

비지도학습 기반 차원 축소 모델링을 통한 한지의 열화특성 분석

이용주¹, 차지은¹, 강나영¹, 권순완¹, 김형진^{2†}, 강광호^{3‡}

접수일(2023년 11월 21일), 수정일(2023년 12월 12일), 채택일(2023년 12월 15일)

Unsupervised Dimensionality Reduction Modeling for Analyzing Aging Characteristics of Hanji

Yong Ju Lee¹, Ji Eun Cha¹, Na Young Kang¹, Soon Wan Kweon¹, Hyoung Jin Kim^{2†}, Kwang-Ho Kang^{3‡}

Received November 21, 2023; Received in revised form December 12, 2023; Accepted December 15, 2023

ABSTRACT

This study investigated the applicability of principal component analysis (PCA) and infrared (IR) spectroscopy in the context of paper aging. Furthermore, it explored the feasibility of employing a t -distributed stochastic neighbor embedding (t -SNE) model for visualizing the aging characteristics of Hanji, a traditional Korean paper, particularly in relation to accelerated aging categories. A total of 130 Hanji IR spectra were collected via attenuated total reflectance IR spectroscopy and categorized into four groups based on the origin, aging time, and aging method (dry and moist heat). The PCA and t -SNE analysis were conducted on the first derivative of the IR spectra. This process enabled the classification and visualization of aged Hanji. Additionally, PCA identified carboxyl and carbonyl groups as critical components influencing cellulose aging. Overall, unsupervised dimensionality reduction models, such as PCA and t -SNE, along with IR spectral data, can be used to characterize the chemical properties of aged papers.

Keywords: PCA, t -SNE, dimensionality reduction, chemometrics, Hanji, aging, aged paper

1 국민대학교 과학기술대학 임산생명공학과(Department of Forest Products and Biotechnology, Kookmin University, Graduate Student), 대학원생

2 국민대학교 과학기술대학 임산생명공학과(Department of Forest Products and Biotechnology, Kookmin University, Professor), 교수

3 에이치피프린팅코리아(유) 소재분석팀(Material and Analysis Team, HP Printing Korea), 팀장

† 교신저자(Corresponding Author): E-mail: hyjikim@kookmin.ac.kr (Address: Dept. of Forest Products & Biotechnology, College of Science and Technology, Kookmin Univ. Seoul, Republic of Korea)

‡ 공동교신저자(Corresponding Author): E-mail: kwangho.kang@hp.com (Address: HP Printing Korea, 26, Yeonnaegaeul-ro, Sujeong-gu, Seongnam-si, Gyeonggi-do, Republic of Korea)

1. 서론

화학계량학(chemometrics)이란 스펙트럼이나 크로마토그램과 같은 분석화학 데이터로부터 잠재된 정보를 도출하여 정성 및 정량분석을 가능하게 하는 수학적 통계 알고리즘을 총칭하는 것으로 다변량 분석(multivariate analysis)과 함께 제품의 종류 및 원산지 판별, 진위 분석 등에 적용된다.¹⁾ 화학계량학에서 주로 이용되는 분광 분석은 분석 대상에 대한 파괴를 요하지 않아 다양한 재료의 특성화에 있어 효과적인 접근법으로 인식되고 있다.^{2,3)} 최근 보고된 다수의 연구를 통하여 분광 스펙트럼과 다변량 분석의 결합은 전통적 화학 분석법을 대체할 수 있는 효과적 도구임을 짐작할 수 있다.²⁻⁸⁾

종이의 열화란 시간이 경과함에 따라 종이 성분 자체 혹은 보존환경의 영향으로 물리·광학적 특성이 저하되는 현상을 일컫는다. 자연 노화 진행에 따른 종이의 열화 특성은 비교적 단시간에 가혹한 조건을 설정하여 종이의 가속 열화를 진행함을 통해 분석 가능하다.⁹⁻¹¹⁾ 가속 열화특성의 정밀 분석 수단으로서는 분광 스펙트럼 데이터가 과거부터 활용되어 왔음을 여러 문헌을 통해 확인할 수 있다. Nelson 등¹²⁾은 종이의 열화 과정에 따라 변화하는 셀룰로오스의 결정화도(crystallinity) 분석을 위한 IR 스펙트럼 상 wavenumber(1,430 및 900 cm⁻¹)를 발견한 바 있으며, Yousef 등¹³⁾과 Tasker 등¹⁴⁾은 IR 스펙트럼 peak intensity 비율의 계산을 통한 열화에 따른 결정화도 정량화 식을 제안하였다. 국내 연구 사례로는 Kim 등⁹⁾이 주성분 분석(principal component analysis, PCA)과 IR 스펙트럼 데이터의 결합을 통해 종이의 열화 유효 성분 peak를 도출하여 산화지수(oxidation index, OI)를 제시한 사례와 Kim 등¹⁵⁾에 의해 분광학적 특성 적용을 통한 연대별 학술지의 펄프 식별연구가 진행된 바

있음을 확인할 수 있다.

PCA(principal component analysis)는 상호 연관된 여러 변수의 차원을 축소하는 분석기법으로 각 변수들의 선형결합을 통해 무상관(uncorrelated) 변수의 집합을 계산하여 고차원인 원본 데이터를 소수의 주성분 값만으로 변환한다. 원본 데이터를 임의의 새로운 축의 값으로 사영(projection)할 때 제1주성분 축(PC1)은 사영된 값들 중 분산이 가장 큰 축이 된다.¹⁶⁾ 상기 과정은 공분산행렬에 대한 고유값(eigenvalue)과 고유벡터(eigenvector)를 구하는 과정과 동일하다. 그러나 PCA의 경우 선형 분석의 일환으로 값을 할당하므로 차원 축소에 따른 군집 데이터들의 식별성이 저하되는 현상이 자주 관찰된다. 이와 같은 한계를 보완한 모델로는 비선형 차원 축소 기법의 t-SNE(t-distributed stochastic neighbor embedding)가 있다. t-SNE에서는 임의의 데이터 포인트로부터 다른 데이터 포인트까지의 거리를 측정한 뒤 t분포의 값을 선택하여 해당 값을 친밀도로 정의하고 친밀도가 유사한 근접 데이터를 하나의 군집으로 할당한다. t-SNE는 다양한 실험 데이터 하에 비지도 학습 가운데 월등히 우월한 성능을 나타내는 모델임을 문헌을 통해 확인할 수 있다.¹⁷⁻¹⁹⁾ 일반적으로 데이터 분석 및 평가에 사용되는 알고리즘은 입력 값에 대한 결과 값을 사전에 정의된 카테고리 할당하는 지도 학습과 데이터에 유사성에 근거하여 입력 값 내에서 잠재된 패턴을 탐색하는 비지도 학습으로 구분된다.²⁰⁾ PCA와 t-SNE는 고차원의 원본 데이터를 비교적 저차원의 데이터로 변환하는 비지도 학습 기반 차원 축소 알고리즘에 해당된다. Fig. 1에 PCA와 t-SNE의 특징을 비교하여 시각화 하였다.

본 연구에서는 우리나라의 고유 전통적 문화유산인 한지의 열화특성을 분석하는 방법으로서 적외선 분광분석법과 PCA를 시도한 뒤 오분류 및 난분류 케이스에 대한

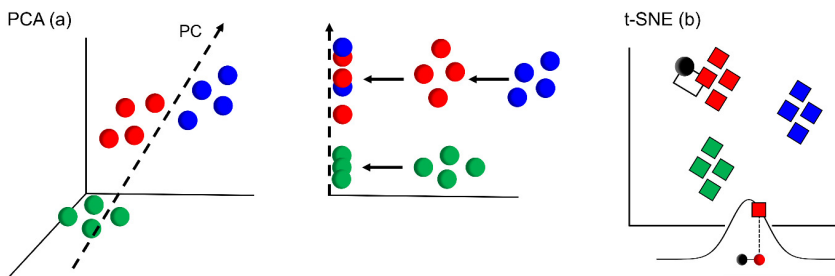


Fig. 1. Visualization of dimensionality reduction process of PCA (a) and t-SNE (b).

보완도구로서 t-SNE의 적용을 시도하였다. 열화특성 탐색을 위한 분석 범주는 건열 처리 및 습윤 가온 처리에 의한 가속열화로 구분하였으며 최종적으로 열화 방법과 기간에 따른 유효 화학적 작용기를 특성화 하여 보존과학 기술 대응을 위한 기초자료를 제공하고자 하였다.

2. 재료 및 방법

2.1 공시재료

열화특성 분석을 위한 공시 재료 정보와 습윤 가온 처리 및 건열 처리에 의한 가속 열화특성 분석 범주는 Table 1과 같다.

2.2 가속 열화

한지의 가속열화는 ISO 5630-1과 ISO 5630-3에 의거하여 각각 105℃에서의 건열 처리와 80℃ 및 65% 상대습도에서의 습윤 가온 처리로 구분하여 실행하였다. 가속열화는 각 조건별로 0일에서 최대 20일까지 진행하였으며 가속열화가 완료된 시료는 물리·화학적 분석에

앞서 ISO 187에 의거하여 23±1℃, 50%의 상대습도 조건에서 조습처리 하였다.

2.3 적외선 스펙트럼 데이터 측정

상기 가속열화 처리된 한지로부터 감쇠 전반사 적외선 분석기 ATR-IR(Alpha-P, Bruker Optics, Germany)를 이용하여 IR 스펙트럼 데이터를 수집하였다. 측정 영역은 4000-400 cm⁻¹ 파장범위에 대하여 4 cm⁻¹ 간격이었으며, 64회 반복 스캔 데이터의 평균치를 사용하였다. 시료 별 스펙트럼 데이터는 각 10회 반복 측정하여 총 130개의 IR 스펙트럼 데이터 세트를 구성하였다.

2.4 스펙트럼 데이터의 전처리

측정된 IR 스펙트럼 데이터를 주성분 분석에 앞서 Norris-Williams 알고리즘에 의거하여 1차 미분 처리하였다. Norris-Williams 알고리즘²¹⁾을 통한 1차 미분 시 파장영역 구간의 길이를 나타내는 segment 변수와 두 segment 간 거리를 나타내는 gap 변수는 Lee 등²²⁾의 연구를 참조하여 각각 7과 3으로 설정하였다.

비지도학습 기반 차원 축소 모델링을 위한 데이터 세트

Table 1. Summary of categories and classes type of accelerated aging

Category	Class	Origin	Cooking chemical	Mucilage	
Origin	K1 (n=10)	Korea	NaOH	<i>Hibiscus manihot</i> L.	
	K2 (n=10)		Ash		
	T0 (n=10)	Thai	NaOH	PAM	
Accelerated aging time (dry heat)	D0 (n=10)	Korea	NaOH	<i>Hibiscus manihot</i> L.	
	D3 (n=10)				
	D6 (n=10)				
	D10 (n=10)				
Accelerated aging time (moist heat)	D20 (n=10)	Korea	NaOH	<i>Hibiscus manihot</i> L.	
	M0 (n=10)				
	M3 (n=10)				
	M6 (n=10)				
Accelerated aging method	M10 (n=10)	Korea	NaOH	<i>Hibiscus manihot</i> L.	
	M20 (n=10)				
	KD1 (n=10)				NaOH
	KD2 (n=10)				Ash
Accelerated aging method	KM1 (n=10)	Korea	NaOH	<i>Hibiscus manihot</i> L.	
	KM2 (n=10)		Ash		

는 1800–1200 cm^{-1} 영역의 스펙트럼 데이터로 구성하였다. 본래 IR 스펙트럼의 측정 영역은 4000–400 cm^{-1} 이었으나 Hwang 등⁷⁾이 제시한 결과를 참조하여 입력 변수를 축소하였다. 과도한 입력변수의 설정은 모델의 복잡도를 증가시키고 효율성을 저해한다.^{2,3,6)} Kim 등⁹⁾에 의하여 1800–1200 cm^{-1} 영역은 종이의 화학적 특성에 있어 적합한 스펙트럼 범위임이 확인된 바 있다.

2.5 주성분 분석과 t-SNE

주성분 분석과 t-SNE는 R software(Ver. 4.3.1, Auckland, New Zealand)와 오픈소스 라이브러리를 활용하여 시행하였다. t-SNE 구현 시 군집 형성 범위에 대한 영향력을 반영하는 perplexity 변수는 초기 5–10으로 설정한 뒤 시각화 특성을 고려하여 최적화하였다.

2.6 기계적 강도 특성 분석

한지의 가속열화 시행에 따른 기계적 강도 특성 변화는 내절도 분석을 통해 평가하고자 하였으며 내절도는 TAPPI Standard T-511 om-20에 의거하여 MIT형 내절도 시험기(Model 3, Tinius Olsen, USA)로 측정하였다.

3. 결과 및 고찰

3.1 원산지에 따른 주성분 분석

Fig. 2는 Table 1에 제시한 원산지 범주에서 1800–1200 cm^{-1} 영역의 한지 IR 스펙트럼을 나타낸 것이다. Fig. 2(a)는 1800–1200 cm^{-1} 영역에서의 원본 IR 스펙트럼 데이터를 도시한 것이며 Fig. 2(b)는 Fig. 2(a)의 1차 미분 스펙트럼 데이터를 나타낸다. Fig. 2(a)로부터 1647–1635 cm^{-1} (water)²²⁾, 1422 cm^{-1} (CH₂ bending)²³⁾, 1370 cm^{-1} (CH₂ bending)²⁴⁾, 1335 cm^{-1} (amorphous cellulose)⁷⁾, 1315 cm^{-1} (OH bending)²⁵⁾, 1265 cm^{-1} (C–O stretching)²⁶⁾ 및 1205 cm^{-1} (C1–O–C4' symmetric or OH bending)²⁷⁾이 검출된 것을 확인하였다. Fig. 2(b)의 1차 미분 스펙트럼의 경우 1500 cm^{-1} 영역 이후 증폭된 노이즈가 지배적인 경향을 나타냈다. 스펙트럼 데이터의 평활화를 위해 선행연구²⁾를 참조하여 1차 미분을 위한 segment와 gap의 조절을 시도하였으나 노이즈의 완벽한 정제는 이뤄지지 않는 것으로 분석되었다. Fig. 2(b)의 1차 미분 스펙트럼 데이터로부터 Fig. 2(a)에서 검출한 피크 이외에 1510 cm^{-1} (aromatic ring)²⁷⁾, 1462 cm^{-1} (CH bending)²⁴⁾ 및 1355 cm^{-1} (OH bending)²⁸⁾의 특징 피크를 추가로 확인하였다.

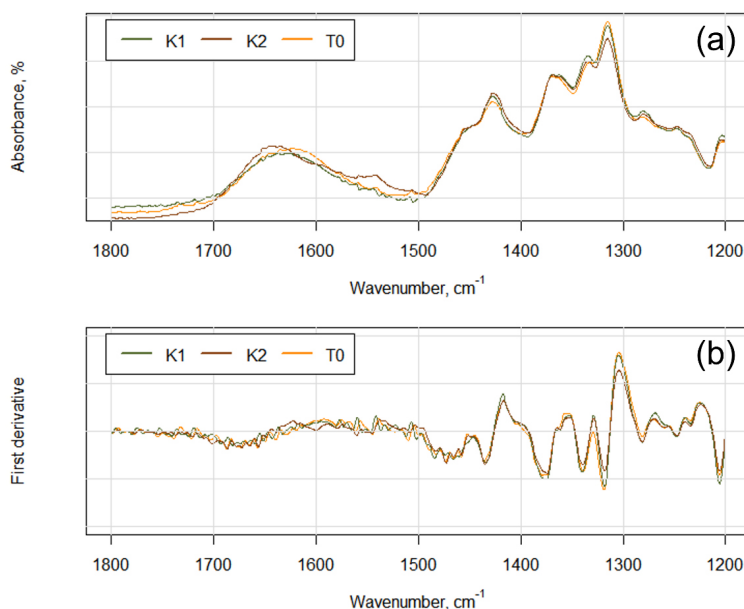


Fig. 2. Raw IR spectra (a) and first derivative IR spectra (b) in the spectral range of 1800–1200 cm^{-1} .

Fig. 3은 Fig. 2의 1800–1200 cm^{-1} 영역 IR 스펙트럼 데이터를 통해 시행한 주성분 분석 결과를 나타내는 PCA score plot을 도시한 것이다. Fig. 3(a)와 Fig. 3(b)는 각각 Fig. 2(a)의 원본 스펙트럼 데이터와 Fig. 2(b)의 1차 미분 스펙트럼으로부터 실시한 주성분 분석 결과를 나타낸다. Lee 등²⁾이 보고한 바와 같이 Norris-Williams 알고리즘에 의거한 스펙트럼 데이터의 1차 미분 처리는 주성분 분석을 통한 시료의 특성화에 있어 효과적 접근임을 Fig. 3으로부터 확인할 수 있었다. 스펙

트럼 데이터의 1차 미분처리 이후 시행한 주성분 분석 결과에서 전처리 이전의 결과보다 각 클래스의 주성분 데이터 포인트의 군집이 보다 명확하게 분류되는 것으로 나타났다. Fig. 3(b)에서 T0를 기준으로 시료 별 특성화에 있어 K2의 경우 PC1, K1의 경우 PC2에 의해 구분되어지는 것을 확인할 수 있다.

Fig. 4는 1800–1200 cm^{-1} 영역에서의 1차 미분 IR 스펙트럼과 Fig. 3(b)에 작성된 PCA score plot에 대한 구동력을 설명하는 PC 로딩 값을 제시한 것이다. Fig.

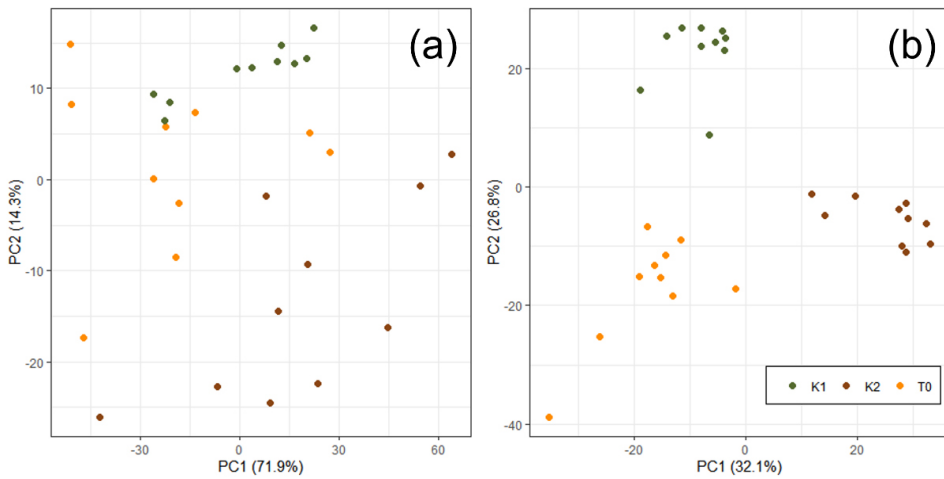


Fig. 3. PCA score plots generated from raw (a) and first derivative (b) IR spectra in the spectral range of 1800–1200 cm^{-1} .

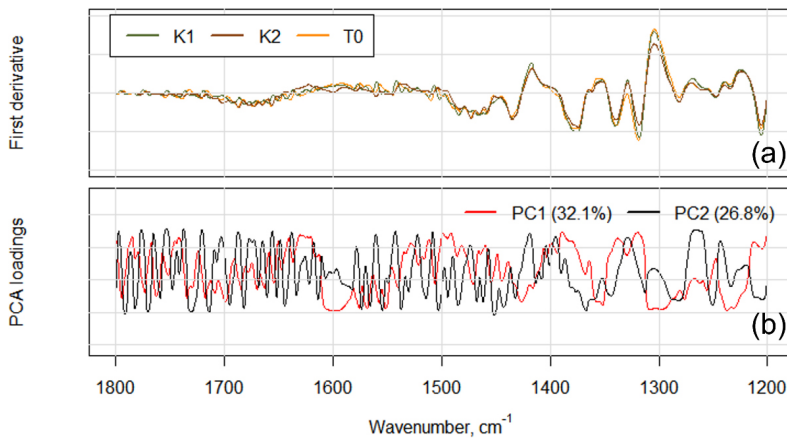


Fig. 4. First derivative (a) IR spectra in the spectral range of 1800–1200 cm^{-1} , along with the loading values of the first two principal components for all data points in the first derivative spectra (b) for origin category.

4(b)의 PC 로딩 데이터로부터 각 PC에 대한 영향력 정보를 도출할 수 있다. 다수의 지배적 노이즈가 작용한 1500 cm^{-1} 영역 이후의 작용기를 제외하면 T0로부터 K2를 PC1의 양(+)의 영역으로 이동시키는 구동력을 대별하는 특징 피크는 1462 cm^{-1} , 1315 cm^{-1} 및 1205 cm^{-1} 흡수대인 것으로 분석되었다. 또한 T0로부터 K1을 PC2의 양(+)의 방향으로 위치시키는 유효 피크는 1422 cm^{-1} , 1265 cm^{-1} 흡수대로부터 발생하는 것으로 확인되었다. 상기 제시된 특징 영역은 대부분 셀룰로오스, 헤미셀룰로오스 및 수분과 연관된 흡수대로 제조 지역 차이로 유래된 원료 자체의 이질성과 증해 공정에 따른 화학적 특성 변화에 기인하여 각 주성분에 해당 유효 피크가 할당되었을 것으로 사료된다.

3.2 가속열화 기간에 따른 주성분 분석

Fig. 5는 Table 1의 가속열화 시간 범주에 속하는 클래스의 1차 미분 IR 스펙트럼 데이터를 나타낸 것이다. 시료명으로 기재된 영문 D(dry heat)와 M(moist heat)은 가속열화 방법을 대표하기 위한 것이며 숫자의 경우 가속열화 시행 일(days)을 나타낸다. Fig. 5(a)와 Fig. 5(b)

는 각각 건열 처리 및 습윤 가온 처리에 대한 분석 범주의 1차 미분 스펙트럼을 제시한 것이다. Fig. 5에 표기한 수평 방향 중심의 음영 보조선은 1차 미분 계수 0을 나타내는 지점이다. 원본 데이터에 대한 극대 혹은 극소는 1차 미분 계수를 0으로 할당한다. 따라서 가속열화 처리에 따라 원본 스펙트럼으로부터 도출된 열화 특징 피크의 경우 1차 미분 스펙트럼 상에서 0에 근접한 수치를 기록할 가능성이 높은 것으로 해석할 수 있다.

앞서 원산지 범주에서 특징 스펙트럼 영역을 제시한 바와 같이 가속열화 기간에 따른 유효 피크를 도출하려고 시도하였으나 원본 스펙트럼과 1차 미분 스펙트럼 상에서 모두 특징 피크의 육안적 식별이 어려웠다. 따라서 가속열화 기간에 따른 특성화 피크는 주성분 분석 결과를 참고하여 추출하고자 하였다.

Fig. 6은 Fig. 5의 스펙트럼 데이터를 통해 시행한 주성분 분석 결과를 나타내는 PCA score plot을 도시한 것이다. Fig. 6(a)와 Fig. 6(b)는 각각 건열 처리 및 습윤 가온 처리 범주에 대한 주성분 분석 결과를 나타낸다. Fig. 6(a)에서 건열 처리에 의한 열화의 경우 가속열화 처리 시간이 증가함에 따라 데이터 포인트가 PC1 축을

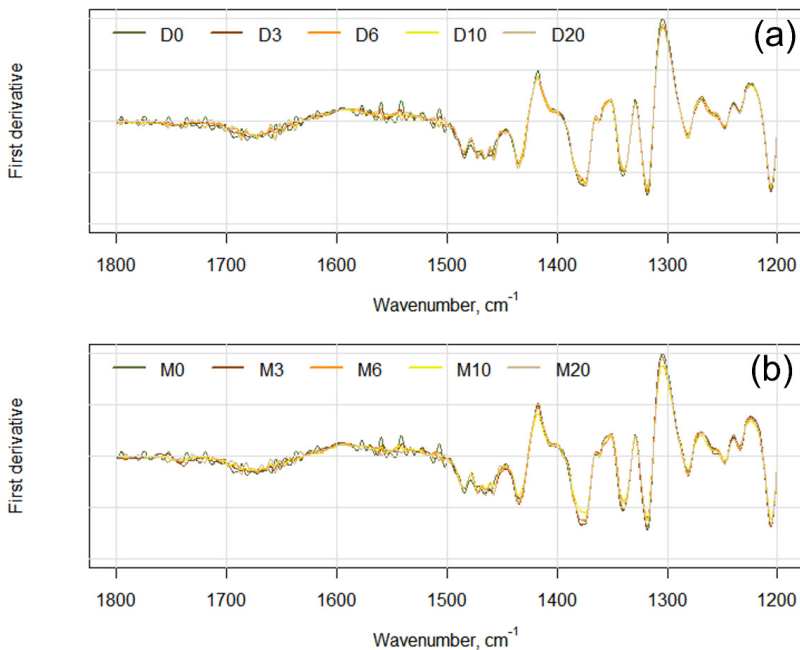


Fig. 5. First derivative IR spectra of accelerated aged Hanji in the spectral range of 1800–1200 cm^{-1} (a: category of aging time for dry heat, b: for moist heat).

기준으로 양(+)¹⁵의 방향으로 이동하는 것으로 관측되었다. 이는 PC 로딩 값으로부터 PC1을 양(+)¹⁵의 영역으로 할당하는 구동력과 연관된 흡수대를 탐색한다면 건열 처리에 의한 한지의 열화 거동을 설명할 수 있는 화학적 작용기를 특정화할 수 있음을 시사한다. 한편 Fig. 6(b)의 습윤 가온 처리 범주에 대하여 작성된 PCA score plot에서는 데이터 포인트가 열화 진행 시간이 경과함에 따라 PC2 축의 음(-) 영역으로 배치되는 것으로 확인되었다. 따라서 PC 로딩 값에 의거하여 PC2의 음(-)의 방향에 대한 구동력을 할당하는 흡수대를 추출한다면 습윤 가온 열화에 따른 유효 화학적 작용기를 설명할 수 있을 것이라 판단된다.

Fig. 7은 Fig. 5에 도시한 가속열화 기간 범주에 속하는 클래스의 1차 미분 IR 스펙트럼 데이터와 이에 대한 PCA score plot의 PC로딩 값을 비교한 결과를 나타낸 것이다. Fig. 7(a)와 Fig. 7(c)의 건열 처리 범주에 대한 스펙트럼 데이터 및 PC로딩 값 비교 결과 Fig. 6(a)의 데이터 포인트를 PC1의 양(+)¹⁵의 영역으로 배치하는 구동력은 1768 cm⁻¹(C=O stretching)²⁹⁾, 1462 cm⁻¹, 1335 cm⁻¹, 1250 cm⁻¹(C=O stretching)¹⁵⁾ 및 1205 cm⁻¹의 흡수대인 것으로 분석되었다. 또한 Fig. 7(b)와 Fig. 7(d)의 습윤 가온 처리 범주의 스펙트럼 데이터와 PC로딩 값 비교에서는 PC2 축 상에 음(-)의 영역 배치에 대한 구동력이 1768 cm⁻¹, 1640 cm⁻¹, 1552 cm⁻¹(COO⁻ stretching)²⁸⁾, 1462 cm⁻¹ 및 1370 cm⁻¹ 흡수대

로 할당되는 것으로 나타났다.

Fig. 7의 PC 로딩 값의 비교로부터 건열 및 습윤 가온 처리에 의한 열화와 관련된 화학적 특징화 피크로서 대부분 카르보닐기와 카르복실기 혹은 C-H bending 및 수분에 기인한 흡수대가 검출되는 것으로 확인되었다. 종이의 열화는 가수분해와 산화가 동시 발생됨을 통해 진행되며 이들은 서로 촉매와 같이 작용한다. 가수분해는 알데하이드 말단기의 변형을 통해 진행되며 산화에 의해 형성된 카르복실기는 가수분해를 촉진하고 이후 다시 산화를 위한 환원제를 제공한다.²⁹⁾ 카르보닐기와 카르복실기는 종이에 열화가 진행됨에 따라 셀룰로오스에 생성되는 작용기로서 C₂, C₃ 및 C₆에 주로 생성되며 경우에 따라 C₁에 나타나기도 한다. C₁에 산화가 발생되면 -COOH가 생성되며 C₂와 C₃의 경우 환의 분열(ring cleavage) 유무에 따라 C=O와 -COOH가 형성된다. C₆의 산화 과정은 수산기가 공격받아 알데하이드기가 생성되고 생성된 알데하이드기가 다시 산화되어 -COOH를 형성함으로써 진행된다.^{9,28,29)} C-H bending의 경우 셀룰로오스의 결정화도와 관련된 작용기로 열화가 진행됨에 따라 셀룰로오스 결정화도가 감소하면서 특징 흡수대로 검출된 것으로 판단된다.¹⁵⁾ 건열 처리 범주 내 PC2 축 상의 배치를 설명하였던 1640 cm⁻¹(OH bending)의 경우 건열에 의한 가속열화가 시료의 함수율에 영향하여 특정된 피크라 사료된다.

Fig. 8은 건열 및 습윤 가온 처리 및 열화 시간 경과에

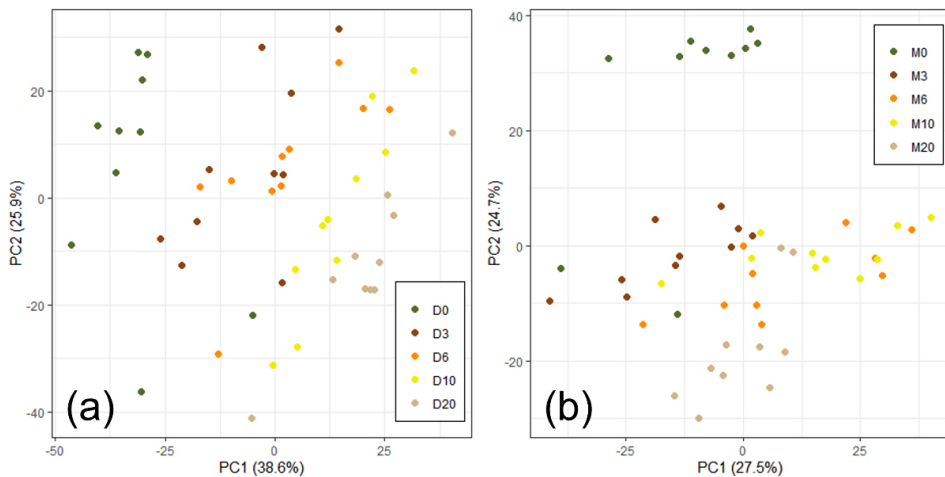


Fig. 6. PCA score plots generated from first derivative IR spectra of dry heated (a) and moist heated (b) Hanji in the spectral range of 1800–1200 cm⁻¹.

따른 내절도 및 열화지수(aging index, AI) 변화를 나타낸 것이다. AI는 Kim 등³⁰⁾의 방법을 검토하여 Fig. 7에서 건열 및 습윤 가온 열화 특징 피크에 모두 해당하였던 카르보닐기(1768 cm⁻¹, C=O)와 메틸렌기(2900 cm⁻¹, -CH₂)의 피크 intensity 비율로 계산하였다. 건열 처리에 의해 열화된 종이의 경우 1740-1710 cm⁻¹에서의 카르보닐기 관련 특징 피크가 검출되지 않을 수 있다는 Hajji 등³¹⁾의 연구결과를 검토하여 건열 및 습윤 가온 열화 특성을 동시에 반영할 수 있는 새로운 지표가 필요할 것이라 판단하여 AI의 산출을 시도하였다. C-H bending에 의거한 특성화 피크가 다수 검출되고 있으나 다중의 신호가 겹쳐 형성되는 IR 스펙트럼 분석 능의 한계로 인하여 본래 셀룰로오스 구조 내에서 결정질과 비결정질 상에 열화로부터 발생한 결정 변화의 구분 및 정량화가 불가능하다고 판단하였기에 결정화 지수(total crystallinity index, TCI)에 대한 정의는 시도하지 않았다.

Fig. 8(a)의 내절도 결과의 경우 Walkden³²⁾이 보고한 바와 일치하게 열화가 진행됨에 따라 강도가 감소하는 경향을 나타냈다. 건열 처리에 의한 열화 한지의 경우 6일(144 hr)의 가속 열화처리까지 강도가 급격히 감소하다가 6일 이후 감소폭이 다소 줄어드는 경향을 나타냈다. 한편 습윤 가온 처리에 의한 가속 열화 한지의 경우 3일(72 hr) 경과 이후 강도 감소 폭이 상대적으로 미미하였다. 이러한 경향은 AI 결과에서도 유사하게 드러났다. Fig. 8(b)에 AI 결과에서 건열 처리 한지의 경우 3에서

6일의 가속 열화 진행 기간 까지 AI가 급속히 증가한 뒤 6일 경과 이후 다시 감소 및 유지되는 경향을 나타냈으며 습윤 가온 처리의 경우 열화 진행 3일까지 AI가 급격히 증가한 후 서서히 감소하는 경향을 나타냈다. 이는 가속 열화에 의한 기계적 물성 저하가 주로 진행되는 기간을 추측할 수 있는 결과라고 사료된다. 각 데이터 포인트의 편차는 수작업으로 제조되는 한지의 제조 과정에 기인한 결과로 추정된다.

Fig. 9는 열화 원인 별 열화 진행도에 따른 분류 분석 가능성을 확인하기 위하여 시행한 t-SNE 결과를 도시한 것이다. t-SNE는 독립 변수 상호 간 유사도를 조건부 확률로 측정하여 상대적으로 가까운 클래스에 높은 확률을 부여하고 반대로 먼 클래스에 0에 근접한 확률을 부여하여 원본 데이터로부터 확률 분포를 계산하여 차원을 축소한다. t-SNE의 경우 선형 변환을 이용하는 PCA와 다르게 비선형 특성을 갖는 데이터의 특성을 추출함에 있어 유리하다.¹⁷⁾ Fig. 6과 Fig. 9를 비교하였을 때 PCA score plot에 비하여 t-SNE plot이 월등히 우월한 분류 성능과 시각화 특성을 나타내고 있는 것을 확인할 수 있다. 그러나 t-SNE의 경우 PCA와 비교하였을 때 계산 비용(computational cost)이 높고 분류규칙에 대한 논리적 설명을 제공하지 않는다. 따라서 특정 모델이 각 분석 도구의 효용성을 완전히 대체할 수 있는 것은 아니며 모델의 선택에 있어 데이터의 규모 및 과적합 등을 종합적으로 고려한 상호 보완적 검토가 요구된다.^{2,3)}

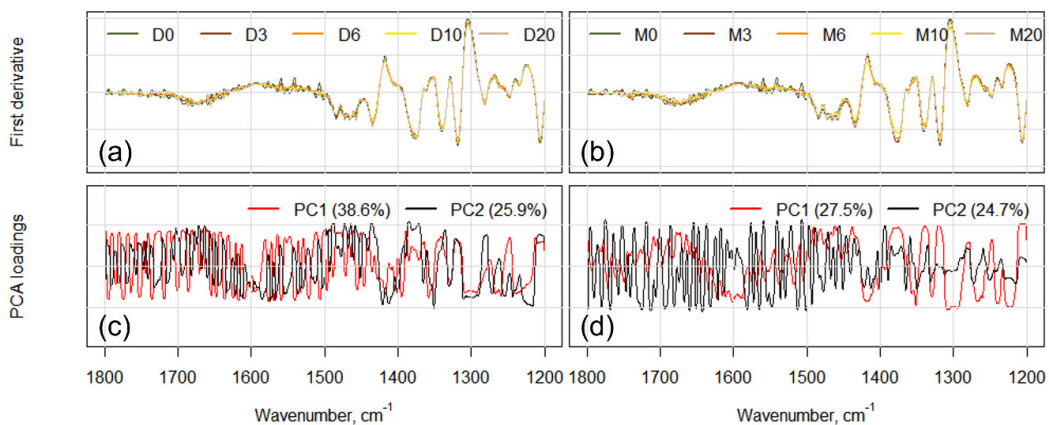


Fig. 7. First derivative IR spectra of dry heated (a) and moist heated (b) Hanji, along with the loading values of the first two principal components for all data points of dry heat (c) and moist heat (d) category.

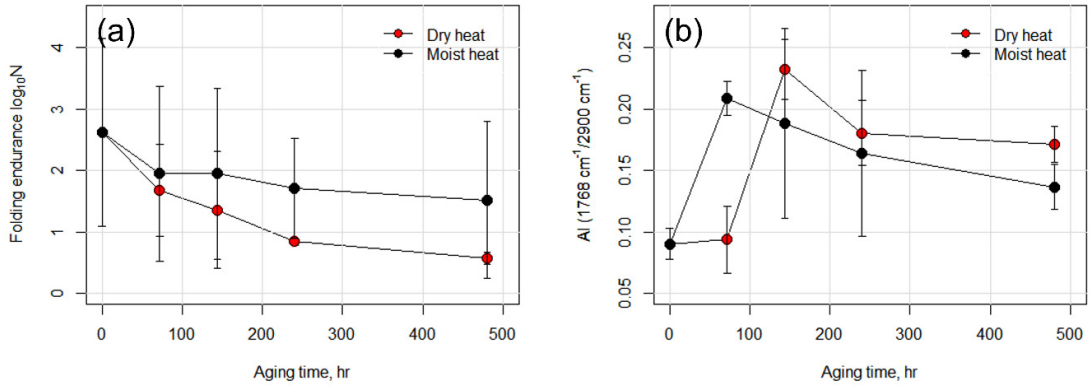


Fig. 8. Folding endurance (a) and oxidation index (b) with respect to aging time.

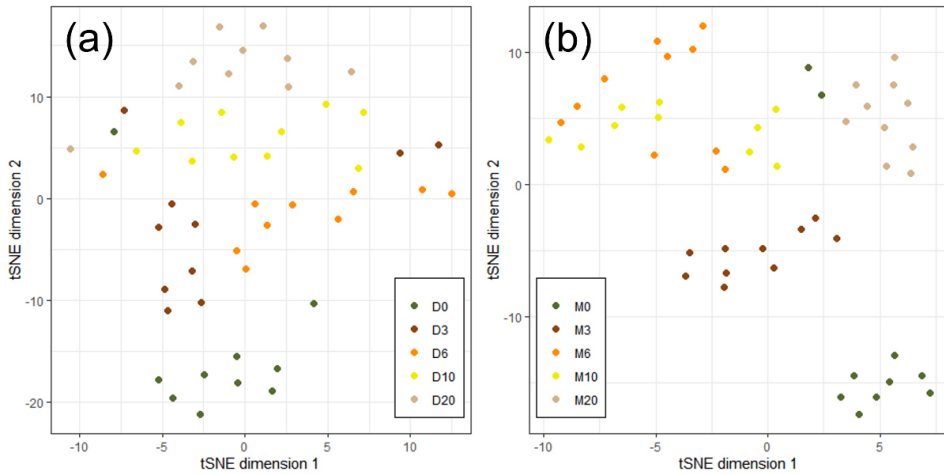


Fig. 9. T-SNE plots generated from first derivative IR spectra of dry heated (a) and moist heated (b) Hanji in the spectral range of 1800–1200 cm⁻¹.

3.3 가속 열화 방법에 따른 주성분 분석

Fig. 10(a)는 Table 1에 나타냈던 가속 열화 방법 범주에 대한 주성분 분석 결과를 도시한 것이다. 가속 열화 방법 범주에서의 각 클래스는 K1 및 K2 시료에 대하여 20일의 건열 및 습윤 가온 열화 처리한 시료로 구성하였다. Fig. 10(a)로부터 열화 처리 방법에 기인하여 각 클러스터가 PC1 축을 따라 양(+)과 음(-)의 방향으로 분류되는 것을 확인할 수 있다. 따라서 PC 로딩 값에 의거하여 PC1의 양(+)과 음(-)의 방향에 대한 구동력을 할당하는 흡수대를 추출한다면 가속 열화 방법에 따른 화학적 작용기의 변화를 특정화 할 수 있을 것이라 판단된다.

Fig. 10(b)는 Fig. 10(a)에서와 같이 PCA를 우선 진행한 뒤 t-SNE를 시행한 결과를 나타낸 것이다. 앞서 언급하였듯이 t-SNE의 경우 계산 비용이 높아 결과 출력에 있어 상대적으로 많은 시간이 요구될 수 있다. 그러나 PCA를 우선 실행함으로써 1차적으로 차원을 축소시킨 뒤 t-SNE를 시행하면 계산 비용 절감과 시각화 특성을 개선할 수 있다.¹⁷⁻¹⁹⁾ Fig. 10(b)의 t-SNE plot은 이상의 PCA score plot들과 Fig. 9에서 확인하였던 t-SNE 결과보다 우월한 분류 성능과 시각화 특성을 나타내고 있는 것을 확인할 수 있다. 따라서 PCA와 t-SNE의 결합은 한지의 열화 요인 추적을 위한 분류 분석의 도구로

서 적용이 가능할 것이라 사료된다.

Fig. 11은 Fig. 10(a)에 도시한 가속열화 방법 범주에 속하는 클래스의 1차 미분 IR 스펙트럼 데이터와 이에 대한 PCA score plot의 PC로딩 값을 비교한 결과를 나타낸 것이다. Fig. 11(b)로부터 PC1 축을 기준으로 데이터 포인트를 양(+)의 영역에 할당하는 흡수대가 1768 cm^{-1} , 1640 cm^{-1} , 1620 cm^{-1} (C=O stretching), 1462 cm^{-1}

및 1370 cm^{-1} 임을 확인하였다. 또한 음(-)의 영역에 대한 구동력과 관련된 피크는 1552 cm^{-1} , 1422 cm^{-1} , 1355 cm^{-1} , 1337 cm^{-1} (C-OH bending)³³ 및 1315 cm^{-1} (CH₂ wagging)³⁴인 것으로 분석되었다. 추가적으로 1740–1710 cm^{-1} (C=O stretching)의 흡수대에서 양(+)과 음(-)의 영역에 대한 구동력이 변동함은 건열에 의한 열화의 경우 해당 흡수대에 산화 관련 작용기가 할당되지 않음에

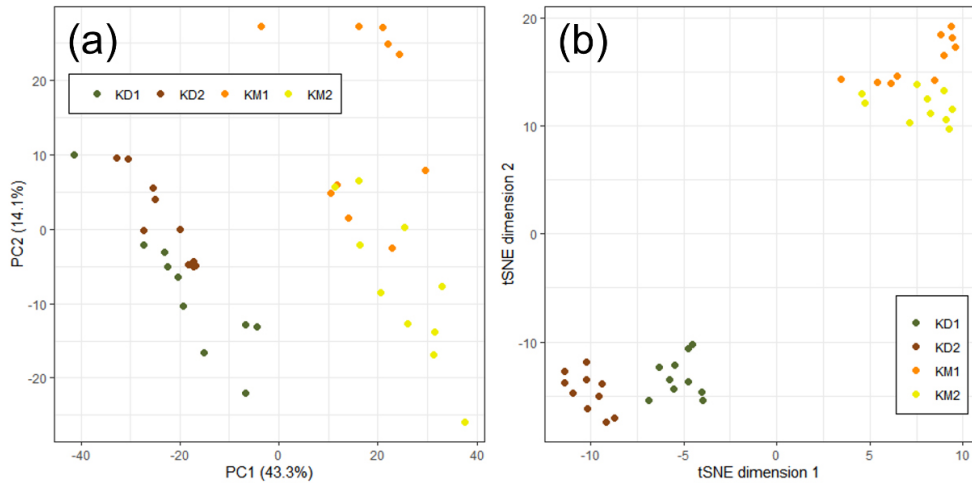


Fig. 10. PCA score (a) and t-SNE (b) plots generated from first derivative IR spectra of aging method category.

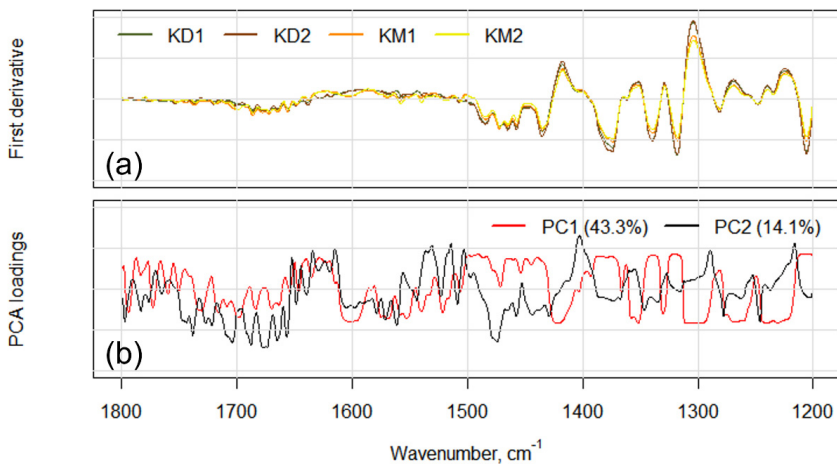


Fig. 11. First derivative (a) IR spectra in the spectral range of 1800–1200 cm^{-1} , along with the loading values of the first two principal components for all data points in the first derivative spectra (b) for aging method category.

기인한 결과로 사료된다.³¹⁾

앞서 카르보닐기와 카르복실기 혹은 C-H bending 및 수분과 관련된 흡수대가 열화에 의한 특성화 피크임을 확인하였으며 본 범주에서도 이에 대응하는 분석결과가 도출되었다.

4. 결론

본 연구에서는 우리나라의 고유 전통적 문화유산인 한지의 열화특성을 분석하는 방법으로서 비지도 학습 기반 차원 축소 모델의 일종인 PCA와 t-SNE를 적용하였으며 각각 원산지, 열화 기간 및 열화 방법에 따른 범주로 구분하여 분류 분석을 진행하였다. 원산지 범주의 경우 제조 지역 차이로 유래된 원료 자체의 이질성과 증해 공정에 따른 화학적 특성 변화에 기인하여 셀룰로오스, 헤미셀룰로오스와 연관된 IR 피크가 데이터 클래스에 대별성을 부여함에 따라 분류 분석이 가능한 것으로 확인되었다. 열화 기간에 대한 범주 내에서는 PCA 시행 결과를 바탕으로 카르보닐기와 카르복실기 혹은 C-H bending 및 수분에 기인한 흡수대가 데이터 클래스에 차별점을 나타내는 특징 피크임을 확인할 수 있었다. 또한 건열 및 습식 가온 처리 의한 열화 방법 양측에서 동시에 특징 피크로 검출되었던 카르보닐기(1768 cm^{-1})의 피크 intensity를 통해 열화 지수 산출을 시도하였으며 기계적 특성 데이터와 일부 부합함을 확인하였다. 그러나 IR 스펙트럼의 다중공선성을 고려하였을 때 신뢰도에 대한 검토가 요구된다고 판단된다. 가속열화 방법 범주에서 역시 카르보닐기와 카르복실기 혹은 C-H bending을 중심으로 한 특성화 피크를 통해 성공적으로 열화 요인 별 분류 분석이 가능한 것으로 분석되었다. PC 로딩 결과로부터 건열 및 습윤 가온 열화로부터 기인되는 화학적 작용기를 구분하여 구체화하고자 하였으나 가수분해와 산화 그리고 결정화도의 변화로부터 할당되는 공통적인 스펙트럼 영역 이외에 열화 요인 별 특성화 피크 도출은 어려웠으며 향후 추가적인 연구가 진행되어야 할 것이라 판단된다. 이상의 결과로부터 한지의 열화 특성을 추적 및 분류하는 방법으로서 PCA와 t-SNE의 적용이 합리적임을 확인하였으며 본 분석 프로세스는 보존 과학 분야 전반에 걸쳐 다양한 범주에 있어 응용이 가능하다고 사료된다.

사사

본 연구는 정부(과학기술정보통신부)의 재원으로 한국연구재단(RS-2023-00301889)의 지원에 의하여 이루어진 것입니다.

Literature Cited

1. Kim, K. J. and Eom, T. J., Classification of papers using IR and NIR spectra and principal component analysis, *Journal of Korea TAPPI* 48(1):34-42 (2016).
2. Lee, Y. J., Kweon, S. W., Kim, J. H., Cha, J. E., and Kim, H. J., Spectral preprocessing and machine learning modeling for discriminating manufacturing origins of mulberry bast fiber, *Journal of Korea TAPPI* 55(5):61-74 (2023).
3. Lee, Y. J., Lee, J. E., Gwon, J. G., Lee, T. J., and Kim, H. J., Predictive modeling for degree of substitution of cellulose acetate using infrared spectroscopy and machine learning, *Journal of Korea TAPPI* 55(5):83-95 (2023).
4. Lee, Y. J., Kim, J. H., and Kim, K. J., Application of infrared spectroscopy to analyze the degree of substitution in acetylated wood powder, *Journal of Korea TAPPI* 54(4):19-24 (2022).
5. Lee, Y. J., Cha, J. E., Kang, N. Y., Kim, J. H., and Kim, H. J., Application of spectroscopy to analyze the degree of substitution of carboxymethyl cellulose, *Journal of Korea TAPPI* 55(1):103-111 (2023).
6. Lee, Y. J., Lee, T. J., and Kim, H. J. (2024). Classification analysis of copy papers using infrared spectroscopy and machine learning modeling, *BioResources* 19(1):160-182 (2024).
7. Hwang, S. W., Park, G. Y., Kim, J. H., and Jeong, M. J., Predictive modeling of traditional korean paper characteristics using machine learning approaches (part 1): discriminating manufacturing origins with artificial

- neural networks and infrared spectroscopy, *Journal of Korea TAPPI* 55(4):57–69 (2023).
8. Kim, K. J., Kim, J. H., Park, G. Y., and Jeong, M. J., Predictive modeling of Korean traditional paper characteristics using machine learning approaches (part2): prediction of carbonyl content and analysis of variable importance using random forest, *Journal of Korea TAPPI* 55(5):13–23 (2023).
 9. Kim, K. J. and Eom, T. J., Study on the aging characteristics of paper with principal component analysis, *Journal of Korea TAPPI* 48(6):144–149 (2016).
 10. ISO 5630–1, Paper and board –Accelerated aging– Part 1: Dry heat treatment at 105 °C.
 11. ISO 5630–3, Paper and board –Accelerated aging– Part 3: Moist heat treatment at 80 degrees C and 65% relative humidity.
 12. Nelson, M. L. and O’Connor, R. T., Relation of certain infrared bands to cellulose crystallinity and crystal lattice type, Part II: A new infrared ratio for estimation of crystallinity in celluloses I and II, *Journal of Applied Polymer Science* 8(3):1325–1341 (1964).
 13. Yousef, M. A., Shabaka, A. A., Nada, A. M. A., and Abdei–Nour, K. N., Physical properties and infrared spectroscopy of thermally treated bagasse paper sheets, *Indian Journal of Pure and Applied Physics* 29(1):6–8 (1991).
 14. Tasker, S., Badyal, J. P. S., Backson, S. C. E., and Richards, R. W., Hydroxyl accessibility in celluloses, *Polymer* 35(22):4717–4721 (1994).
 15. Kim, K. J., Ahn, E. B., Ryu, J. A., and Eom, T. J., Identification of pulp fibers in the chronological academic journals using spectroscopic method, *Journal of Korea TAPPI* 50(4):61–70 (2018).
 16. Mattias S., Approaches to analyse and interpret biological profile data, Ph. D. thesis, University of Potsdam, Germany (2006).
 17. Van Der Matten, L. and Hinton, G., Visualizing data using t–SNE, *Journal of Machine Learning Research* 9(2605):2579–2605 (2008).
 18. Zhou, Q., Huang, W., Fan, S., Zhao, F., Liang, D., and Tian, X., Non–destructive discrimination of the variety of sweet maize seeds based on hyperspectral image coupled with wavelength selection algorithm, *Infrared Physics & Technology* 109:103418 (2020).
 19. Grabowski, B., Masarczyk, W., Głomb, P., and Mendys, A., Automatic pigment identification from hyperspectral data, *Journal of Cultural Heritage* 31:1–12 (2018).
 20. Lee, Y. J., A Study on the quality characterization of printing and writing paper using surface parameters and data mining techniques, Master thesis, Kookmin University, Korea (2022).
 21. Norris, K. and Williams, P., Optimization of mathematical treatments of raw nearinfrared signal in the measurement of protein in hard red spring wheat. I. influence of particle size, *Cereal Chem* 61(2):158–165 (1984).
 22. Polovka, M., Polovková, J., Vizárová, K., Kirschnerová, S., Bielíková, L., and Vrška, M., The application of FTIR spectroscopy on characterization of paper samples, modified by Bookkeeper process, *Vibrational Spectroscopy* 41(1):112–117 (2006).
 23. Ciolacu, D., Kovac, J., and Kokol, V., The effect of the cellulose–binding domain from *Clostridium cellulovorans* on the supramolecular structure of cellulose fibers, *Carbohydrate Research* 345(5):621–630 (2010).
 24. Horikawa, Y., Hirano, S., Mihashi, A., Kobayashi, Y., Zhai, S., and Sugiyama, J., Prediction of lignin contents from infrared spectroscopy: chemical digestion and lignin/biomass ratios of *Cryptomeria japonica*, *Applied Biochemistry and Biotechnology* 188:1066–1076 (2019).
 25. Delmotte, L., Ganne–Chedeville, C., Leban, J. M., Pizzi, A., and Pichelin, F., CP–MAS 13C NMR and FT–IR investigation of the degradation reactions of polymer constituents in wood welding, *Polymer degradation and stability* 93(2):406–412 (2008).

26. Shimizu, S., Yokoyama, T., and Matsumoto, Y., Effect of type of aromatic nucleus in lignin on the rate of the β -O-4 bond cleavage during alkaline pulping process, *Journal of wood science* 61:529-536 (2015).
27. Pandey, K. K., A study of chemical structure of soft and hardwood and wood polymers by FTIR spectroscopy, *Journal of applied polymer science* 71(12):1969-1975 (1999).
28. Fengel, D., Influence of water on the OH valency range in deconvoluted FTIR spectra of cellulose, *Holzforschung* 47(2):103-108 (1993).
29. Dupont, A. L., Degradation of cellulose at the wet/dry interface-II - An approach to the identification of the oxidation compounds, *Restaurator* 17:145-164 (1996).
30. Kim, K. J. and Eom, T. J., Aging behavior of beeswax coated Hanji (I) - Thermal aging test of beeswax coated Hanji, *Journal of Korea TAPPI* 42(2):46-52 (2010).
31. Hajji, L., Boukir, A., Assouik, J., Pessanha, S., Figueirinhas, J. L., and Carvalho, M. L., Artificial aging paper to assess long-term effects of conservative treatment. Monitoring by infrared spectroscopy (ATR-FTIR), X-ray diffraction (XRD), and energy dispersive X-ray fluorescence (EDXRF), *Microchemical Journal* 124:646-656 (2016).
32. Walkden, S. A., Permanence and durability of paper, TAPPI Press, Atlanta, Ga, pp. 81-84 (1988).
33. Colom, X. and Carrillo, F., Crystallinity changes in lyocell and viscose-type fibres by caustic treatment, *European Polymer Journal* 38(11): 2225-2230 (2002).
34. Boukir, A., Hajji, L., and Zghari, B., Effect of moist and dry heat weathering conditions on cellulose degradation of historical manuscripts exposed to accelerated aging: ^{13}C NMR and FTIR spectroscopy as a non-invasive monitoring approach, *Journal of Materials and Environment Science* 9:641-654 (2018).