this study, we investigated the possibility of predicting the MOE, MFA, and cellulose crystallinity of cedar by using multiple linear regression (MLR) model built with the absorption coefficient and refractive index extracted from the THz region.

4.2 Method and materials

4.2.1 Sample preparation

One single Japanese cedar log was cut from bark to bark and through the pith, size about 64 mm (L)×30 mm (T)×400 mm (R). The log was cut in same two bars along the R direction, the two bars matched by annual rings that make them have the same properties. One of them was sent to measure density, MOE, MFA, crystallinity, etc. as the reference values by a Silviscan analysis system (FPInnovations, Canada) which consists of three primary components based on three principles: image analysis (optical microscopy), X-ray attenuation (densitometry), and X-ray diffraction

(diffractometry). Particularly, the MOE measured by Silviscan is obtained by indirect calculation as follow:

$$MOE = A(I_{cv}D)^B \qquad (4-1)$$

The I_{cv} includes the scattering from S2 layer and the background scattering from other cell wall constituents such as the S1 and S3 layers, parenchyma, and amorphous cellulose and lignin present in the fiber wall. The model contains two statistically determined calibration constants (A and B), that have been shown to be insensitive to species, and relate to the sonic resonance method used for calibration. And another one was cut into 57 pieces along the L direction of the approximate size of 30 mm (L) ×30 m (T)×3 mm (R) as samples, the wood grain is parallel to the edge. All samples were set in a room with 20℃ ± 1℃ for 48 hours to an equilibrium MC, before the THz measurement, the three-dimensional sizes of samples were recorded using a caliper (±0.01 mm), and the mass was weighted with an electronic balance (±0.001g).

4.2.2 THz-TDS measurement

After recording the size and mass of samples, a transmission measurement was taken by a THz time-domain spectrometer (Tera Prospector-Kit model, Nippo Precision Co,. Ltd.) through LT-plane of samples, the samples were rotated from 0° to 180° with an interval of 5° (0° and 90° are defined as the polarization perpendicular and parallel respectively, to the grain orientation). Reference signals were taken twice before and after the measurement without a sample in the beam path, respectively. A total of 39 original spectra were measured for each sample. The workflow of the experiment is shown in Fig. 4-1.

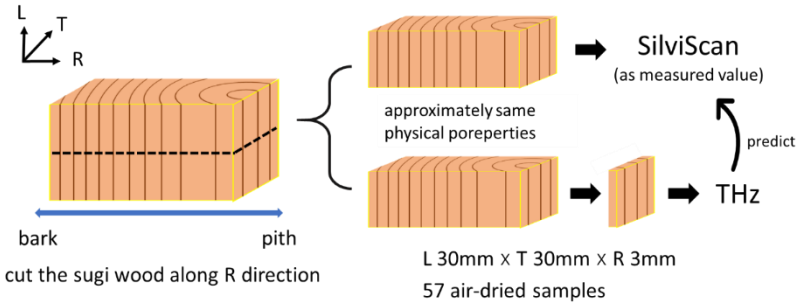

Fig. 4-1 The workflow of the experiment

4.3 Model construction

4.3.1 Birefringence of wood

The refractive index and absorption coefficient changes with the polarization, which means the

optical properties of the material are anisotropic, which is said birefringence. Incident light is decomposed into two waves of light having different polarization directions in the material. In the THz region, wood material shows large birefringence (Todoruk et al. 2012). As shown in Fig. 4-2. It can be seen that the detection time of peak and intensity change as the angle θ between the polarization direction of the THz and the wood fibers changes, which means the refractive index and absorption coefficient also changes at the same time. The detection time of the peak of 90° (the wood fiber is parallel to the THz polarization) is longer than 0° and 180° (the wood fiber is perpendicular to the THz polarization). As the fiber axis is a slow axis and the axis that is perpendicular to the fiber orientation is the fast axis, thus the THz wave arrives earlier. Particularly, as shown in Fig. 4-3, the refractive index changes with the angle between wood grain with THz polarization (the rotation angle) as a sine function approximated relation, where the difference between maximum and minimum of refractive index is termed as the birefringence. The large birefringence shown by wood material comes from two different sources (Todoruk et al. 2008), the form birefringence and the intrinsic birefringence, which is caused by the regular alignment of the wood cell wall and due to the regular alignment of the wood cell wall and the cellulose crystalline zone in the cell wall. Therefore, THz radiation can be used to probe the gross physical structure and alignment of the wood fibers, the cellulose crystallinity, and the MFA.

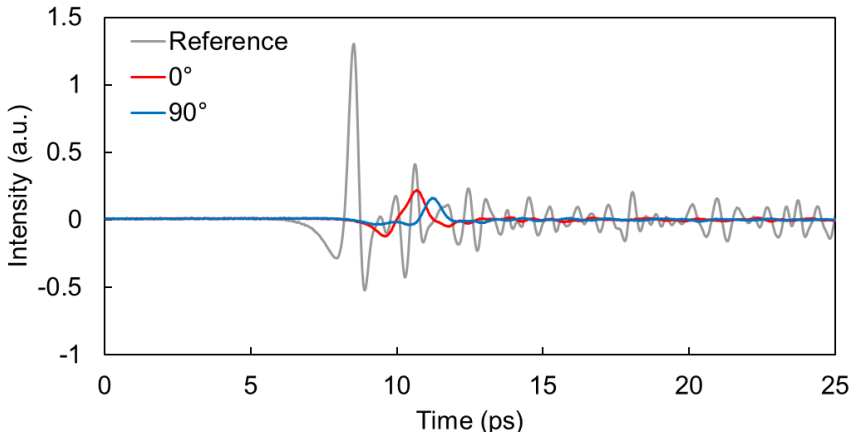

Fig. 4-2 THz transmission spectrum change as angle θ between the polarization direction of the THz and the wood fibers

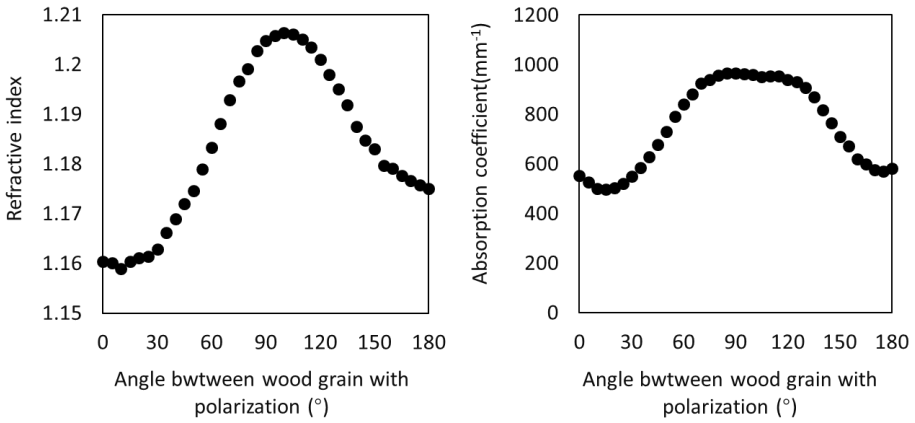

Fig. 4-3 Correlations of refractive index (left) and absorption coefficient (right) with the angle between wood grain with THz polarization.

4.3.2 Multiple linear regression model

The different time delays of peaks that were observed between reference and measured pulses of THz, which mainly caused by the pulses propagating parallel and perpendicular to the grain orientation. After the Fourier transformation, the refractive index and absorption coefficient averaged from a THz frequency of 0.2 THz to 0.4 THz by using Eq. (2-6) and Eq. (2-7). Three MLR models were built by using maximum, minimum values of the refractive index and the corresponding absorption coefficient, and the averaged values of the refractive index and absorption coefficient were used as explanatory variables, respectively. Density, MOE, MFA, and cellulose crystallinity were obtained from Silviscan as the objective variables with multiple linear regression analysis.

4.4 Results and discussion

4.4.1 The properties of cedar measured by Silviscan

The samples for the THz measurement are approximately 3mm, whereas, the reference data measured by Silviscan are very detailed (0.025mm/one measurement), since the samples for Siliviscan and THz measurements were cut from one single log, which means the boundaries (form the bark side) are in the same position, therefor the data were averaged simply by the sample thickness of the THz measurement for representing the properties of the wood sample.

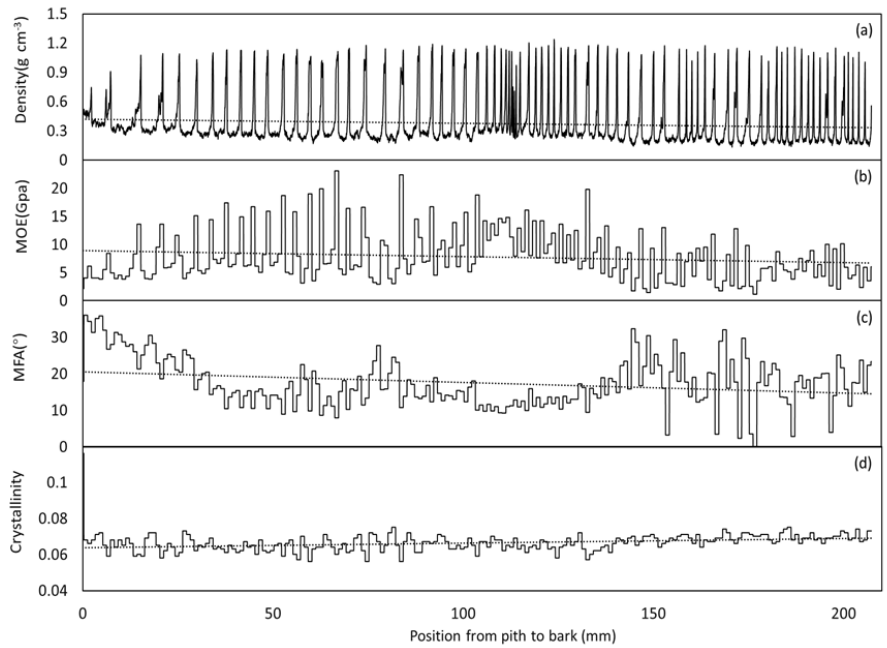

Fig. 4-3 Correlations of (a) density, (b) MOE, (c) MFA, and (d) cellulose crystallinity with the position from pith to bark obtained from Silviscan

Fig. 4-3 shows the correlations of (a) density, (b) MOE, (c) MFA, and (d) cellulose crystallinity with the position from pith to bark, which was obtained from Silviscan. Where the figure of density shows a lot of sharp peaks where is the late wood with a higher density than early wood in an annular ring, which means with the density is higher in bark than pith, particularly, different from other positions, near 110 mm, the figure of density shows some dense peaks, which can be considered caused by the reaction wood. MOE shows a tendency of decreasing with the position near to bark, and MFA also shows the same tendency as MOE, on the other hand, cellulose crystallinity does not have a significant change with growth.

4.4.2 Predictions of density, MOE, MFA and cellulose crystallinity by THz-TDS

As described in advance, the maximum, minimum, and averaged values of the refractive index and the corresponding absorption coefficient were used as explanatory variables, density, MOE, MFA, and crystallinity were obtained from Silviscan as the objective variables with MLR models to do the predictions. The results are shown as follows.

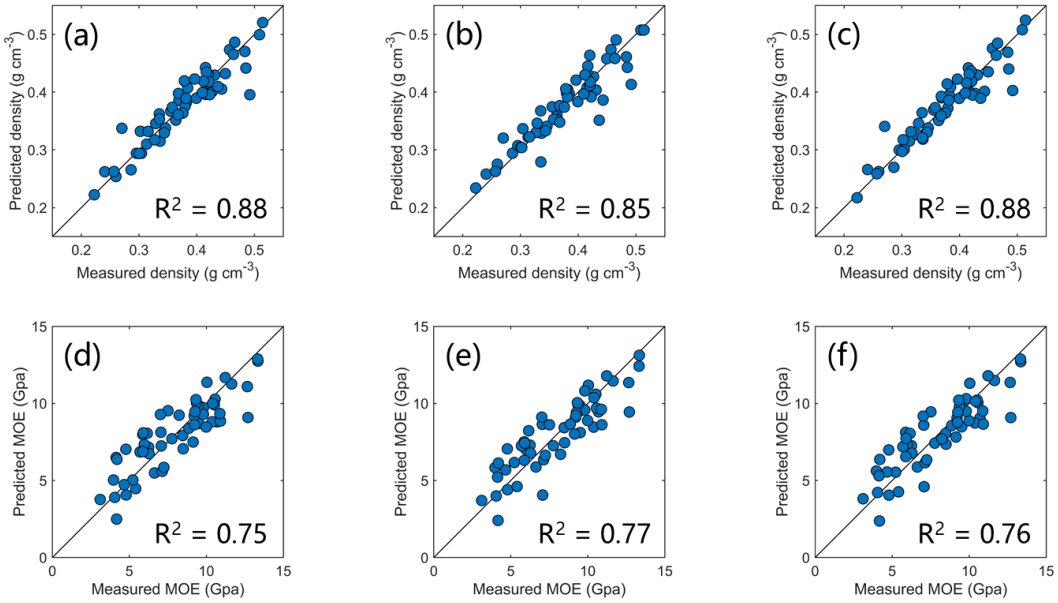

Fig. 4-4 Prediction of density (a)-(c) and MOE (d)-(f) using MLR models, where (a) and (d) using the maximum refractive index and the corresponding absorption coefficient; (b) and (e) using the minimum refractive index and the corresponding absorption coefficient, (c) and (f) using the averaged values of the refractive index and the absorption coefficient as explanatory variables. R2, determination coefficient.

Fig. 4-4 shows the correlation between measured density obtained from Silviscan and predicted density using different values of refractive index and absorption coefficient. All prediction shows a relatively high accuracy with R^2 of (a) 0.88, (b) 0.85, (c) 0.88. And the correlation between the measured MOE and the predicted MOE is also shown in Fig.4-4 with relatively high accuracy with R^2 of (a) 0.75, (b) 0.77, and (c) 0.76.

Fig. 4-5 shows correlation between measured MFA and predicted MFA using different values of refractive index and absorption coefficient. The determination coefficients R^2 of (a) 0.24, (b) 0.33, (c) 0.36. And the predictions of the cellulose crystallinity are shown in Fig. 4-5 (d)-(e) using different MLR models with R^2 of 0.44, 0.42, and 0.43, respectively.

It has been demonstrated that THz-TDS is able to predict the density and MC in our previous study, however, the EMT model is complicated, in this study, we applied a simple MLR model, and

the prediction accuracy of density and MOE are acceptable. Different from the predictions of density and MOE, the predictions of MFA and crystallinity show unreliable results with low R^2, which means to predict MFA and cellulose crystallinity, further research is necessary, and more details will be discussed in the next part.

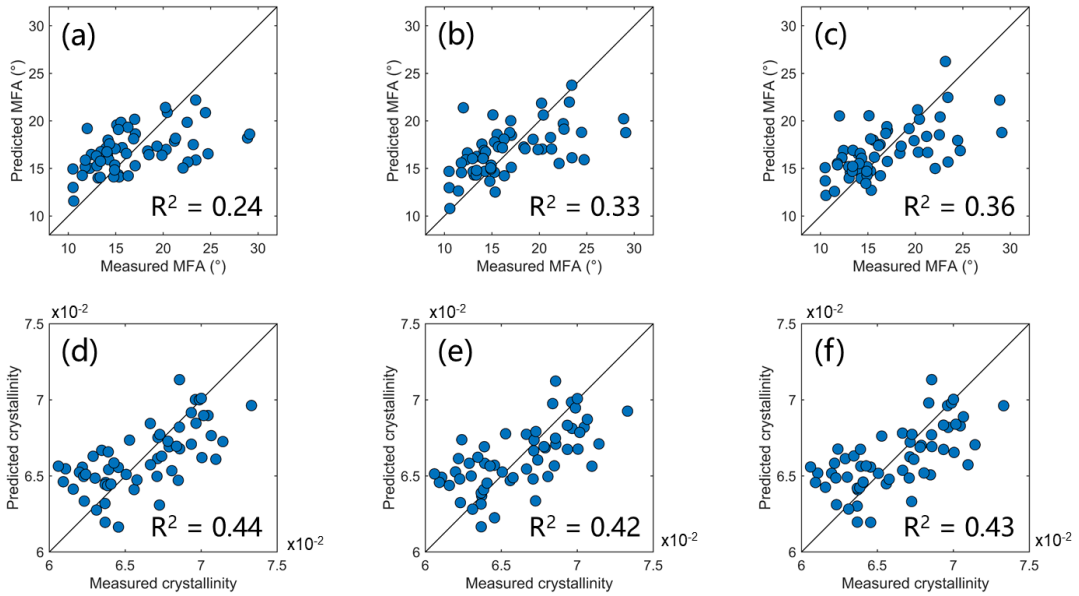

Fig. 4-5 Prediction of MFA (a)-(c) and cellulose crystallinity (d)-(f) using MLR models, where (a) and (d) using the maximum refractive index and the corresponding absorption coefficient; (b) and (e) using the minimum refractive index and the corresponding absorption coefficient, (c) and (f) using the averaged values of the refractive index and the absorption coefficient as explanatory variables. R2, determination coefficient.

4.5 Conclusions

With the strong birefringence exhibited by wood, the complex refractive index and absorption coefficient can be used to predict many properties. Particularly, for the predictions of density and MOE, the predicted values and measured values showed a high correlation, however, the result of MFA and cellulose crystallinity showed a worse result than density and MOE. As argued in chapter 4.3.1, the birefringence of wood that shows in the THz region has two different resources, the form birefringence which is caused by the regular alignment of the cell wall, and the intrinsic birefringence which is caused by the crystalline region in the cell wall. Therefore, to improve the prediction of MFA and cellulose crystallinity that both directly correlated with the cellulose crystalline region in the cell wall, further research on this in the THz region is necessary. And it will be the main theme of the next three chapters of this thesis.

5. THz-TDS as a novel tool for crystallographic analysis in cellulose: characteristic absorption of I_α and I_β allomorph

5.1 Introduction

Cellulose is the most abundant and widely distributed polysaccharide in nature. It is an important component of the structural framework of plant cell walls and is synthesized by cellulose synthetases on cytoplasmic membranes. As described above, the prediction accuracy for MFA and crystallinity of wood by using THz-TDS is very limited, therefore, to improve the prediction accuracy, it is necessary to investigate the cellulose (basic unit of microfibril) in the THz region. In addition, THz radiation is responsive to optical phonons in crystal lattices; therefore, THz-TDS has potential utility in crystallographic analysis. There have been some reports that used THz-TDS to investigate the polysaccharide crystalline such as D-glucose, fructose, and sucrose (Ren et al. 2019), but the characteristic absorption of different cellulose allomorphs are still uncertain. Not only in the field of cellulose but the XRD is also considered to some extent a standard tool for other crystallographic studies. However, it is hard for XRD to distinguish the cellulose I_α and I_β allomorphs (Horikawa 2017). Where cellulose I is the crystalline cellulose that is naturally produced by plants, algae, and bacteria. Cellulose I has been found to consist of two polymorphs: I_α has a triclinic structure and contains one cellulose chain, and I_β has a monoclinic structure and contains two cellulose chains (Sugiyama et al. 1991), coexist according to the different proportions of cellulose sources (O'sullvian 1997; Nishiyama 2009)

In this study, THz-TDS was used to investigate cellulose I with different fractions of I_α and I_β extracted from various sources, and the characteristic absorption of cellulose I_α and I_β allomorphs was confirmed for the first time.

5.2 Method and materials

5.2.1 Preparation of the cellulose samples

The same cell wall samples with various I_α and I_β fractions were extracted from different sources, and the XRD profiles employed in previous research (Horikawa 2017) were still used in this study. Cell wall samples of *Glaucocystis nostochinearum*, *Valonia macrophysa*, *Cladophora* spp., *Boergesenia forbesii*, and *Oocystis apiculata* were boiled in 1% NaOH for 8–10 h and treated in 0.05 NHCl at room temperature. The samples were then bleached in 0.3% $NaClO_2$ with a pH of 4.9

in an acetate buffer at 70 °C for 3 h. The obtained sample was then thoroughly washed with distilled water. At room temperature, *Halocynthia roretzi* was treated with 5% KOH and washed with distilled water. The samples were incubated for 2 h in a 70 °C bleach solution, and the procedure was repeated several times. After the sample became colorless, it was washed with distilled water. All purified cellulose samples were freeze-dried and made into three tablets for each source with a diameter of 7.3 mm and a thickness of 0.5 to 1.0 mm by hand pressing with a special die set (KYOTO PASTEC Co., Ltd.).

5.2.2 THz-TDS measurement

THz transmission spectra were acquired using a Tera Prospector-Kit model (NIPPO PRECISION Co., Ltd.). The THz beam was horizontally polarized with a bandwidth that extended from approximately 0.01 to 4.00 THz. The diameter of the THz beam spot on the sample was around 3 mm. Each measurement was recorded by averaging 100 scans to improve the signal-to-noise ratio (one scan takes about 5 s). The reference THz signal without the samples was obtained prior to and after each transmission measurement. For reproducibility, all measurements were conducted three times. Because THz is sensitive to water vapor, the THz optical system was placed in an almost-closed acrylic box that was continuously filled with dry air to ensure the stability of the humidity; all samples were placed in the box for 24 h before measurements to balance the ambient humidity. The thickness of samples was measured using a caliper (±0.01 mm), and the mass was measured with an electronic balance (±0.0001 g) before and after all the THz measurements, which was designed to ensure that during the measurements, there was no water absorption.

5.3 Results and discussion

5.3.1 Calculation of crystalline structure values with XRD1

The XRD spectra were cut out with a scattering angle (2θ) from 10° to 30° as shown in Fig. 5-1(a), different colors to distinguish the source of the samples. All the spectra could be considered composites of a baseline, a background intensity curve which may attributed to the amorphous scattering, and the three main crystalline peaks for I_α had Miller indices of (100), (010), and (110), which corresponded to the $(1\bar{1}0)$, (110), and (200) peaks in the I_β pattern, which denoted as peak1, peak2, and peak3 in Fig.5-1 (a). The positions and FWHM varied for the samples from different sources, which was due to the varying I_α fractions of the cellulose sample, to conduct further analysis,

the baseline was first fitted as a first-order polynomial, and the separation of peaks was carried out with Gaussian peaks using the following equation with a curve-fitting process:

$$I_G(2\theta) = I_{max} \cdot exp\left\{-4\,ln(2) \cdot \left(\frac{2\theta - 2\theta_{max}}{\beta}\right)^2\right\} \quad (5-1)$$

where I_{max} is intensity of peaks, $2\theta_{max}$ is the peak position, and β is the FWHM.

The d-spacings, crystalline size (Cs), and crystallinity index (CrI) were calculated using the following equations:

$$d = \frac{\lambda}{2\,sin\,\theta} \quad (5-2)$$

$$C_s = \frac{K\lambda}{\beta \cdot cos\theta} \quad (5-3)$$

$$CI = \frac{S_{cr}}{S_{cr} + S_{Am}} \quad (5-4)$$

where λ is the wavelength of the X-ray radiation (0.1542 nm); θ is the Bragg angle (Bragg and Bragg 1913); K is the Scherrer constant, which was around 0.9 when assuming Gaussian peaks; and β is the FWHM in radians (Zsigmondy and Scherrer 1912).

In this study, the crystallinity index was calculated with the integral intensity, where S_{cr} is the sum of the integral intensity of each crystal peak and S_{Am} is the integral intensity of the amorphous peak. All the calculated results are summarized in Table 5-1, where the cellulose data are supported by other research (Wada and Okano 2001; Wada et al. 2001), and the fraction of I_α was referenced with data estimated from IR spectra (Belton et al. 1989; Imai and Sugiyama 1998; Imai et al. 1999).

For peak 1, the peak position (2θ) range was from 14.53° to 15.03° and the FWHM was from 0.74° to 1.20°. For peaks 2 and 3, the peak position range was 16.70° to 17.07° and 22.72° to 23.10°, respectively, and the FWHM range was 0.61° to 1.20° and 0.58° to 0.98°, respectively. As shown in Fig. 5-1(a), the *Halocynthia* (black line) which only had 0.05 fraction of cellulose I_α showed observable peak shift at peak1 and peak2, the directions of peak shifts were indicated by the black arrow in Fig. 5-1(a).

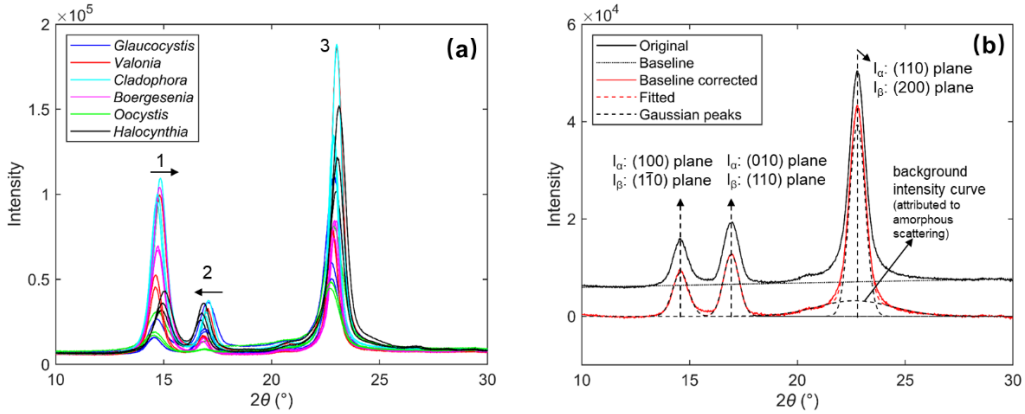

Fig. 5-1 The original and preprocessing of the XRD patterns of cellulose and classification of the type of cellulose. **(a)** Original XRD patterns of cellulose samples with different $I_\alpha$fractions, three peaks corresponding to the Miller indices of (100), (010), and (110) in the I_α type and (1$\bar{1}$0), (110), and (200) peaks in the I_β pattern, black arrows indicated the directions of peak shifts of *Halocynthia.* **(b)** Baseline correction and deconvolution process of XRD patterns. Black line: the original pattern of *Glaucocystis*. Black dot dotted line: baseline. Red line: baseline-corrected pattern of *Glaucocystis*. Black dashed line: separation of three peaks corresponding to the Miller indices of (100), (010), and (110) in the I_α type and (1$\bar{1}$0), (110), and (200) peaks in the I_β pattern, and the background intensity curve. Red dashed line: three separated peaks and amorphous scattering fitted pattern.

The range of d-spacing values for peaks 1, 2, and 3 were from 0.589 to 0.610 nm, 0.519 to 0.531 nm, and 0.385 to 0.391 nm, respectively. The d-spacing value for the I_β-dominant type (i.e., *Halocynthia*) and I_α-rich type (i.e., *Oocystis*) showed a clear difference, where the d1 was lower in value in the I_β-dominant type, and d2 was higher in value than the I_α-rich type. On the other hand, the d3 value showed no obvious changes related to the I_α fractions, and the crystalline size and crystallinity index did not show significant changes related to the I_α fractions. As shown in Fig. 5-2 (a), using the plot of d1 and d2, I_α fractions of the samples could be distinguished, and the results were consistent with previous research (Wada et al. 2001). In order to find the correlation between d1, d2 values and the fraction of I_α, principal component analysis (PCA) was carried out to compres the standardized d1 and d2, the results were shown in Fig. 5-2(b), where the fraction of I_α showed a high correlation with the PC1 score, the correlation coefficient of *r* was 0.96, which indicated the fraction of I_α can be estimated using XRD patterns.

Table 5-1. Calculated values of the crystalline structure from the XRD pattern and published I_α fractions

Position of the peaks (°)			FWHM (°)			d-spacing (nm)			Crystalline size (nm)				
2θ	2θ	2θ	β1	β2	β3	d1	d2	d3	cs1	cs2	cs3	CI	fa
Glaucocystis													
14.71	17.04	22.89	1.05	1.00	0.90	0.602	0.520	0.389	7.63	8.01	8.97	76.15	0.81[a]
14.58	16.93	22.78	0.90	0.85	0.86	0.607	0.524	0.390	8.95	9.41	9.45	79.90	
14.57	16.91	22.77	0.93	0.88	0.86	0.608	0.524	0.391	8.61	9.11	9.38	76.81	
Valonia													
14.83	17.02	23.02	0.78	0.79	0.64	0.598	0.521	0.386	10.22	10.15	12.58	84.99	0.64[a]
14.63	16.83	22.82	0.77	0.61	0.63	0.606	0.527	0.390	10.48	13.07	12.87	87.34	
14.62	16.82	22.82	0.74	0.56	0.61	0.606	0.527	0.390	10.86	14.39	13.24	87.27	
Cladophora													
14.85	17.07	23.01	0.77	0.66	0.60	0.597	0.519	0.386	10.42	12.11	13.55	85.59	0.76[b]
14.73	16.95	22.90	0.77	0.66	0.59	0.602	0.523	0.388	10.40	12.20	13.77	86.23	
14.68	16.90	22.85	0.77	0.65	0.58	0.604	0.525	0.389	10.48	12.30	14.06	86.25	
Boergesenia													
14.82	16.87	22.98	0.82	0.70	0.62	0.598	0.525	0.387	9.82	11.48	13.00	88.53	0.55[b]
14.72	16.80	22.90	0.81	0.62	0.63	0.602	0.528	0.388	9.90	12.94	12.94	89.15	
14.73	16.82	22.91	0.82	0.65	0.63	0.602	0.527	0.388	9.79	12.32	12.95	86.53	
Oocystis													
14.69	16.92	22.87	1.20	1.20	0.98	0.603	0.524	0.389	6.68	6.70	8.31	72.02	0.83[a]
14.53	16.81	22.72	1.09	0.93	0.98	0.610	0.527	0.391	7.33	8.60	8.25	75.26	
14.56	16.83	22.75	1.10	1.20	0.94	0.609	0.527	0.391	7.30	6.70	8.59	72.22	
Halocynthia													
14.89	16.70	22.98	0.99	0.79	0.79	0.595	0.531	0.387	8.06	10.23	10.25	79.85	0.05[c]
14.94	16.75	23.04	0.99	0.79	0.78	0.593	0.529	0.386	8.08	10.23	10.46	78.92	
15.03	16.84	23.10	1.02	0.83	0.79	0.589	0.527	0.385	7.88	9.72	10.22	78.36	

fa: Cellulose I_α fractions

[a]Imai et al. 1999

[b]Imai and Sugiyama 1998

[c]Mean value of fractions reported by Belton et al. 1989; Larsson et al. 1995

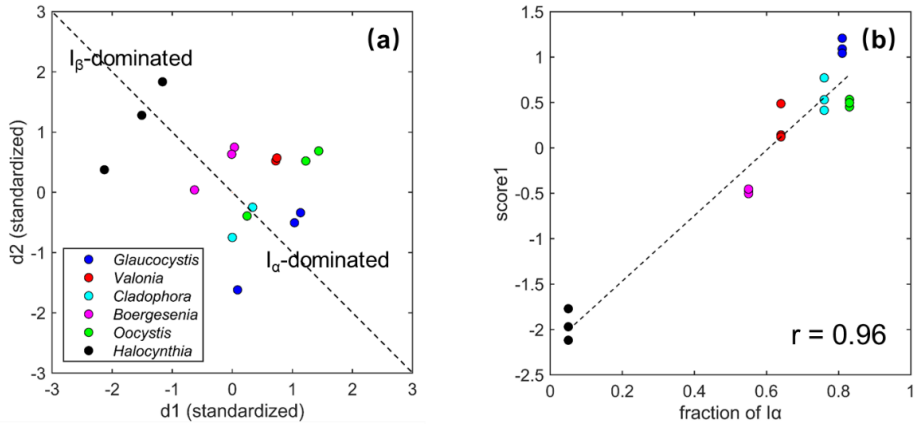

Fig. 5-2 (a) Scatter plot of standardized d1 and d2, where the top left corner corresponds to the I_β-dominant pattern and the bottom right corner corresponds to the I_α rich pattern, **(b)** relationship between fraction of I_α and the PC1 score, r, correlation coefficient composites

5.3.2 Absorption coefficient of cellulose with different crystalline structures at the THz region

All the measured THz time-domain signals were Fourier transformed then calculated the absorption coefficient using the Eq. (2-6) and Eq. (2-7). The original absorption coefficient spectra are shown in Fig. 5-3(a), all the samples showed absorption peaks around 3 THz, where the intensity of the absorption peaks has been reported to be correlated with the cellulose crystallinity index (Vieira and Pasquini 2014). On the contrary, the insanities of these peaks did not show a clear correlation with the I_α fractions, whereas the peak positions showed some fluctuations with the different I_α fractions of the samples. The peaks around 2.1 and 2.4 THz in the absorption spectra are noteworthy, which may correlate to the cellulose crystalline structure. One abnormal spectrum of *Cladophora* showed obviously different absorption characteristics with two other *Cladophora* samples (indicated by the arrow in Fig. 5-3), the abnormal spectrum was eliminated from the analysis here.

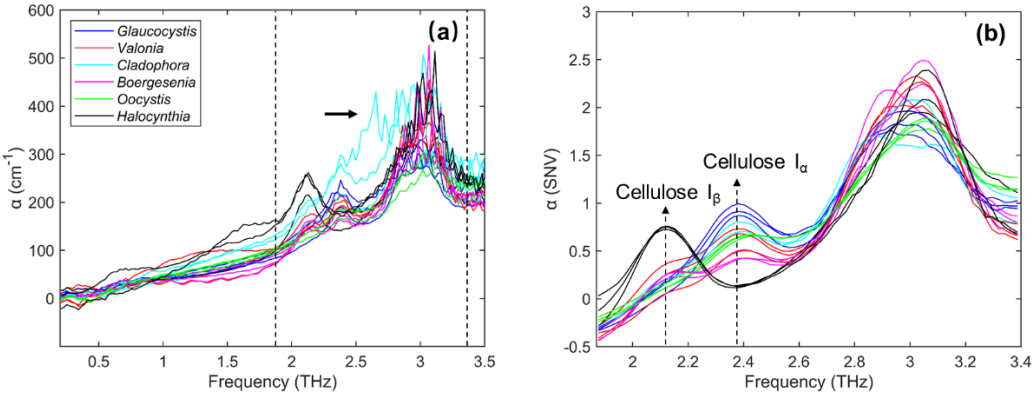

Fig. 5-3 (a) Original THz absorption coefficient spectra, the abnormal spectrum of *Cladophora* is indicated by the black arrows and was eliminated in subsequent processing, **(b)** SNV and smoothed absorption coefficient spectra of the frequency range of 1.88–3.40 THz.

All the spectra were treated with a standard normal variate algorithm (SNV) to correct baseline fluctuations of the absorption spectra, and the bands from 1.88 to 3.40 THz were cut out and smoothed by applying a Savitzky–Golay filter with a second-order polynomial and fifteen smoothing points. As shown in Fig. 5-3(b), the I_β-dominant type showed absorption peaks at 2.11 THz, whereas the I_α-rich type showed absorption peaks at 2.38 THz. It can be concluded that the absorptions at 2.11 and 2.38 THz were due to cellulose I_β and I_α types, respectively.

The absorption coefficients at 2.11 and 2.38 THz showed a considerable linear correlation with the I_α fractions as shown in Fig. 5-4 (a) and (b) with correlation coefficient values, r, of −0.94, 0.90,

respectively. This indicates that absorption coefficients at 2.11 and 2.38 THz reflected the difference in the proportion of components in cellulose I_α and suggested that the absorption spectra could correspond to the crystalline structure to some extent.

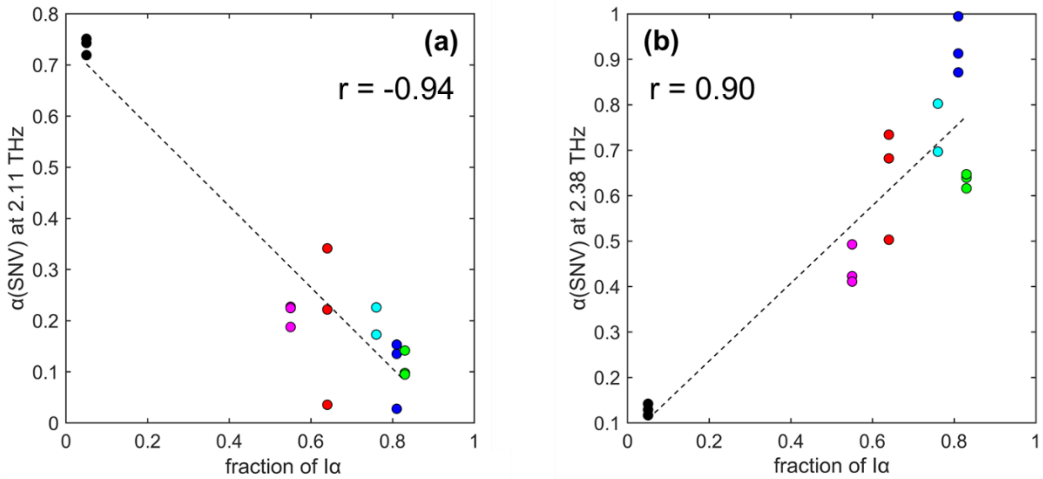

Fig. 5-4 Relationship between fraction of I_α and absorption coefficients. **(a)** 2.11 THz, **(b)** 2.38 THz, *r*, correlation coefficient.

5.3.3 Correlations between THz absorption spectra and the crystalline structural values of cellulose

To further investigate the correlations between the THz absorption spectra and the crystalline structure and crystallinity index of cellulose, the absorption coefficient spectra of the frequency range of 1.88–3.40 THz were first detrended by subtracting the baseline, and the separation of peaks as that in XRD patterns was carried out with three Gaussian peaks using the Eq. (5-1), where the peak position $2\theta_{max}$ were changed into f_{max} that indicate the peak positions in THz region.

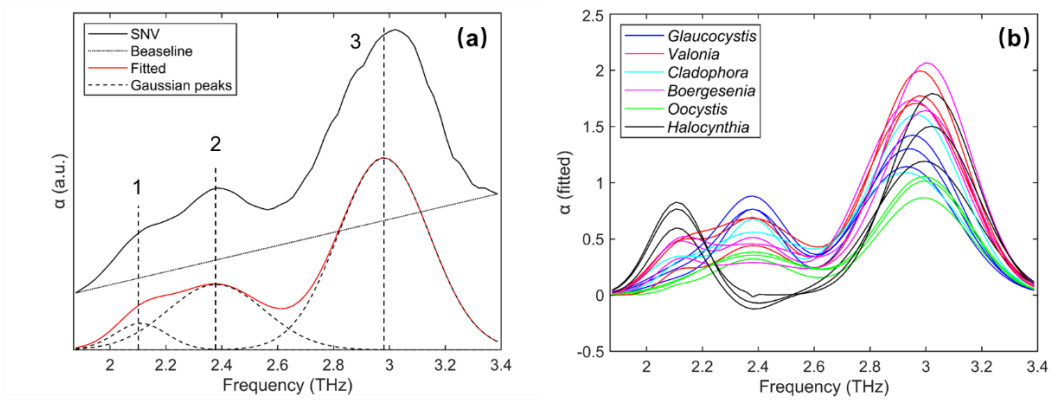

Fig. 5-5 The processing of THz absorption coefficients spectra, (a) the curve-fitting process, black line: the SNV spectra of *Valonia*. Black dot line: baseline. Black dashed line: separated three Gaussian peaks. Redline: three separated peaks fitted absorption coefficients spectra. (b) the fitted absorption coefficient spectra of cellulose samples with different I_α fractions.

Taking *Valonia* with I_α fraction of 0.64, which showed absorption peaks at both 2.11 and 2.38

THz as an example, the curve-fitting process was shown in Fig .5-5(a). Peak1 and peak2 were fitted by only using I_{max} and FWHM β, where peak positions were fixed at 2.11 and 2.38 THz that corresponded to the absorption of cellulose I_β and I_α patterns as described above, peak3 were fitted with all three parameters, and the fitted spectra were shown in Fig. 5-5(b), where the peak shift of peak3, which seemed to be correlated with the fraction of I_α can be observed.

The correlations between the parameters calculated from detrended THz absorption coefficient spectra (integrated intensity, peak positions of peaks 3) and the parameters calculated from XRD patterns (the d-spacing, crystalline size, and crystallinity index) were investigated, and the correlation coefficient value was more significant than 0.6 were summarized in Table 5-2.

As shown in Table 5-2, the integrated intensities of peak1 and peak2 in THz region were correlated with the fraction of I_α, and the d-spacing values (d1 and d3), that corresponds to the lattice planes of (100), (110) in the cellulose I_α type and $(1\bar{1}0)$, (200) in the cellulose I_β type. Different from the peak1 and peak2, the integrated intensity of peak3 correlated with all three crystalline sizes (Cs1, Cs2 and Cs3) including Cs2 correspond to lattice planes of (010) in the cellulose I_α type and (110) in the cellulose I_β type, and the integrated intensity also showed a good correlation with the crystallinity index of cellulose. It is worth noting that the peak positions of peak3 were correlated with the fraction of I_α and the d2 values which correspond to d-spacing of lattice planes of (010) in the cellulose I_α type and (110) in the cellulose I_β type.

The correlations may indicate that the absorption peaks of 2.11 and 2.38 THz reflected the intermolecular vibrations along the direction between the lattice planes of (100), (110) in the cellulose I_α type and $(1\bar{1}0)$, (200) in the cellulose I_β type. On the other hand, the absorption peaks around 3 THz might reflect the size of the crystalline region, since the integrated intensity of the peak3 was not only correlated with the crystalline size but also correlated with the crystallinity index. It is worth mentioning that the absorption peak around 3 THz showed peak shift, probably due to the absorption peaks here are composed of two peaks that are very close to each other, and the intensity change of the two peaks caused such a slight peak shift. These two peaks may correspond to the cellulose I_α and I_β types, just like the two absorption peaks shown in the 2.11 and 2.38 THz. Under such assumption, the peak position around 3 THz showed a correlation with the fraction of I_α might be reasonable. However, we were unable to effectively separate the peaks at 3 THz which

may be due to the relative low S/N ratio.

Table 5-2. Correlations between the THz parameters and the crystallite values calculated from XRD patterns and published I_α fractions

THz parameters	crystallite values	correlation coefficient
integrated intensities of peak1	d1	-0.86
	d3	-0.81
	fa	-0.88
integrated intensities of peak2	d1	0.71
	d3	0.62
	fa	0.81
integrated intensities of peak3	Cs1	0.78
	Cs2	0.79
	Cs3	0.76
	CrI	0.86
peak positions of peak3	d2	0.79
	fa	-0.61

Cs: Crystallite size
fa: Cellulose I_α fractions
CrI: Crystallinity index

5.4 Conclusions

In this study, various sources of cellulose samples were investigated by THz-TDS and XRD. After PCA processing the XRD pattern can determine the fractions of I_α directly, on the other hand, the THz absorption coefficient spectra showed two peaks at 2.11 and 2.38 THz, which corresponded to the absorptions due to cellulose I_β and I_α types, respectively. Furthermore, the intensity at 2.11 and 2.38 THz showed a considerable linear correlation with fractions of I_α. The parameters obtained from the detrended THz spectra (integrated intensity and peak position) were shown correlations with that calculated from XRD patterns (the d-spacing, crystalline size, and the crystallinity). However, the assignments of the absorption peaks in the THz region of cellulose are still not certain. This study demonstrated the potential of THz-TDS as a powerful tool for crystallographic analysis with rapid measurement, especially the cellulose I_α and I_β allomorphs, and the fractions of the component can be investigated easier than XRD.

6. THz-TDS as a novel tool for crystallographic analysis in cellulose: the potentiality of being a new standard for evaluating crystallinity

6.1 Introduction

The crystallinity of cellulose reflects the physical and chemical properties of cellulose-based materials. Therefore, the crystallinity of cellulose must be investigated. Generally using the CrI which XRD measures. The most used was the Segal peak height method before, and the peak deconvolution method was most discussed in recent studies of CrI. However, when the cellulose crystal is relatively small, the “amorphous peak” probably overlaps the adjacent crystalline peaks, which means the accurate expression of amorphous curves using only profiles such as Gaussian function, Lorentzian function and Voigt function is difficult (del Cerro et al. 2020). In the previous study, the CrI of microcrystalline cellulose (MCC) was determined by THz-TDS combined with a partial least squares (PLS) model (Vieira and Pasquini 2014). In this study, we follow up on the previous research described in Chapter 5, THz-TDS was used to investigate the crystallinity of cellulose-based materials, including MCC and wood, both ball-milled at different times, and pseudo-wood which was a mixture of MCC and lignin.

6.2 Method and materials

6.2.1 Samples preparation

To obtain cellulose and wood powder with different CrI, the MCC powder (EMD Millipore 1.02331.0500) were ball-milled with a benchtop ball-mill (AV-2, Asahi Rika Factory. Ltd) at 200 rpm using ceramic spheres and jars for 0, 12, 24, 48, 72, and 144 hours. An air-dried softwood, Hinoki cypress (*Chamaecyparis obtuse*), was first crushed by a rotary crusher and sieved with 10 US standard mesh to collect wood powder with a particle size of 200 μm. The wood powder was then ball-milled for 0, 6, 12, 24, 32, 48, 72, and 144 hours using the same procedure as the MCC powder.

The pseudo-wood powder was a mixture of MCC and organic-solvent lignin powder (Guangzhou Yinnovator Bio-tech Co., Ltd), where the lignin powder was extracted from Eucalyptus (*Eucalyptus*) by a high-concentration glacial acetic acid (>70%). To obtain different CrIs, we used the following mass fractions of lignin in the powder: 0%, 25%, 50%, 75%, and 100%. The 100% organic-solvent lignin powder can be considered as a totally amorphous powder.

The prepared powders with different CrIs were collected at a mass of 0.075 g by an electronic balance (±0.0001 g). All the powders were compressed into tablets with a diameter of 14 mm and a thickness of approximately 0.35 mm using a hand-pressing tableting kit (IMC-180C, Imoto Machinery Co., LTD). Three tablets were prepared for each powder. The thickness of these tablet samples was measured using a micrometer (± 0.001 mm).

6.2.2 XRD and THz-TDS measurement

The XRD measurements of all the tablet samples were performed with Cu-Kα radiation (λ = 0.1542 nm) using an X-ray diffractometer (Ultima IV, Rigaku) at a voltage of 40 kV and a current of 40 mA. Diffractograms were recorded from 5° to 40°, where the scan range included the peak of the main crystal lattice of cellulose and wood. The scan speed was set to 5° min^{-1}, and the sampling step was 0.05°. The background diffractogram was obtained from an empty sample holder.

The THz transmission spectra of all the tablet samples were measured by using a Tera Prospector-Kit model (NIPPO PRECISION Co., Ltd.), and the reference signals were obtained by measurement of air before and after the sample measurement. The THz beam was horizontally polarized with the bandwidth from about 0.1 THz to 4.00 THz, and the spectral resolution was 0.02 THz, corresponding to the inverse of the temporal scan range (50 ps). The diameter of the THz beam spot on the sample was around 3 mm. Each measurement was recorded by averaging 100 scans to improve the signal-to-noise ratio. For reproducibility, all measurements were conducted thrice. To avoid the influence of water vapor on the measurement caused by THz absorption, we placed the whole THz optical system in an almost-closed acrylic box that was continuously filled with dry air to ensure the stability of humidity. All samples were placed in the box for 24 hours before measurement to balance the ambient humidity. When the samples were placed in the box, dry airflow was filled into the box until all the THz measurements were completed.

6.3 Results and Discussion

6.3.1 Calculation of CrI with XRD pattern

The original XRD pattern was cut out with a scattering range from 10° to 40° of pseudo-wood, MCC, and wood samples, as shown in Fig. 6-1(a), (b), and (c), respectively, with gradient colors expressed changes in the ball-milling time of cellulose and wood samples or mass fractions of lignin of pseudo-wood sample. Background subtraction and baseline correction were performed on all

original XRD patterns before further calculation. The background pattern was obtained as mentioned in the experimental section. The baseline was fitted as a first-order polynomial after background subtraction; the process is shown in Fig. 6-2(a).

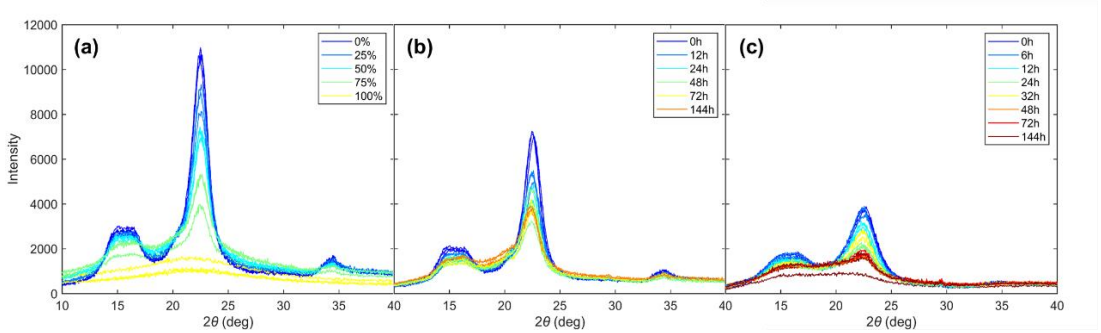

Fig.6-1 Original XRD patterns of **(a)** pseudo-wood with a decreased mass fraction of organic-solvent lignin, **(b)** MCC, and **(c)** wood samples.

After background subtraction and baseline correction, the corrected XRD patterns can be considered as composites of an amorphous intensity curve and five main crystalline peaks for the I_β type cellulose, a dominant type in vascular plants such as wood. The five main crystalline peaks had Miller indices of $(1\bar{1}0)$, (110), (102), (200), and (004). The amorphous intensity curve was determined by two different methods (Park et al. 2010), which will be discussed in detail later.

The deconvolution of all the crystalline peaks was carried out with a curve-fitting process using a pseudo-Voigt profile, which is a linear combination of a Gaussian (Eq. (5-1)) curve and a Lorentzian curve as shown respectively in the following equations:

$$I_L(2\theta) = I_{max}\left\{1 + 4\left(\frac{2\theta - 2\theta_{max}}{\beta}\right)^2\right\}^{-1} \qquad (6-1)$$

$$I_{PV}(2\theta) = \mu I_L + (\mu - 1)I_G \qquad (6-2)$$

where I_{max} is the intensity of peaks, $2\theta_{max}$ is the peak position, and β is the full width at half maximum (FWHM); $I_G(2\theta)$ and $I_L(2\theta)$ are the Gaussian and Lorentzian curves, respectively (de Keijser et al. 1983; Wada et al. 1997)

The positions (2θ) of the crystalline peaks were fixed at 14.9°, 16.7°, 20.5°, 22.6°, and 34.5°, corresponding to the Miller indices of $(1\bar{1}0)$, (110), (102), (200), and (004), respectively. The other parameters of FWHM, the peak intensity, and the coefficient μ were all determined by the curve-fitting process.

In this study, the amorphous intensity curve was first fitted as a pseudo-Voigt profile, as shown in Fig. 6-2(b), the same fitting process as that of other crystalline peaks. Given that recent research

showed that the maximum of the amorphous intensity curve is slightly more than 20°, therefor, the peak position (2θ) of the amorphous intensity curve was fixed at 20.6° (Yao et al. 2020), where the position was very close to the (102) crystalline peak, and the influence caused by this will be discussed in detail later.

Yao et al. reported the Fourier series modeling of the amorphous intensity curve, and this method was used in this study (Yao et al. 2020). The amorphous intensity curve was determined by fitting the averaged XRD pattern of three 100% organic-solvent lignin samples using a 7th-order Fourier series as shown in Fig. 6-2(c). The obtained Fourier series model can be considered as the basic function of the amorphous intensity curve, where different samples only showed different curve intensity that were adjusted by multiplying this Fourier series by a coefficient *k* in the deconvolution of XRD patterns:

$$f(2\theta) = k \cdot \{a_0 + \sum_{i=1}^{7} a_i \cos(2\theta \cdot w) + b_i \cos(2\theta \cdot w)\} \qquad (6-3)$$

where a_0, a_i, b_i and w are all constants determined automatically by the fitting process, and the details of the fitting of the Fourier series of 100% organic-solvent lignin samples is shown in Fig 6-3, the fitted formular of the amorphous intensity curve is also list below:

f(x) = a0 + a1*cos(x*w) + b1*sin(x*w) +
a2*cos(2*x*w) + b2*sin(2*x*w) + a3*cos(3*x*w) + b3*sin(3*x*w) +
a4*cos(4*x*w) + b4*sin(4*x*w) + a5*cos(5*x*w) + b5*sin(5*x*w) +
a6*cos(6*x*w) + b6*sin(6*x*w) + a7*cos(7*x*w) + b7*sin(7*x*w)

Coefficients (with 95% confidence bounds):

a0 = -2.544e+04 (-2.927e+05, 2.418e+05)
a1 = -4.673e+04 (-5.062e+05, 4.127e+05)
b1 = 8965 (-1.768e+05, 1.947e+05)
a2 = -3.391e+04 (-3.235e+05, 2.557e+05)
b2 = 1.262e+04 (-2.577e+05, 2.83e+05)
a3 = -1.978e+04 (-1.395e+05, 9.99e+04)
b3 = 1.192e+04 (-2.24e+05, 2.478e+05)
a4 = -9211 (-3.383e+04, 1.541e+04)
b4 = 7510 (-1.359e+05, 1.51e+05)
a5 = -3234 (-1.095e+04, 4486)
b5 = 3403 (-5.751e+04, 6.431e+04)
a6 = -880.7 (-6200, 4438)
b6 = 1006 (-1.694e+04, 1.895e+04)

a7 = -163.4 (-926.6, 599.8)

b7 = 132.3 (-3087, 3351)

w = 0.1157 (0.02327, 0.2082)

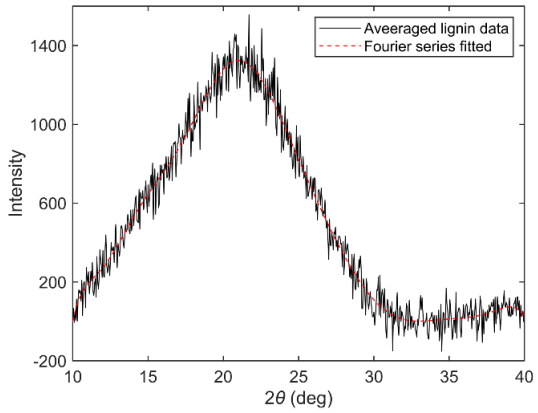

Fig. 6-3 Fitting of the amorphous intensity curve by a 7th order Fourier series modeling, black line: the averaged pattern of three100 % organic solvent lignin samples, red dushed line: fitted 7th order Fourier series modeling.

The deconvolution results were evaluated with a coefficient of determination R^2:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(X_i - Y_i')^2}{\sum_{i=1}^{n}(X_i - \bar{X})^2} \qquad (6-4)$$

where X_i is the intensity after background subtraction and baseline correction, Y_i is the fitted intensity, and $\bar{X}$ is the average of intensities. Table 6-1 summarizes the calculated R^2 of the pseudo-wood fitted by two different amorphous intensity curves. As shown in Table 6-1, the curves fitted by both methods showed good R^2, which indicates that the curve-fitting of the two methods was credible.

Once the peaks were deconvoluted by the curve-fitting process, the CrI was determined by Eq. (5-4). As shown in Fig. 6-2(b) and (c), the amorphous intensity curve fitted with a pseudo-Voigt profile and a Fourier series gave different deconvolution results. The most evident intensity of the (102) peak in Fig. 6-2(b) was almost invisible, since the peak position of (102) peak was fixed at 20.5° and the amorphous peak was fixed at 20.6°, the contribution of the (102) peak was almost systematically ignored in the fitting process. On the other hand, the (102) peak can be clearly observed in Fig. 6-2(c). Due to this, the CrI calculated with a pseudo-Voigt profile amorphous intensity curve was smaller than that calculated with a Fourier series amorphous intensity curve on average.

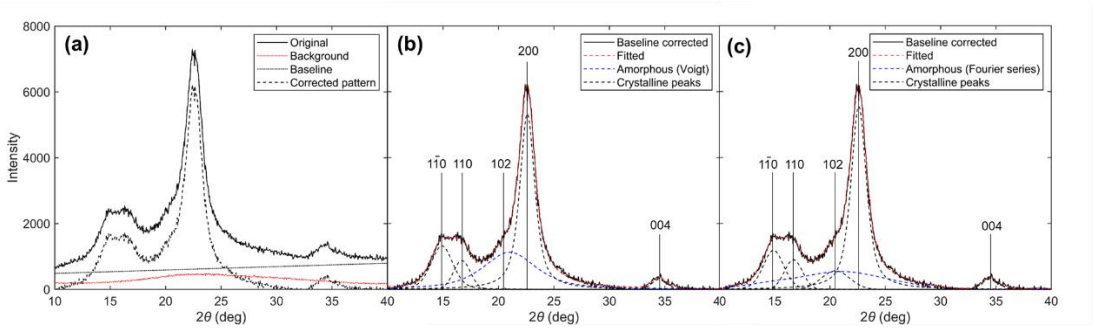

Fig. 6-2 Curve-fitting process of a pseudo-wood with 50% lignin. **(a)** Background subtraction and baseline correction, **(b)** deconvolution of peaks by using a pseudo-Voigt profile only, and **(c)** the deconvolution of peaks where the crystalline peaks fitted by using a pseudo-Voigt profile; the amorphous curve was fitted by a 7th-order Fourier series model

Table 6-1. Calculated R^2 values of the pseudo-wood fitted by two different amorphous intensity curves

Amorphous curve	**Volume fraction of lignin**				
pseudo-Voigt	0%	25%	50%	75%	100%
	0.992	0.993	0.996	0.993	0.909
R^2	0.990	0.987	0.987	0.989	0.960
	0.982	0.996	0.996	0.994	0.898
Fourier series					
	0.991	0.993	0.997	0.994	0.910
R^2	0.988	0.986	0.988	0.991	0.962
	0.979	0.996	0.996	0.995	0.902

Fig. 6-4 shows the CrI calculated from the different methods. In the figure, the amorphous curve for (a) pseudo-wood, (b) MCC, and (c) wood was fitted by the pseudo-Voigt profile, whereas that for (d) pseudo-wood, (e) MCC, and (f) wood was fitted by Fourier series. As shown in Fig. 6-4, two different fittings gave a similar crystallinity, whereas for the pseudo-wood, the crystallinity decreased from about 85% to about 5% with the increase in mass fraction of lignin. For MCC and wood, the crystallinity decreased from about 90% to 60% and 80% to 40% with the increase in ball-milling time, respectively.

As shown in Fig. 6-4 (a) and (d), for the pseudo-wood, the samples with a 25% mass fraction of organic-solvent lignin showed a similar crystallinity to the samples without organic-solvent lignin (100% MCC samples). A non-linear correlation was observed between the CrI and mass fraction of lignin, where the CrI decreased sharply when the mass fraction of lignin was over 50%. Regardless of the fitting method used to obtain the curve, the sample that had 100% organic-solvent lignin,

which was supposed to have no crystalline region, still showed a crystallinity of about 5%. The crystallinity for MMC and wood showed a similar downtrend. In general, the concentration of cellulose in different species of wood is about 30% to 40% (PETTERSEN 1984). However, in this study, for most of the ball-milling times, the CrI of wood was about 10% lower than that of MCC. For the same type of samples, the changes in the CrI calculated from the XRD patterns were easy to observe and evaluate. However, when the cellulose concentration of wood was about 30% to 40%, the original wood samples still showed a relatively similar CrI compared with that of the original MCC samples. Thus, accurate comparison of the differences in CrIs between various types of samples using the same standard can be a challenge. Table 6-2 summarizes the detailed crystallinity of the pseudo-wood sample calculated from the two curve-fittings and the crystallinity calculated from the THz mass absorption coefficient spectrum, which will be discussed in detail in the next section.

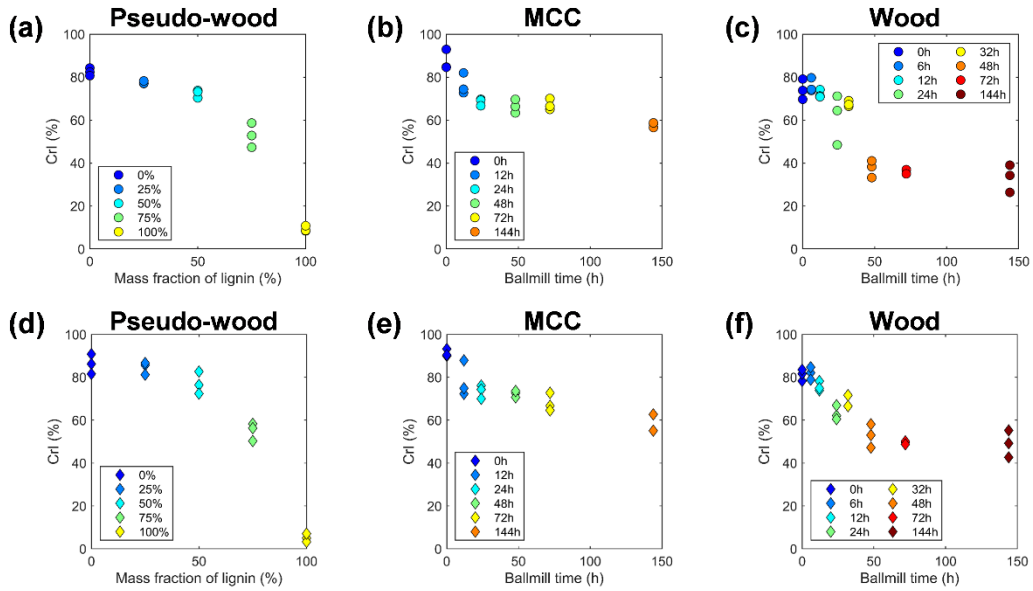

Fig.6-4 CrI calculated from two different curve-fitting processes of the amorphous curve, where **(a) – (c)** with circle markers were fitted by a pseudo-Voigt profile, and **(d) – (e)** with diamond markers were fitted by a 7th-order Fourier series.

6.3.2 Evaluation of CrI by THz mass absorption coefficient spectra

The measured THz time-domain signals were Fourier transformed into the frequency domain, different from the calculation in the chapter. 5, here, to correct the possible inhomogeneity of hand-made samples, the mass absorption coefficient α that also considers the influence of density was calculated using the following equations:

$$\alpha = \frac{1}{\rho L}\ln(\frac{R_r}{R_s}) \qquad (6-5)$$

where ρ is the sample density, L is the sample thickness, R_r, R_s are the phase and amplitude in the frequency domain of the reference and measured samples, respectively (Kore and Pawar 2014).

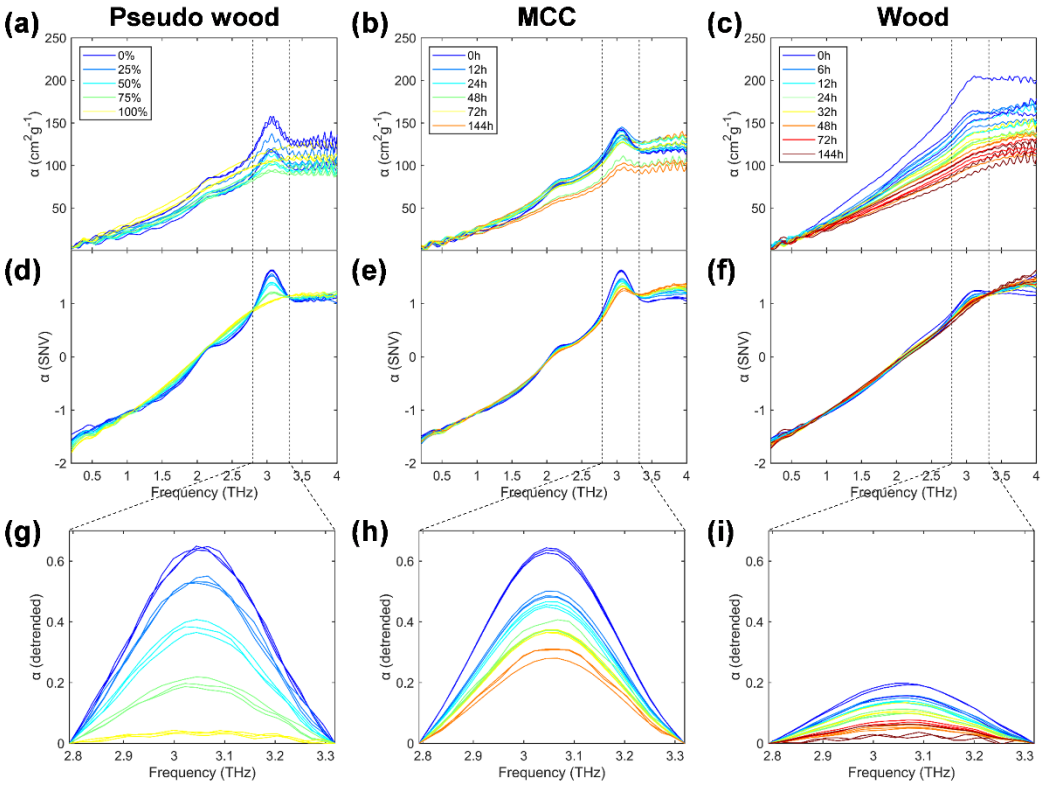

Fig. 6-5 Original THz mass absorption coefficient spectra of **(a)** pseudo-wood, **(b)** MCC, and **(c)** wood samples. **(d)**, **(e)**, and **(f)** SNV and smoothed mass absorption coefficient spectra. **(g)**, **(h)**, and **(i)** Detrended mass absorption coefficient spectra obtained by subtraction of the baseline at the frequency range of 2.79 THz to 3.32 THz

Fig. 6-5 (a), (b), and (c) showed the original THz mass absorption coefficient spectra from 0.2 THz to 4 THz of the pseudo-wood, MCC, and wood, respectively. The same gradient colors as that in Fig. 6-1 were used. The calculated mass absorption coefficient spectra were all first corrected by baseline fluctuations with SNV and then smoothed by the application of a Savitzky–Golay filter with a second-order polynomial and fifteen smoothing points, as shown in Fig. 6-5 (d)-(f). The pseudo-wood and MCC samples showed absorption peaks at around 2.1 THz, and these peaks gradually decreased with the increase in mass fraction of lignin or ball-milling time, whereas at this frequency position, no evident absorption peaks of wood samples were observed. The absorption peaks at 2.1 THz correlated with I_β crystalline structures as described in the chapter. 5. Similar to the case of wood samples, the absorption peaks of pseudo-wood samples with the mass fraction

over 50% of organic-solvent lignin were difficult to observe at around 2.1 THz, whereas the cellulose concentration of wood samples was about 30%. Therefore, no evident absorption peak, such as that of MCC samples at around 2.1 THz can be observed.

As shown in Fig. 6-5 (a)-(f), all types of samples expressed a relatively strong absorption at 3.04 THz without a peak shift. The mass absorption coefficient spectra of wood showed a relatively small intensity compared with other samples. To further study this absorption peak, we cut out the mass absorption coefficient spectra from the range of 2.79 THz to 3.32 THz and then subtracted the baseline that was fitted as a first-order polynomial. Fig. 6-5 (g), (h), and (i) showed the baseline corrected mass absorption coefficient spectra of the pseudo-wood, MCC, and wood, respectively. The integrated intensities of all the detrended mass absorption coefficient spectra were calculated, and their correlations with the mass fraction of lignin and ball-milling time are plotted in Fig. 6-6(a), (b), and (c).

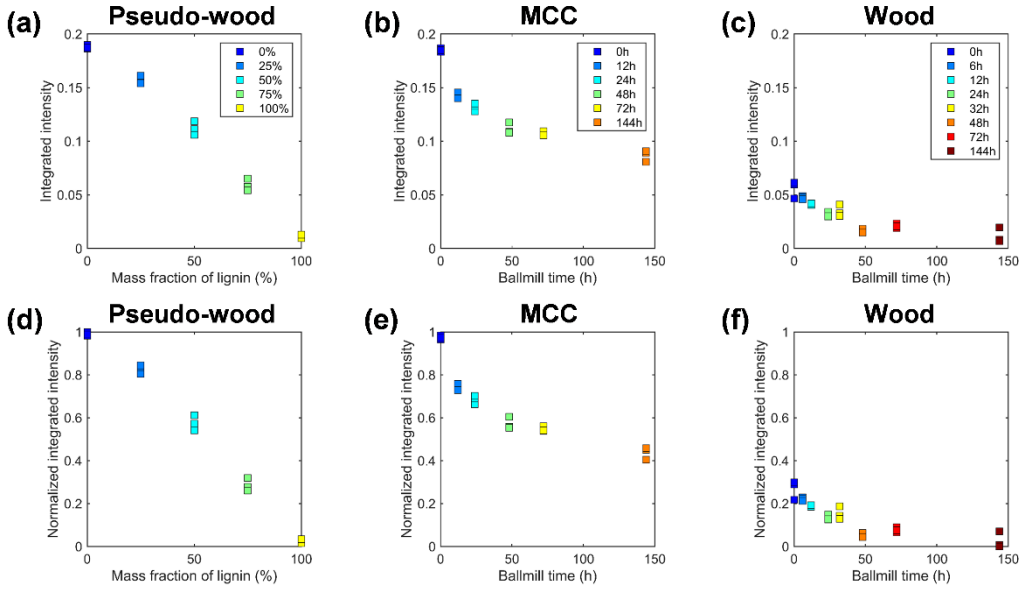

Fig. 6-6 Integrated intensity calculated from the detrended mass absorption coefficient spectra from 2.79 THz to 3.32 THz of **(a)** pseudo-wood, **(b)** MCC, and **(c)** wood samples. **(d)**, **(e)**, and **(f)** Normalized integrated intensity by min-max algorithm

For the pseudo-wood samples, as calculated from the XRD patterns, the integrated intensity of the mass absorption coefficient spectra from 2.79 THz to 3.32 THz showed a linear correlation with the mass fractions of lignin. Given that the mass fraction of lignin was adjusted to increase linearly (+25% each time) during the sample preparation process, the correlation obtained from THz mass absorption coefficient spectra is more reasonable than that obtained from XRD patterns. The

correlations between the ball-milling time with the integrated intensity of the MCC and wood samples showed a similar trend to that of the crystallinity obtained from the XRD patterns. However, the values of 144 h ball-milled MCC samples decreased by about 50% compared with that of the original MCC samples, whereas the CrI of the 144 hours of ball-milled MCC samples was calculated from XRD patterns decreased by about 25%. Furthermore, the original wood samples without ball-milling showed relatively small values compared with the original MCC samples. By contrast, the CrI of the original wood samples calculated from the XRD pattern showed no evident difference from that of the original MCC samples. The integrated intensities of the original wood samples were about 30% of those of the original MCC samples, thus matching with the approximate concentration of cellulose in wood (about 30%). Table 6-2 summarizes the detailed integrated intensities of the pseudo-wood samples and the CrI calculated by two curve-fittings from the XRD patterns.

Table 6-2. CrI of the pseudo-wood calculated from XRD patterns and the integrated intensity obtained from 2.79 THz to 3.32 THz mass absorption coefficient spectra

Calculation method	**Volume fraction of lignin**				
	0%	25%	50%	75%	100%
XRD with pseudo-Voigt Am	84.270	77.024	73.808	47.351	8.393
CrI (%)	82.361	77.163	70.362	58.641	8.586
	80.704	78.268	73.242	52.810	10.794
XRD with Fourier series Am					
CrI (%)	90.870	85.692	76.463	50.256	4.966
	86.272	81.236	72.402	58.293	3.296
	81.576	86.667	82.702	56.254	7.075
Detrended THz mass					
integrated intensity	0.171	0.132	0.095	0.051	0.011
	0.142	0.133	0.093	0.048	0.013
	0.178	0.139	0.083	0.041	0.013
Min-maxed integrated intensity					
	0.957	0.734	0.518	0.261	0.027
	0.789	0.740	0.506	0.246	0.039
	1.000	0.774	0.448	0.206	0.044

Am: Amorphous intensity curve

The integrated intensities of all samples were normalized to the range of 0 to 1 by the min-max algorithm, and the results are presented in Fig. 6-6 (d), (e), and (f). After this process, the normalized

value can be considered as an “index” that can be used to evaluate crystallinity. Different from the CrI, which is a relative value calculated from XRD patterns, an assumption was easily reached from the above results, that is, the THz mass absorption coefficient spectra may reveal the absolute mass fractions of crystalline cellulose in all the samples used in this study. Given that the peaks of the mass absorption coefficient in the THz region did not shift with the changes the type of samples or ball-milling time and mass fraction of lignin, only the intensity changed. Moreover, the THz spectra contain rich physical and chemical information, such as the phonon frequency of crystal lattices, which provides a possibility for the comparison of the mass fraction of crystalline cellulose between different cellulose-based materials.

6.4 Conclusions

All the cellulosic materials were observed absorption at 3.04 THz without a peak shift. For the pseudo-wood, the integrated intensity of detrended THz mass absorption coefficient spectra decreased linearly with the increase in the mass fraction of lignin. And the original wood sample without ball-milling had an integrated intensity value that was about 30% of that in the original MCC sample, matching the cellulose concentration of wood. By contrast, the CrI calculated from XRD patterns is suspicious, due to there are numerous optional functions are available for the deconvolution of XRD patterns that rely on the experience of analysts, and the evaluation of different cellulosic materials in the same standard is almost impossible. Thus, based on the above results, THz mass absorption coefficient spectra may be used to evaluate the mass fraction of crystalline cellulose. The measurement of the THz signal is quick, simple, and safe and requires no sample pretreatment. Thus, THz-TDS has the potential of becoming a new standard for the evaluation of cellulose crystallinity.

7. THz-TDS as a novel tool for crystallographic analysis in cellulose: tracing the structural changes under chemical treatment

7.1 Introduction

Cellulose II can be obtained by mercerization (alkali treatment) or regeneration (recrystallization after solubilization) of native cellulose (Kolpak and Blackwell 1976). After the alkali treatment such as NaOH of cellulose, both the crystalline lattice structure and the crystallinity change, therefore, tracing the transformation of the crystalline structure is essential for a better understanding of the processes of industry and biosynthesis where cellulose is involved, such as viscose rayon manufacturing (O'sullvian 1997). The same as the other studies of cellulose crystalline, XRD is the most used method (Oudiani et al. 2011; Kafle et al. 2014). As a follow-up to the previous study, the transformation of the crystalline lattice of cellulose I to that of cellulose II, and the changing in the crystallinity of cellulose II by ball-milling were traced by using THz-TDS in this study.

7.2 Method and materials

7.2.1 Samples preparation

To observe the transformation from cellulose I to cellulose II, the MCC powder (cellulose I, EMD Millipore 1.02331.0500) were treated with different concentration of NaOH (0, 2, 4, 6, 8, 10, 12, 14, 16, 18, 20 wt%) for 30 min at room temperature. After the reaction, the sample was washed with acetic acid and distilled water up to PH 7 and filtered, then the washed cellulose powders were dried in a desiccator containing phosphorus pentoxide. The NaOH, acetic acid, and phosphorus pentoxide used here were all purchased from KISHIDA CHEMICAL Co.,Ltd.

The cellulose II powder was obtained after the 20 wt% NaOH treatment as described above, to further obtain cellulose II with different CrI, the cellulose II powder were ball-milled with a benchtop ball-mill (AV-2, Asahi Rika Factory. Ltd) at 200 rpm using ceramic spheres and jars for 0, 6, 12, 24, and 48 hours.

The above powders that treated with different concentration of NaOH and ball-milled cellulose II with different time were collected at a mass of 0.075 g by an electronic balance (±0.0001 g). All the powders were compressed into tablets with a diameter of 14 mm and a thickness of approximately 0.35 mm using a hand-pressing tableting kit (IMC-180C, Imoto Machinery Co., LTD). For reproducibility of the experiment, three tablets were prepared for each powder. The

thickness of these tablet samples was measured using a micrometer (± 0.001 mm).

7.2.2 XRD and THz-TDS measurement

XRD measurement of all the tablet samples was performed with Cu-Kα radiation (λ = 0.1542 nm) using an X-ray diffractometer (Ultima IV, Rigaku) at a voltage of 40 kV and a current of 40 mA. Diffractograms were recorded from 5° to 40°. The scan speed was set to 5° min^{-1}, and the sampling step was 0.05°. The background diffractogram was obtained from an empty sample holder.

The THz transmission spectra of all the tablet samples were measured by using a Tera Prospector (Nippo Precision Co., Ltd.), and the reference signals were obtained by measurement of air before and after sample measurement. The THz beam was horizontally polarized with a bandwidth of about 0.1 THz to 4.00 THz, and the spectral resolution was 0.02 THz, which corresponded to the inverse of the temporal scan range (50 ps). The diameter of the THz beam spot on the sample was around 3 mm. Each measurement was recorded by averaging 100 scans to improve the signal-to-noise ratio. For reproducibility, all measurements were conducted thrice. To avoid the influence of the THz absorption of water vapor on the measurement, we placed the whole THz optical system in an almost-closed acrylic box, which was filled with dry air until all the THz measurements were completed to ensure stability of humidity. All samples were placed in the box for 24 h before measurement to balance the ambient humidity.

7.3 Results and Discussion

7.3.1 Tracing the crystalline change after NaOH treatment with XRD pattern and THz absorption spectrum

The original XRD pattern was cut out with a scattering range from 10° to 30°, the range including the main crystalline peaks of cellulose I and cellulose II. Before the further peak deconvolution, the background subtraction and baseline correction were performed on all original XRD patterns, where the background pattern was obtained as mentioned in the experimental section, and the baseline was fitted as a first-order polynomial after background subtraction, the processing was shown in Fig. 7-1 (a). In order to obtain the crystalline index of cellulose I and II, the three main crystalline peaks of cellulose I had Miller indices of $(1\bar{1}0)$, (110), and (200); for cellulose II, the Miller indices were $(1\bar{1}0)$, (110), and (020), total six crystalline peaks as described above, and two amorphous intensity curves for cellulose I and II, respectively, were deconvoluted via a curve-fitting process using a

pseudo-Voigt profile as shown in Fig. 7-1 (b). The positions (2θ) of the crystalline peaks of cellulose I profile were fixed at a range, which was 15° ± 0.5°, 16.5°± 0.5°, and 22.75° ± 0.25°, corresponding to the Miller indices of $(1\bar{1}0)$, (110), and (200), respectively. For cellulose II profile, the crystalline peaks were fixed at a range of 12° ± 0.5°, 20° ± 0.5°, and 21.5° ± 0.5°, corresponding to the Miller indices of $(1\bar{1}0)$, (110), (020). And the amorphous intensity curve of cellulose I and cellulose II were fixed at 20.6° and 16°, respectively. Due to this method has been used in the previous study described in the chapter. 6, thus, the detailed process of this part will not be repeated here again.

In Fig. 7-1 (b), the crystalline peaks of cellulose I are denoted as peak1, peak2, and peak3, and those of cellulose II are denoted as peak4, peak5, and peak6. The changes in the XRD patterns after the NaOH treatment can be easily identified; under treatment with 0% to 8% NaOH, typical cellulose I patterns can be observed. With the increase in NaOH concentration, the pattern treated with 10% NaOH showed a superimposition of the cellulose I and cellulose II patterns, and the samples treated with 12% to 20% NaOH showed patterns dominated by the cellulose II pattern. The XRD patterns after the background subtraction and baseline correction are arranged in Fig. 7-1 (c) with the color gradient expressing the different concentrations of NaOH during the treatment.

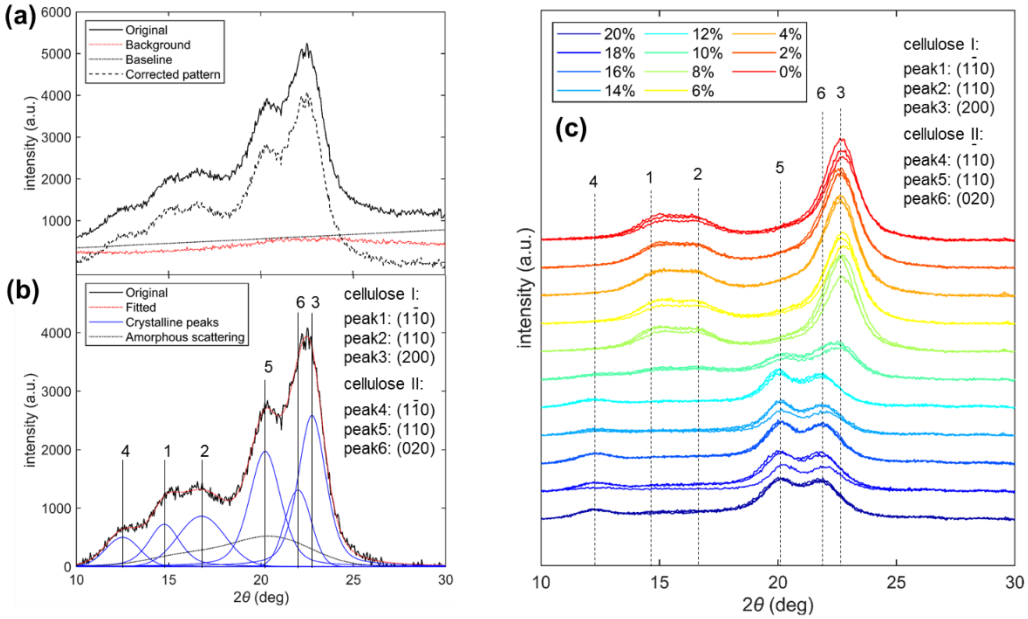

Fig.7-1 Curve-fitting process of a 10 % NaOH treated cellulose sample. **(a)** Background subtraction and baseline correction, **(b)** deconvolution of peaks with a pseudo-Voigt profile, where the amorphous scattering in the figure was the sum of the amorphous intensity curves of cellulose I and II, **(c)** the baseline corrected XRD patterns of all cellulose samples, from top to bottom the gradient colors were used to express the different concentrations of NaOH during the treatment.

The THz absorption coefficient spectra were calculated by using Eq. (2-6) and Eq. (2-7), and the original THz absorption coefficient spectra from 0.2 THz to 3.5 THz were shown in Fig. 7-2 (a), all the absorption coefficient spectra were first corrected the baseline fluctuations with SNV and then smoothed by the application of a Savitzky–Golay filter with a second-order polynomial and twenty-one smoothing points to remove the noise, as shown in Fig. 7-2 (b), with the same color gradient used in Fig. 7-1.

The THz absorption coefficient spectra of all the cellulose samples (including those treated with/without NaOH) showed a similar change trend as that of the XRD patterns. The absorption characteristics are shown in the THz absorption coefficient spectra of cellulose I and cellulose II totally differed from those of other spectroscopy, such as FTIR spectra (which show a peak shift after NaOH treatment of cellulose (Oh et al. 2005)). At NaOH concentrations below 10%, the THz absorption coefficient spectra showed the typical characteristics of cellulose I observed in previous studies in the chapter. 4 and 5: the absorption peak at 2.13 THz corresponded to the cellulose I_β type, and the absorption peak at 3.04 THz correlated with the amount of cellulose I, regardless of the I_α and I_β allomorphs. For the 20% NaOH–treated cellulose samples, which were almost cellulose II, the absorption coefficient spectra showed different characteristics of cellulose I; two absorption peaks (at 1.32 THz and 2.77 THz) can be directly observed. Given that the absorption peaks in the THz region of cellulose II were smaller than those of cellulose I, the Savitzky–Golay second derivative of the THz absorption coefficient spectra was examined to distinguish the absorption peaks clearly as shown in Fig. 7-2 (c). The peaks at 1.32 THz, 1.76 THz, and 2.77 THz could be easily observed, whereas the peak at 1.76 THz was almost invisible in the THz absorption coefficient spectra. The THz absorption coefficient spectra of cellulose treated with 10% NaOH showed the characteristics of both cellulose I and cellulose II as that in the XRD patterns.

The peak positions of the THz absorption coefficient spectra, 2θ values, and d-spacing values calculated from the XRD patterns of different types of cellulose are summarized in Table 7-1. The d-spacing values were calculated using Eq. (5-2). As shown in Table 7-1, the change in the crystalline structures was reflected in the observed d-spacing values, the 2θ of the XRD patterns, and the peak positions of the THz absorption coefficient spectra. Generally speaking, the higher the frequency, the greater the photon energy. Cellulose I—whether I_β or I_α—showed a common

absorption peak at around 3.04 THz (refer to chapter.5). The THz absorption peak at 2.13 THz of cellulose I_β was located at a lower frequency compared with that of cellulose I_α at 2.38 THz, which may be caused by the more stable crystalline structure of cellulose I_β than that of cellulose I_α. The THz absorption peaks of cellulose II were distributed at a lower frequency, which may be related to the more stable and lower overall energy of cellulose II (Langan et al. 1999, 2001). As shown in Table 7-1, the 2θ of the same cellulose was inconsistent in XRD patterns, on the other hand, the position of the absorption peak of the THz absorption coefficient spectrum is very definite, only varying due to the crystalline form of cellulose. However, the specific assignments of these absorption peaks at the molecular level still need further research.

Table 7-1. *2θ* and d-spacing values calculated from the XRD patterns and the peak positions of THz absorption coefficient spectra

	XRD pattern						**THz absorption spectrum**		
	***2θ* (°)**			**d-spacing (nm)**			**peak position (THz)**		
	$(1\bar{1}0)$	**(110)**	**(200)/(020)***	**$(1\bar{1}0)$**	**(110)**	**(200)/(020)***	**peak1**	**peak2**	**peak3**
cellulose I_α (*Glaucocystis*)[a]	14.71	17.04	22.89	0.602	0.520	0.389	2.38	3.04	NaN
	14.58	16.93	22.78	0.607	0.524	0.390	2.38	3.04	NaN
	14.57	16.91	22.77	0.608	0.524	0.391	2.38	3.04	NaN
cellulose I_β (*Halocynthia*) [a]	14.89	16.7	22.98	0.595	0.531	0.387	2.13	3.04	NaN
	14.94	16.75	23.04	0.593	0.529	0.386	2.13	3.04	NaN
	15.03	16.84	23.1	0.589	0.527	0.385	2.13	3.04	NaN
cellulose I_β (without NaOH treated MCC)	14.50	17.00	22.50	0.611	0.522	0.395	2.13	3.04	NaN
	14.54	16.81	22.50	0.610	0.522	0.395	2.13	3.04	NaN
	14.54	16.99	22.50	0.611	0.522	0.395	2.13	3.04	NaN
cellulose II (20% NaOH treated MCC)	12.50	20.24	22.00	0.709	0.434	0.404	1.32	1.76	2.77
	12.50	20.45	21.09	0.708	0.440	0.404	1.32	1.76	2.77
	12.50	20.50	22.00	0.708	0.439	0.404	1.32	1.76	2.77

*The Miller indices for cellulose I and cellulose II were (200) and (020), respectively.

[a]The values were adopted from chapter 5.

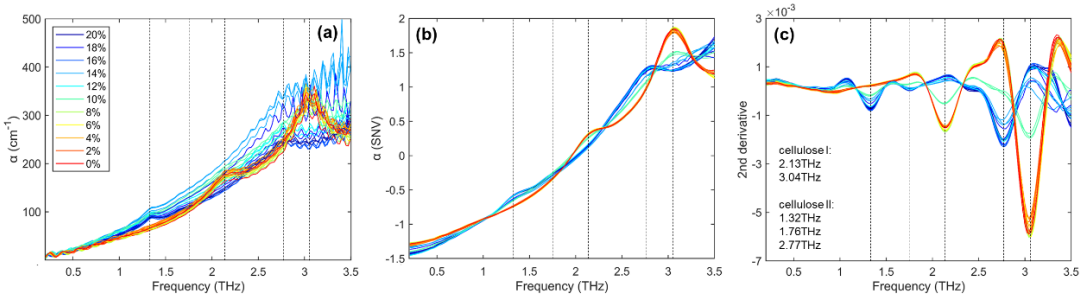

Fig. 7-2 THz spectra of cellulose treated with NaOH of different concentrations. **(a)** The original THz absorption coefficient spectra, **(b)** SNV and smoothed THz absorption coefficient spectra, **(c)** second derivative of the THz absorption coefficient spectra.

7.3.2 Crystallinity and relative content of cellulose I and II (determined by XRD and THz)

The crystallinity index (CrI) calculated from the XRD patterns was determined by Eq. (5-4), and CrI was divided into CrI1 (for cellulose I) and CrI2 (for cellulose II), which were calculated as follows:

$$CrI1 = \frac{S_{Cr1}}{S_{Cr1} + S_{Cr2}} \cdot CrI \qquad (7-1)$$

$$CrI2 = \frac{S_{Cr2}}{S_{Cr1} + S_{Cr2}} \cdot CrI \qquad (7-2)$$

where S_{Cr1} and S_{Cr2} are the sums of the integrated intensity of the crystalline peaks of cellulose I and cellulose II, respectively.

CrI1 and CrI2 can be used to evaluate the relative content of cellulose I and cellulose II in the samples to grasp the progress of the NaOH treatment. The correlations of the concentrations of NaOH with CrI1 and CrI2 are shown in Figs. 7-4 (a) and (b), respectively. CrI1 and CrI2 showed completely opposite trends with the change in the NaOH concentration, where CrI1 decreased with the increase in the NaOH concentration, indicating a decrease in the cellulose I in the samples. In contrast, CrI2 increased with the NaOH concentration, which showed a transformation from cellulose I to cellulose II. Both Fig. 7-4 (a) and (b) showed inflection points at the 10% NaOH concentration, which indicated that the crystalline lattice transformed from cellulose I to cellulose II, as shown by the XRD patterns in Fig. 7-1 (c).

As shown in Fig. 7-2, the THz absorption coefficient spectra of the 10% NaOH–treated sample also showed mixed characteristics of cellulose I and cellulose II, the other samples showed a characteristic of cellulose I (treated with NaOH < 10%) and cellulose II (treated with NaOH > 10%), respectively. The changes of the intensity are more easily observed in the second derivative of the

THz absorption coefficient spectra, where the troughs correspond to the absorption peaks in the THz absorption coefficient spectra, and the reversed second derivative value reflects the intensity changes of the absorption peaks directly. As shown in Fig. 7-4 (c) and (d), the reversed second derivative values at 3.04 THz and 2.77 THz showed a similar trend of change of CrI1 and CrI2 calculated from XRD patterns.

Given the simplicity of the THz absorption coefficient and the above identification of the absorption peaks of cellulose I and cellulose II, the THz absorption coefficient spectra of the cellulose I without NaOH treatment and the cellulose II with 20% NaOH treatment can be written as two seventh-order Fourier series (denoted as $f(0)$ and $f(20)$). Furthermore, the expression of the THz absorption coefficient spectra of the samples treated with NaOH of other concentrations can be written as $f = r1 \cdot f(0) + r2 \cdot f(20)$, where *r1* and *r2* are coefficients determined by the curve-fitting process, and *r1* + *r2* = 1, since the samples only have two types of cellulose crystalline. As shown in Fig. 7-3, the fitted results reproduced the measured THz absorption coefficient spectra (only after SNV to correct the baseline fluctuations) to a certain degree. The specific parameters of the fitted Fourier series ($f(0)$ and $f(20)$) and the residual distribution are provided below:

coefficients for *f(0)* fitted of cellulose I (0% NaOH treated samples):

Coefficients (with 95% confidence bounds):

a0 = -7.899e+07 (-1.646e+09, 1.488e+09)

a1 = 1.153e+08 (-2.26e+09, 2.491e+09)

b1 = 7.835e+07 (-1.342e+09, 1.499e+09)

a2 = -3.512e+07 (-9.34e+08, 8.637e+08)

b2 = -8.857e+07 (-1.759e+09, 1.582e+09)

a3 = -1.072e+07 (-6.136e+07, 3.992e+07)

b3 = 4.85e+07 (-9.439e+08, 1.041e+09)

a4 = 1.389e+07 (-1.98e+08, 2.258e+08)

b4 = -1.32e+07 (-3.354e+08, 3.09e+08)

a5 = -5.094e+06 (-9.923e+07, 8.904e+07)

b5 = 8.716e+05 (-4.52e+07, 4.694e+07)

a6 = 7.979e+05 (-1.698e+07, 1.858e+07)

b6 = 3.541e+05 (-1.242e+06, 1.95e+06)

a7 = -3.763e+04 (-1.216e+06, 1.141e+06)

b7 = -5.867e+04 (-8.837e+05, 7.663e+05)

w = 0.381 (-0.2115, 0.9735)

coefficients for *f(20)* fitted of cellulose II (20% NaOH treated samples):

Coefficients (with 95% confidence bounds):

a0 = 1425 (-4525, 7375)

a1 = -418.7 (-3193, 2355)

b1 = -2555 (-1.301e+04, 7902)

a2 = -1834 (-8864, 5197)

b2 = 616.2 (-3402, 4635)

a3 = 549.4 (-2914, 4013)

b3 = 1035 (-2458, 4529)

a4 = 444.8 (-737.4, 1627)

b4 = -341.1 (-2382, 1700)

a5 = -147.8 (-967.7, 672.1)

b5 = -136.6 (-353.1, 79.92)

a6 = -26.23 (-46.92, -5.542)

b6 = 41.31 (-163.3, 245.9)

a7 = 5.707 (-18.15, 29.57)

b7 = 2.23 (-6.933, 11.39)

w = 0.8129 (0.5954, 1.03)

The correlations of the NaOH concentrations with the coefficients *r1* and *r2* are shown in Fig. 7-4 (e) and (f). These correlations were highly similar with those of CrI1, CrI2 and reversed second derivative values of THz absorption coefficient spectra. This indicated that the coefficients obtained from the fitting of the THz absorption coefficient can be used to evaluate the relative content of cellulose I and cellulose II. Furthermore, for the samples which only have two types of cellulose crystalline components, the coefficients *r1* and *r2* still may give a more accurate relative content of each crystalline component, since the integrated intensity of the THz mass absorption coefficient

peak was very possible to directly reflect the amount of the cellulose crystalline without the interference of the amorphous region on the THz absorption coefficient spectra.

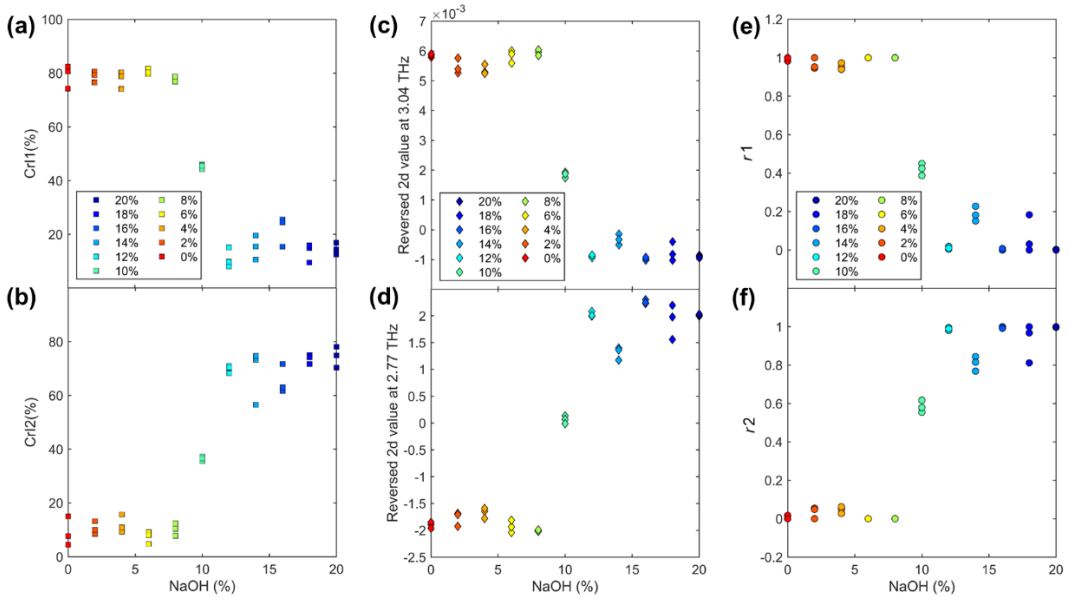

Fig. 7-4 Correlation of the concentrations of NaOH with **(a)** CrI1 and **(b)** CrI2 calculated from the XRD patterns, reversed second derivative at **(c)** 3.04 THz and **(d)** 2.77 THz, and coefficients **(e)** *r*1 and **(f)** *r*2 fitted from the THz absorption coefficient spectra.

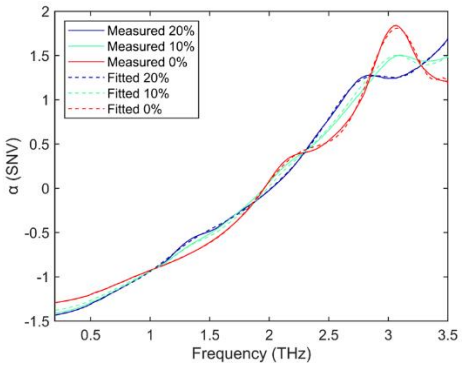

Fig. 7-3 THz absorption coefficient spectra of cellulose treated with 20% (blue), 10% (green), and 0% (red) NaOH that are measured (after SNV) and curve fitted using a 7th Fourier series.

7.3.3 Evaluation the crystallinity of ball-milled cellulose II with XRD and THz

The cellulose II obtained from the 20% NaOH treated cellulose I were ball-milled at different times to change the crystallinity, and the original XRD patterns cut from 10° to 30° were shown in Fig. 7-5 (a), the intensity was decreased as an increasing ball-milled time. Same as described in the last section, the measured XRD patterns were first subtracted the background, and the baseline which fitted by a first-order polynomial as shown in Fig.7-1 (a). After this process, the crystalline peaks and the amorphous scattering curve were deconvoluted by using a pseudo-Voigt profile. Where the specific procedure was the same as described in the last section, the crystalline peaks was

fixed at a range of 12° ± 0.5°, 20° ± 0.5°, and 21.5° ± 0.5°, corresponding to the Miller indices of $(1\bar{1}0)$, (110), (020), and the amorphous scattering curve was fixed at 16° as shown in Fig. 7-5 (b).

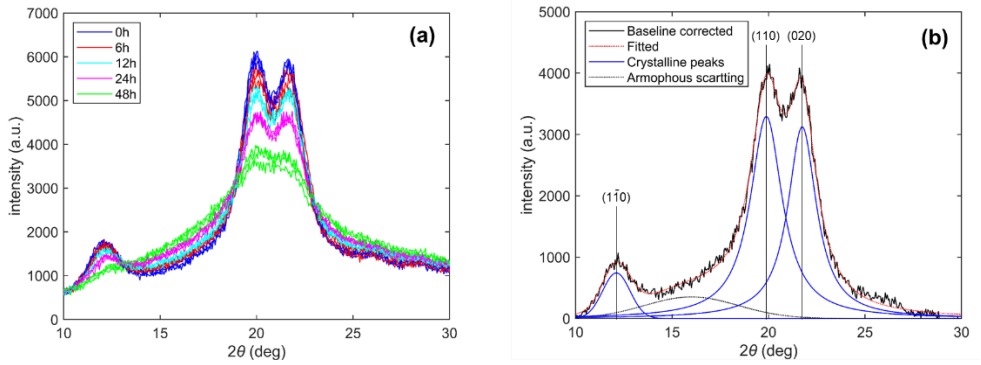

Fig. 7-5 The original and preprocessing of the XRD patterns of the ball-milled cellulose II, **(a)** original XRD patterns, and **(b)** the peak deconvolution process.

The crystallinity index was determined by Eq. (5-4), and the correlation between the ball-milled time and the crystallinity index was shown in Fig. 7-7 (a). The crystallinity index of ball-milled cellulose II was in a range from about 90% to 80%, where the crystallinity index decreased as an increasing ball-mill time. A non-linear correlation was observed, where for the ball-milled time from 0h to 12h, the crystallinity index was decreased more sharply than that of the ball-milled time from 24h to 48h.

The THz absorption coefficient spectra of the ball-milled cellulose II was calculated used Eq. (2-6) and Eq. (2-7), and the THz absorption coefficient spectra after SNV and smoothing were shown in Fig.7-6 (a). The intensity of the absorption peaks at 2.77 THz of the cellulose II was relatively small than that of the cellulose I at 3.04 THz (Fig. 6-5). Similar with the satiation of the ball-milled cellulose I, without the peak shift, only the intensity of the absorption peaks decreased as an increase of the ball-milled time. The peaks at 1.32 THz and 1.76 THz were almost invisible. The THz absorption coefficient spectra from 2.29 THz to 3.02 THz was cut out and then subtracted the baseline that was fitted as a first-order polynomial as shown in Fig.7-6 (b).

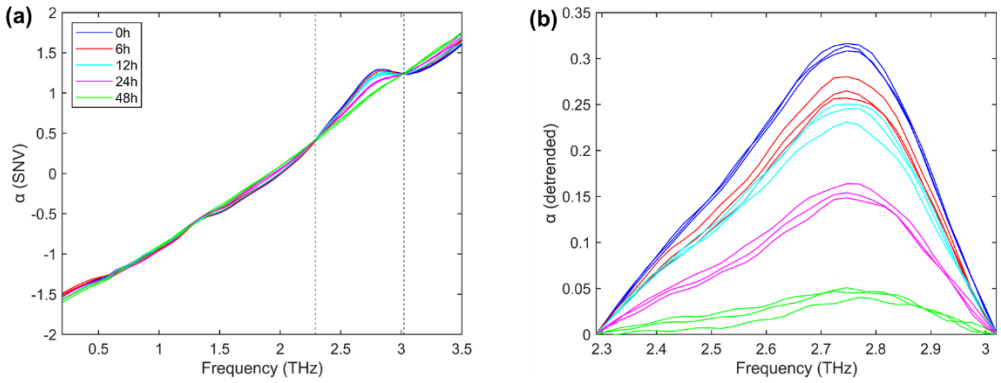

Fig. 7-6 THz absorption coefficient spectra of cellulose II, **(a)** SNV and smoothed absorption coefficient spectra, and **(b)** detrended THz absorption coefficient spectra obtained by subtraction of the baseline form 2.29 THz to 3.02 THz

The integrated intensity of the detrended absorption coefficient spectra was calculated, the correlation between the ball-milled time and the integrated intensity was shown in Fig. 7-7 (b). Where the trend of the changes of the integrated intensity calculated from the THz absorption coefficient spectra was similar to that of the crystallinity calculated from the XRD patterns, where the integrated intensity decreased as an increase of the ball-milled time, indicated the ability of the THz absorption coefficient spectra to evaluate the crystallinity of cellulose II. However, compared with the correlation obtained from the XRD patterns, the correlation obtained from the THz absorption coefficient spectra was linear-like, and for each ball-milled time, the float of the integrated intensity was also smaller than the crystallinity calculated from the XRD patterns. The reason for this result was discussed in the previous chapters, where the peak deconvolution process of the XRD patterns was more complicated compared with the calculation of the integrated intensity of THz absorption peaks, and small fluctuations in parameters will bring about large changes in results. Further, the peak deconvolution process of the XRD patterns depends on the experience of the analyst, since the profiles of the peaks and the parameters such as the peak position and the limit of the intensity will change by the choice made by the analyst. On the other hand, once the cutting range of the THz absorption peaks was determined, the integrated intensity will not change, that is, the analysis operation will be easier.

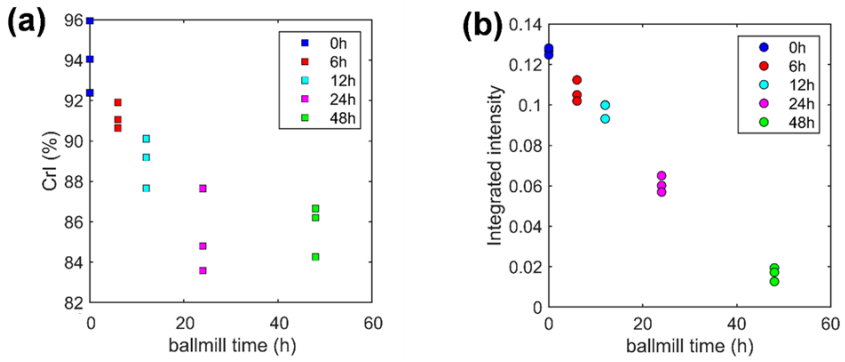

Fig. 7-7 Correlations between the ball-milled time with **(a)** the crystallinity calculated from the XRD patterns, and **(b)** the integrated intensity calculated from the THz absorption coefficient spectra.

7.4 Conclusions

In this study cellulose I was treated with NaOH of different concentrations (from 0% to 20%) to change the crystalline structure. For the THz absorption coefficient spectra of cellulose I, only two absorption peaks (at 2.13THz and 3.04 THz) could be observed. By contrast, three absorption peaks (at 1.32 THz, 1.76 THz, and 2.77 THz) could be observed for cellulose II. The correlations between the concentrations of NaOH and *r1* and *r2* (coefficients calculated from a fitting process of the THz absorption spectra) showed similar changing trends of CrI1 and CrI2 calculated from XRD. This indicated that the coefficients obtained from the fitting of the THz absorption coefficient can be used to evaluate the relative content of cellulose I and cellulose II. Further, the 20% NaOH treated cellulose, that is, the cellulose II samples were ball-milled at different times to change the crystallinity. Same as the absorption peaks of the ball-milled cellulose I at 3.04 THz without a peak shift, only the intensity of the THz absorption peaks at 2.77 THz for the ball-milled cellulose II was decreased. The integrated intensities were calculated from detrended THz absorption coefficient spectra from 2.29 THz to 3.02 THz, the integrated intensities showed a similar changing trend with the crystallinity index obtained from XRD patterns, which were decreased with an increase of ball-milled time. This means the THz absorption coefficient spectra can be used to evaluate not only the transformation of crystalline lattice from cellulose I to cellulose II after the NaOH treatment but also the crystallinity of cellulose I and cellulose II.

8. General conclusions

In this study, THz-TDS was used to investigate various cellulosic materials, overall:

1. By applying an MC-dependent dielectric function of water in the THz region of the wood cell wall, the prediction accuracy of the density and MC of wood was improved significantly, which indicates that THz-TDS can be used to predict the physical properties of wood with good accuracy.

2. With an MLR model that is easier to build by the refractive index and absorption coefficient in the THz region (compare with the EMT model using dielectric function), the properties of wood, such as the density and MOE can be predicted well. However, the prediction accuracy of MFA and cellulose crystallinity was limited. This result indicated the possibility of using THz-TDS to predict the physical properties of cellulosic materials simultaneously. Thus, in order to improve the prediction accuracy of MFA and cellulose crystallinity, the following study is focused on the optical properties in the THz region of various cellulose crystallines.

3. Native cellulose I_α and I_β allomorphs were investigated by THz-TDS and XRD, in the XRD patterns, the difference between these allomorphs was difficult to identify since there are only slight changes in peak positions that can be observed. On the other hand, cellulose I_α and I_β showed a characteristic absorption at 2.38 THz and 2.11 THz, respectively. Furthermore, different from the relative content can not be calculated directly in the XRD patterns, the intensity of the absorption peaks in the THz region was directly correlated with the relative content of the component, which makes the evaluation of the cellulose allomorphs in the THz region is easier and simpler.

4. Following the previous research, cellulose crystallinity was investigated by THz-TDS. The standard of evaluating the cellulose crystallinity has been used, CrI which is just a relative parameter calculated from XRD patterns and the calculation was dependent on the experience of the analyst, where the available profile for the peak deconvolution is not unique. On the other hand, cellulose I regardless of I_α or I_β all showed absorption peaks around 3.04 THz, the integrated intensity of the absorption peaks at 3.04 THz was directly correlated with the amount of cellulose I, which showed a more reasonable result of the cellulose crystallinity compared with CrI.

5. The changing of the cellulose crystalline under a chemical treatment was traced by using THz-TDS, cellulose I and cellulose II was shown a different characteristic absorption profile, by fitting the THz absorption spectra, the coefficient r1 and r2 can be used to describe the relative content of

cellulose I and II, respectively. Furthermore, the changing of cellulose crystallinity of cellulose II after ball-milling was investigated with the integrated intensity of the absorption peaks in the THz region, where the peak position was at 2.77 THz. Compared with the CrI that was calculated from the XRD pattern, the integrated intensity showed less scatter of data. Considering the result of cellulose I in the chapter.6, it is believed that the integrated intensity in the THz region is the preferred method when evaluating the crystallinity of cellulose. However, the absorption peaks assignment in the THz region at a molecular dynamic level is still not certain, further research is needed to understand the behavior of absorption in the THz region in principle.

In summary, THz-TDS was used for investigating various cellulosic materials. The THz signal can be measured at room temperature and does not require sample pretreatment, the measurement and analysis processes are rapid and simple compared with the XRD patterns. As the results shown in this thesis, we believe THz-TDS has the potential to become a useful tool in the research and understanding of cellulose crystallography. Especially when analyzing the cellulose crystallinity, the results are better than XRD, which demonstrates the great potential of THz-TDS in the research of cellulosic materials.

9. Acknowledgement

First and foremost, I want to thank my primary supervisor, Associate Professor. Tetsuya Inagaki, who is also the man led me down the academic path. He gave me not only the most careful guidance in academics but also show me a way of knowing the world. When I decided to continue to get a Ph.D., he told me the most important thing during the Ph.D. course is not how many publications I achieved, but in this lonely process to maintain the stability of my mental state. As it happens, not long after the Ph.D. course started, the world was hit by the COVID-19, I cannot go back to my hometown, and was isolated from my friends and family. In fact, when I write down this, I still cannot go back to my homeland. But I am not alone here, we have talked about so many things, including philosophy, history, society, and the universe. This has undoubtedly brought me a lot of new ideas about the world, and I have gradually matured in thinking like never before. Although we will be separated in the end, the way you left me about knowing the world, about your wisdom, will always walk with me and ever glow on my way.

Similar gratitude belongs to my secondary supervisor, Professor. Satoru Tsuchikawa. As an international student, he gave me much more care than I could have ever imagined both in life and academically. When I needed to work part-time to pay tuition and living expenses, he always cared about my health and life situation, just like a kind elder in my family. He is very hard working and optimistic, and I regard him as a role model. Without his initial permission, I could not have come to this laboratory to start this research, nor would I have achieved the current results. Maybe my academic path will take a completely different direction.

Besides, I want to express my gratitude to Dr. Te Ma, he is like my big brother, cares about my life, and discussed the experimental plan with me in detail. He always guides us with the sincerest advice and concern.

I also want to thank my thesis committee: Professor. Kazuhiko Fukushima, Professor. Hiroyuki Yamamoto, Associate Professor. Masato Yoshida, for their insightful comments and encouragement. And I want to thank Associate Professor. Yoshiki Horikawa (Tokyo University of Agriculture and Technology), and Masako Seki (AIST) with the help in my research.

And, I want to thank the members of the System Engineering for Biological Resources laboratory, the staffs at Nagoya University Graduate School of Bioagricultural Sciences for their continued

support. I also owe a great debt of gratitude to the THERS Interdisciplinary Frontier Next Generation Researcher program, and Mizutani scholarship for their financial support, which allows me to focus on research.

I am very grateful to my parents and friends. Even though we haven't seen each other for a really long time, because of your trust, support and love, despite many difficulties in the process, I was still able to persevere and have not given up until now.

At last, special thanks to the radios, movies, music, and games. Lu Zhao and Laoju Wang, my favorite voice actor and YouTuber, respectively, is always a joy for me to wait for their new works. Ran Zuo (Libra, from Tears of Themis), Diluc Ragnvindr, and Ayato Kamisato (from Genshin Impact), even if the characters are virtual, they still brought me a lot of irreplaceable happiness.

Because of you all, I can be as sure as in the book that my 20s were my precious golden age, and I will always miss those days that shine like a reflection on the sea.

2022.07.06

王 晗

@Nagoya

10. Reference

Atalla RH, VanderHart DL (1984) Native cellulose: A composite of two distinct crystalline forms. Science (80-) 223:283–285.

Barnett JR, Bonham VA (2004) Cellulose microfibril angle in the cell wall of wood fibres. Biol Rev Camb Philos Soc 79:461–472.

Baxter JB, Guglietta GW (2011) Terahertz spectroscopy. Anal Chem 83:4342–4368.

Belton PS, Tanner SF, Cartier N, Chanzy H (1989) High-Resolution Solid-State 13C Nuclear Magnetic Resonance Spectroscopy of Tunicin, an Animal Cellulose. Macromolecules 22:1615–1617.

Bhuiyan MTR, Hirai N, Sobue N (2000) Changes of crystallinity in wood cellulose by heat treatment under dried and moist conditions. J Wood Sci 46:431–436.

Born B, Weingärtner H, Bründermann E, Havenith M (2009) Solvation dynamics of model peptides probed by terahertz spectroscopy. observation of the onset of collective network motions. J Am Chem Soc 131:3752–3755.

Bragg WH, Bragg WL (1913) The reflection of X-rays by crystals. Proc R Soc London Ser A, Contain Pap a Math Phys Character 88:428–438

de Keijser T, Mittemeijer EJ, Rozendaal HCF (1983) The determination of crystallite-size and lattice-strain parameters in conjunction with the profile-refinement method for the determination of crystal structures. J Appl Crystallogr 16:309–316.

del Cerro DR, Koso T V., Kakko T, et al (2020) Crystallinity reduction and enhancement in the chemical reactivity of cellulose by non-dissolving pre-treatment with tetrabutylphosphonium acetate. Cellulose 27:5545–5562.

Duvillaret L, Garet F, Coutaz JL (1996) A reliable method for extraction of material parameters in terahertz time-domain spectroscopy. IEEE J Sel Top Quantum Electron 2:739–745.

Ebbinghaus S, Seung JK, Heyden M, et al (2007) An extended dynamical hydration shell around proteins. Proc Natl Acad Sci U S A 104:20749–20752.

Federici JF, Schulkin B, Huang F, et al (2005) THz imaging and sensing for security applications - Explosives, weapons and drugs. Semicond Sci Technol 20:.

Ferguson B, Zhang XC (2002) Materials for terahertz science and technology. Nat Mater 1:26–33.

Fischer BM, Walther M, Jepsen PU (2002) Far-infrared vibrational modes of DNA components studied by terahertz time-domain spectroscopy. Phys Med Biol 47:3807–3814.

Fujimoto T, Yamamoto H, Tsuchikawa S (2007) Estimation of wood stiffness and strength properties of hybrid larch by near-infrared spectroscopy. Appl Spectrosc 61:882–888.

Hashin Z (1983) Analysis of Composite Materials—A Survey. J Appl Mech 50:481.

Horii F, Hirai A, Kitamaru R (1987) CP/MAS13C NMR Spectra of the Crystalline Components of Native Celluloses. Macromolecules 20:2117–2120.

Horikawa Y (2017) Assessment of cellulose structural variety from different origins using near infrared spectroscopy. Cellulose 24:5313–5325.

Imai T, Sugiyama J (1998) Nanodomains of Iα and Iβ cellulose in algal microfibrils. Macromolecules 31:6275–6279.

Imai T, Sugiyama J, Itoh T, Horii F (1999) Almost pure I(α) cellulose in the cell wall of Glaucocystis. J Struct Biol 127:248–257.
Inagaki T, Ahmed B, Hartley ID, et al (2014a) Simultaneous prediction of density and moisture content of wood by terahertz time domain spectroscopy. J Infrared, Millimeter, Terahertz Waves 35:949–961.
Inagaki T, Hartley ID, Tsuchikawa S, Reid M (2014b) Prediction of oven-dry density of wood by time-domain terahertz spectroscopy. Holzforschung 68:61–68.
Inagaki T, Mitsui K, Tsuchikawa S (2009) Near-Infrared Spectroscopic Investigation of the Hydrothermal Degradation Mechanism of Wood as an Analogue of Archaeological Objects. Part I: Softwood. Appl Spectrosc 63:753–758.
Jepsen PU, Møller U, Merbold H (2007) Investigation of aqueous alcohol and sugar solutions with reflection terahertz time-domain spectroscopy. Opt Express 15:14717.
Kafle K, Greeson K, Lee C, Kim SH (2014) Cellulose polymorphs and physical properties of cotton fabrics processed with commercial textile mills for mercerization and liquid ammonia treatments. Text Res J 84:1692–1699.
Kawase K, Ogawa Y, Watanabe Y, Inoue H (2003) Non-destructive terahertz imaging of illicit drugs using spectral fingerprints. Opt Express 11:2549.
Keunecke D, Novosseletz K, Lanvermann C, et al (2012) Combination of X-ray and digital image correlation for the analysis of moisture-induced strain in wood: Opportunities and challenges. Eur J Wood Wood Prod 70:407–413.
Kirkpatrick S (1971) Classical Transport in Disordered Media: Scaling and Effective-Medium Theories. Phys Rev Lett 27:1722–1725
Koch M, Hunsche S, Schuacher P, et al (1998) THz-imaging: A new method for density mapping of wood. Wood Sci Technol 32:421–427.
Kolpak FJ, Blackwell J (1976) Determination of the Structure of Cellulose II. Macromolecules 9:273–278.
Kore PS, Pawar PP (2014) Measurements of mass attenuation coefficient, effective atomic number and electron density of some amino acids. Radiat Phys Chem 98:86–91.
Langan P, Nishiyama Y, Chanzy H (1999) A revised structure and hydrogen-bonding system in cellulose II from a neutron fiber diffraction analysis. J Am Chem Soc 121:9940–9946.
Langan P, Nishiyama Y, Chanzy H (2001) X-ray structure of mercerized cellulose II at 1 Å resolution. Biomacromolecules 2:410–416.
Langkilde FW, Svantesson A (1995) Identification of celluloses with Fourier-Transform (FT) mid-infrared, FT-Raman and near-infrared spectrometry. J Pharm Biomed Anal 13:409–414.
Lindgren LO (1991) Medical CAT-scanning: X-ray absorption coefficients, CT-numbers and their relation to wood density. Wood Sci Technol 25:341–349.
Ma T, Inagaki T, Tsuchikawa S (2017) Calibration of SilviScan data of Cryptomeria japonica wood concerning density and microfibril angles with NIR hyperspectral imaging with high spatial resolution. Holzforschung 71:341–347.
Mantsch HH, Naumann D (2010) Terahertz spectroscopy: The renaissance of far infrared spectroscopy. J Mol Struct 964:1–4.

Mourou G, Stancampiano C V., Antonetti A, Orszag A (1981) Picosecond microwave pulses generated with a subpicosecond laser-driven semiconductor switch. Appl Phys Lett 39:295–296.
Mwaikambo LY, Ansell MP (2001) The determination of porosity and cellulose content of plant fibers by density methods. J Mater Sci Lett 20:2095–2096.
Nahata A, Weling AS, Heinz TF (1996) A wideband coherent terahertz spectroscopy system using optical rectification and electro-optic sampling. Appl Phys Lett 69:2321–2323.
Nakamura K, Wada M, Kuga S, Okano T (2004) Poisson's Ratio of Cellulose Iα and cellulose II. J Polym Sci Part B Polym Phys 42:1206–1211.
Nishiyama Y (2009) Structure and properties of the cellulose microfibril. J Wood Sci 55:241–249.
O'sullvian AC (1997) Cellulose: the structure slowly unravels. Cellulose 4:173–207
Oh SY, Dong IY, Shin Y, et al (2005) Crystalline structure analysis of cellulose treated with sodium hydroxide and carbon dioxide by means of X-ray diffraction and FTIR spectroscopy. Carbohydr Res 340:2376–2391.
Oudiani A El, Chaabouni Y, Msahli S, Sakli F (2011) Crystal transition from cellulose i to cellulose II in NaOH treated Agave americana L. fibre. Carbohydr Polym 86:1221–1229.
Park S, Baker JO, Himmel ME, et al (2010) Cellulose crystallinity index: Measurement techniques and their impact on interpreting cellulase performance. Biotechnol Biofuels 3:1–10.
PETTERSEN RC (1984) The Chemical Composition of Wood. In: The Chemistry of Solid Wood. pp 57–126
Pickwell E, Wallace VP (2006) Biomedical applications of terahertz technology. J Phys D Appl Phys 39:.
Popescu MC, Popescu CM, Lisa G, Sakata Y (2011) Evaluation of morphological and chemical aspects of different wood species by spectroscopy and thermal methods. J Mol Struct 988:65–72.
Reid M, Fedosejevs R (2006) Terahertz birefringence and attenuation properties of wood and paper. Appl Opt 45:2766–2772
Ren A, Zahid A, Fan D, et al (2019) State-of-the-art in terahertz sensing for food and water security – A comprehensive review. Trends Food Sci Technol 85:241–251.
Rice A, Jin Y, Ma XF, et al (1994) Terahertz optical rectification from 〈110〉 zinc-blende crystals. Appl Phys Lett 64:1324–1326.
Schenzel K, Fischer S (2001) NIR FT Raman spectroscopy - A rapid analytical tool for detecting the transformation of cellulose polymorphs. Cellulose 8:49–57.
Shen YC, Upadhya PC, Beere HE, et al (2004) Generation and detection of ultrabroadband terahertz radiation using photoconductive emitters and receivers. Appl Phys Lett 85:164–166.
Sugino H, Sugimoto H, Miki T, Kanayama K (2007) Fine Structure Changes of Wood during Moisture Adsorption and Desorption Process Analyzed by X-ray Diffraction Measurement. Mokuzai Gakkaishi/Journal Japan Wood Res. Soc. 53:82–89
Sugiyama J, Vuong R, Chanzy H (1991) Electron Diffraction Study on the Two Crystalline Phases Occurring in Native Cellulose from an Algal Cell Wall. Macromolecules 24:4168–4175.
Takeuchi I, Tomoda K, Nakajima T, et al (2012) Estimation of crystallinity of trehalose dihydrate microspheres by usage of terahertz time-domain spectroscopy. J Pharm Sci 101:3465–3472.

Tanaka S, Shiraga K, Ogawa Y, et al (2014) Applicability of effective medium theory to wood density measurements using terahertz time-domain spectroscopy. J Wood Sci 60:111–116.
Todoruk TM, Hartley ID, Reid ME (2012) Origin of birefringence in wood at terahertz frequencies. IEEE Trans Terahertz Sci Technol 2:123–130.
Todoruk TM, Schneider J, Hartley ID, Reid M (2008) Birefringence of wood at terahertz frequencies - art. no. 70992Q. Photonics North 2008 7099:Q992-Q992\r36.
VanderHart DL, Atalla RH (1984) Studies of Microstructure in Native Celluloses Using Solid-State 13C NMR. Macromolecules 17:1465–1472.
Vieira FS, Pasquini C (2014) Determination of cellulose crystallinity by terahertz-time domain spectroscopy. Anal Chem 86:3780–3786.
Wada M, Okano T (2001) Localization of Iα and Iβ phases in algal cellulose revealed by acid treatments. Cellulose 8:183–188.
Wada M, Okano T, Sugiyama J (2001) Allomorphs of native crystalline cellulose I evaluated by two equatorial d-spacings. J Wood Sci 47:124–128.
Wada M, Okano T, Sugiyama J (1997) Synchrotron-radiated X-ray and neutron diffraction study of native cellulose. Cellulose 4:221–232.
Walther M, Plochocka P, Fischer B, et al (2002) Collective vibrational modes in biological molecules investigated by terahertz time-domain spectroscopy. Biopolym - Biospectroscopy Sect 67:310–313.
Williams GP (2006) Filling the THz gap - High power sources and applications. Reports Prog Phys 69:301–326.
Yao W, Weng Y, Catchmark JM (2020) Improved cellulose X-ray diffraction analysis using Fourier series modeling. Cellulose 27:5563–5579.
Zeitler JA, Kogermann K, Rantanen J, et al (2007) Drug hydrate systems and dehydration processes studied by terahertz pulsed spectroscopy. Int J Pharm 334:78–84.
Zsigmondy R, Scherrer P (1912) Bestimmung der inneren Struktur und der Größe von Kolloidteilchen mittels Röntgenstrahlen. Kolloidchem Ein Lehrb 277:387–409.

List of publications

1. **Wang, H.**, Inagaki, T., Hartley, I. D., Tsuchikawa, S., & Reid, M. (2019). Determination of dielectric function of water in THz region in wood cell wall result in an accurate prediction of moisture content. *Journal of Infrared, Millimeter, and Terahertz Waves*, *40*(6), 673-687.
2. **Wang, H.**, Horikawa, Y., Tsuchikawa, S., & Inagaki, T. (2020). Terahertz time-domain spectroscopy as a novel tool for crystallographic analysis in cellulose. *Cellulose*, *27*(17), 9767-9777.
3. **Wang, H.**, Tsuchikawa, S., & Inagaki, T. (2021). Terahertz time-domain spectroscopy as a novel tool for crystallographic analysis in cellulose: The potentiality of being a new standard for evaluating crystallinity. *Cellulose*, *28*(9), 5293-5304.
4. **Wang, H.**, Kataoka, H., Tsuchikawa, S., & Inagaki, T. (2022). Terahertz time-domain spectroscopy as a novel tool for crystallographic analysis in cellulose: cellulose I to cellulose II, tracing the structural changes under chemical treatment. *Cellulose*, *29*(6), 3143-3151.