FORENSIC DISCRIMINATION OF NATURALLY AND ARTIFICIALLY AGED PAPER USING ATR-FTIR SPECTROSCOPY AND CHEMOMETRIC TECHNIQUES

CHIONG SU PENG

SCHOOL OF HEALTH SCIENCES

UNIVERSITI SAINS MALAYSIA

2024

FORENSIC DISCRIMINATION OF NATURALLY AND ARTIFICIALLY AGED PAPER USING ATR-FTIR SPECTROSCOPY AND CHEMOMETRIC TECHNIQUES

by

CHIONG SU PENG

Dissertation submitted in partial fulfillment of the requirement for the Bachelor of Science in Forensic Science with Honours

February 2024

CERTIFICATE

This is to certify that the dissertation entitled "Forensic Discrimination of Naturally and Artificially Aged Paper using ATR-FTIR Spectroscopy and Chemometric Techniques" is the bonafide record of research work done by Chiong Su Peng, Matric Number 160065 during the period from **October 2023 to February 2025** under my supervision. I have read this dissertation and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation to be submitted in partial fulfillment for the degree of Bachelor of Science in Forensic Science with Honours.

Supervisor,

…………………

(ChM Dr. DZULKIFLEE ISMAIL)

Lecturer

School of Health Sciences

Universiti Sains Malaysia

Health Campus

16150 Kubang Kerian

Kelantan, Malaysia

Date: 5/2/2025

DECLARATION

I hereby declare that this dissertation is the result of my investigations, except where otherwise stated and duly acknowledged. I also declare that it has not been previously concurrently submitted as a whole for any other degrees at Universiti Sains Malaysia or other institutions. I grant Universiti Sains Malaysia the right to use the dissertation for teaching, research, and promotional purposes.

. ….

(CHIONG SU PENG)

Date:5/2/2025

ACKNOWLEDGEMENTS

First and foremost, I would like to express my heartfelt gratitude to my supervisor, Dr. Dzulkiflee Ismail, for the thoughtful comments and constant support throughout this research journey. I would equally want to express profound appreciation to my co-supervisor, Dr. Wan Nur Syuhaila Mat Desa, whose meticulous guidance and strategic recommendation have quite enhanced this project. They are not only great advisers but also a supportive couple whose encouragement has been vital throughout my research work. Their combined expertise, motivation, and patience have enabled me to go through all the processes necessary to produce this piece of work.

Moreover, I owe an additional debt of gratitude to the staff members of the Analytical Laboratory and Forensic Laboratory who provided the needed help. I am also thankful to May, a PhD student of Dr. Dzulkiflee, for her help with FTIR. Their professionalism, attention to my practical work with the equipment, and readiness to share knowledge were very important for the successful accomplishment of my practical work. They made my stay in the laboratory both productive and educational, for which I am deeply grateful.

This message is for my family, I am grateful for all the love, support, and faith you had in me as no amount of words is ever going to be enough to thank you all for making me what I am today. You have always been there for me as my strength through every obstacle which helped me gain emotional and motivational power to keep going. No encouragement would have been possible without you, and this success could not have been achieved without you.

It is also more than any other feeling that I appreciate the presence of my friends and classmates and my housemates “101” Ethel Lim, Koh Pei Yee, Nayomi Khoo, and

Seow Zhi Lin, who have been with me throughout the project. With your faithful support and camaraderie, it is always easy to get inspired. Through late-night study sessions, shared jokes, and cheering each other up at the most depressing moments, you make this journey an ordinary yet extraordinary one.

To all of you who have contributed to this journey, whether directly or indirectly, I extend my heartfelt thanks. Your support has been a beacon of hope and motivation, and I am forever grateful for your part in the successful completion of this thesis.

Thank you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS AND SYMBOLS

ATR	Attenuated Total Reflectance
C-C	Carbon-Carbon single bond
C-Cl	Carbon-Chlorine bond
C-H	Carbon-Hydrogen bond
C-O	Carbon-Oxygen bond
C=C	Carbon-Carbon double bond
C=O	Carbon-Oxygen double bond
$cm-1$	Wavenumber, reciprocal of wavelength
dpt	Data file format for storing spectral data in ATR-FTIR analysis.
FTIR	Fourier-Transform Infrared Spectroscopy
HCA	Hierarchical Cluster Analysis
IR	Infrared
LDA	Linear Discriminant Analysis
MCT	Mercury Cadmium Telluride
O-H	Hydroxyl group (Oxygen-Hydrogen bond)
O=C=O	Carbon dioxide
PCA	Principal Component Analysis
PCs	Principal Components
Py-GC/MS	Pyrolysis Gas Chromatography-Mass Spectrometry
SD	Standard Deviation
TOF-SIMS	Time-of-Flight Secondary Ion Mass Spectrometry
UV	Ultraviolet
UV-Vis	Ultraviolet-Visible Spectroscopy
xi	Original intensity value
$\bar{x}$	Mean intensity value
Z	Z-score, standardized value
%	Percent
2D	Two-dimensional
3D	Three-dimensional

DISKRIMINASI FORENSIK KERTAS YANG MENUA SECARA SEMULA JADI DAN BUATAN MENGGUNAKAN SPEKTROSKOPI ATR-FTIR DAN TEKNIK KEMOMETRIK

ABSTRAK

Penuaan adalah proses semula jadi yang mengubah sifat kimia dan fizikal bahan, termasuk kertas, yang sering menyukarkan analisis dalam konteks forensik dan pemuliharaan. Kajian ini bertujuan untuk menganalisis sampel kertas yang telah mengalami penuaan di bawah keadaan berbeza (pengeringan udara, pengeringan dalam ketuhar, dan pendedahan UV) menggunakan spektroskopi ATR-FTIR dan analisis kemometrik bagi membezakan kertas yang telah menua secara semula jadi dengan kertas yang telah menua secara buatan. Sampel kertas baharu telah diwarnakan menggunakan empat agen pewarna (kopi, teh, cuka hitam, dan kicap gelap) dengan cara mencelup kapas ke dalam cecair agen pewarna dan menekapnya ke atas permukaan kertas bagi memastikan sapuan yang seragam. Sampel kertas yang telah diwarnakan ini kemudiannya tertakluk kepada tiga keadaan penuaan yang berbeza: pengeringan udara pada suhu bilik selama 30 hari, pengeringan dalam ketuhar pada suhu 60°C selama 1 hari, dan pendedahan UV selama kira-kira 4.3 jam bagi mensimulasikan penuaan semula jadi selama 30 hari. Kertas baharu yang tidak diwarnakan digunakan sebagai kawalan untuk perbandingan. Analisis ATR-FTIR menunjukkan perubahan spektrum yang ketara pada kawasan getaran O–H, C–H, dan C=O serta pada kawasan cap jari (1500–600 cm^{-1}). Kertas yang mengalami penuaan semula jadi menunjukkan tanda kimia unik yang terhasil daripada degradasi oksidatif beransur-ansur, membezakannya daripada sampel kertas penuaan buatan. Di antara kaedah penuaan buatan, pengeringan udara paling hampir menyerupai perubahan kimia yang diperhatikan dalam penuaan semula jadi, manakala pengeringan dalam ketuhar dan pendedahan UV menghasilkan perubahan yang lebih intensif tetapi berbeza. Teknik kemometrik, termasuk Analisis Kluster Hierarki (HCA) dan Analisis Komponen Utama (PCA), digunakan untuk mengklasifikasikan sampel dan mendedahkan perbezaan kimia asas. Hasil analisis menunjukkan bahawa, walaupun agen pewarna memperkenalkan ciri spektrum tambahan, kertas yang mengalami penuaan semula jadi sentiasa membentuk kluster yang jelas, manakala sampel penuaan buatan menunjukkan pertindihan yang ketara. Ini mencadangkan bahawa walaupun kaedah penuaan dipercepatkan boleh menghasilkan perubahan kimia yang serupa dengan penuaan semula jadi, ia tidak dapat sepenuhnya meniru proses degradasi kompleks yang berlaku selama beberapa dekad. Penemuan ini menekankan potensi ATR-FTIR dan teknik kemometrik sebagai alat analitik yang kukuh untuk analisis penuaan kertas, sambil menyerlahkan keterbatasan dalam protokol penuaan buatan ketika meniru degradasi semula jadi. Kajian ini menyediakan asas untuk memperhalusi kaedah penuaan buatan dan pendekatan analitikal, sekaligus menyumbang kepada peningkatan ketepatan pengesahan dokumen dan strategi pemeliharaan kertas berusia.

FORENSIC DISCRIMINATION OF NATURALLY AND ARTIFICIALLY AGED PAPER USING ATR-FTIR SPECTROSCOPY AND CHEMOMETRIC TECHNIQUES

ABSTRACT

Aging is a natural process that alters the chemical and physical properties of materials including papers, often complicating its analysis in forensic and preservation contexts. This study aimed to analyse paper samples aged under different conditions (air-drying, oven-drying, and UV exposure) using ATR-FTIR spectroscopy and chemometrics analysis to distinguish naturally aged paper from artificially aged samples. Fresh paper samples were stained using four staining agents (coffee, tea, black vinegar, and dark soy sauce) by dabbing the liquid staining agents onto the paper with cotton wool to ensure uniform application. The stained samples were then subjected to three different aging conditions: air-drying at room temperature for 30 days, oven-drying at 60°C for 1 day, and UV exposure for approximately 4.3 hours to simulate 30 days of natural aging. The fresh, unstained paper served as the baseline for comparison. ATR-FTIR analysis revealed significant spectral changes across the different groups, particularly in the regions corresponding to O–H, C–H, and C=O stretching vibrations, as well as in the fingerprint region (1500–600 cm^{-1}). Naturally aged paper exhibited unique chemical signatures resulting from prolonged, gradual oxidative degradation, distinguishing it from artificially aged samples. Among the artificial methods, air-drying most closely mimicked the chemical changes observed in natural aging, while oven-drying and UV exposure caused more intense however distinct alterations. Chemometrics methods, including Hierarchical cluster analysis (HCA) and Principal component analysis (PCA), were employed to classify the samples and identify the underlying chemical differences. The analyses demonstrated that, despite the additional

spectral features introduced by the staining agents, naturally aged paper consistently formed a distinct cluster, whereas artificially aged samples displayed considerable overlap. This suggests that while accelerated aging methods can induce chemical changes similar to those of natural aging, they do not fully replicate the complex degradation processes occurring over decades. These findings highlight the potential of ATR-FTIR spectroscopy and chemometric techniques as tools for forensic document analysis, while also highlighting the limitations of current artificial aging protocols in mimicking natural degradation. The study provides a foundation for refining artificial aging methods and improving the accuracy of document authentication and preservation strategies.

CHAPTER 1: INTRODUCTION

1.1 Background of Study

Paper is a commonplace form of documentation, which makes its study beneficial in forensic science (Zięba-Palus et al., 2020). However, little or great as the time may be, the paper will undergo degradation due to environmental and chemical factors, such as exposure to light, heat, moisture, and air pollutants (Kaldhone and Kumar, 2023; Silva et al., 2022). These factors cause changes at the molecular level where the structural integrity and appearance of the paper will be altered. For example, cellulose, the primary component of paper, undergoes oxidation and hydrolysis which cause its fibres to be weakened, and colouration alters (Bicchieri et al., 2006). While lignin, another paper component, is particularly sensitive to UV light and oxidizing agents undergo photodegradation that contributes to the yellowing commonly observed in aged or old paper (Kaldhone and Kumar, 2023).

The age and authenticity of documents play a crucial role in forensic investigations. While not always the primary focus, they are particularly valuable in cases involving fraud, forgery, and the preservation of historical documents (Kumar et al., 2024). Accurately determining a document's age can help verify its legitimacy, detect alterations, and provide critical evidence in legal and forensic contexts. Since both the processes of naturally aged paper and artificially aged paper involve physical methods, there is a challenge to observe uncontrolled similarities in more than just visual processes (Silva et al., 2022). Accurate identification of aging methods can help trace the provenance of documents and detect fraud. Silva et al. (2022), also described that paper degradation mechanisms, such as cellulose oxidation and the formation of carbonyl and carboxylic groups can be identified through infrared spectroscopy thus attribution of the modes of aging is possible. Similarly, Kumar and Rajasekaran's (2020) approaches emphasize the need for systematic investigations to detect artificial aging in fraudulent scenarios such as forged historical documents.

In forensic document examination, understanding the differences between naturally and artificially aged papers is important. The naturally aged paper includes papers that have become aged or weathered through heat exposure and time. On the other hand, artificially aged papers undergo weathering processes such as thermal treatment or exposure to UV in a controlled environment but within a short time period (Zięba-Palus et al., 2017). However, such artificial aging processes often leave detectable markers, such as irregular oxidation patterns and chemical traces from aging agents like coffee or soy sauce (Kumar and Rajasekaran, 2020; Silva et al., 2022).

Failing to differentiate between these two aging processes can lead to serious problems, including misinterpretation of a document's age or authenticity. For instance, in legal or historical disputes, documents that have been intentionally artificially aged may be mistakenly presented as older artifacts. Thus, infrared spectroscopy paired with HCA and PCA, has been established to effectively classify aging processes (Kaldhone and Kumar, 2023; Silva et al., 2022). The difference between these two processes can be defined through specific markers, so forensic scientists are expected to make accurate and reliable assessments of a document's authenticity.

Attenuated Total Reflectance Fourier Transform Infrared (ATR-FTIR) spectroscopy is a widely used non-destructive analytical technique in forensic science for characterizing organic materials (Farid et al., 2021). It enables the detection of specific vibrational modes of functional groups within a sample's chemical structure, effectively providing a molecular fingerprint of the sample (Zięba-Palus et al., 2020).

In the case of paper, ATR-FTIR spectroscopy may point out the changes in cellulose, lignin, and other components that occur due to aging processes.

Several of the key spectral markers such as shifts in hydroxyl (-OH) and carbonyl (C=O) groups, offer insights into the oxidation, hydrolysis, and various degradation pathways that the paper experiences over time (Zięba-Palus et al., 2017). ATR-FTIR is particularly beneficial in forensic analysis because it requires minimal sample preparation and can analyse small or delicate samples without causing significant damage (Wang et al., 2024). Forensic scientists use ATR-FTIR spectroscopy on aged paper to capture and compare chemical changes resulting from natural and artificial aging processes (Kaldhone and Kumar, 2023). These comparisons can help determine if the aging profile of a document aligns with its purported history, aiding in authentication claims.

Whereas chemometric techniques like HCA and PCA serve as valuable statistical tools for interpreting complex spectral data. The interdisciplinary field of chemometrics uses various statistical and mathematical techniques to extract meaningful information from massive chemical and biological data sets (Granato et al., 2018). In the context of forensic paper analysis, these methods allow scientists to classify and compare samples based on their chemical profiles, revealing subtle differences that might not be apparent through visual inspection or direct spectral comparison (Kaldhone and Kumar, 2023).

PCA helps to simplify multidimensional data by reducing it into a few principal components that account for the majority of variance in the data while HCA groups samples into clusters based on their similarities (Bartholomew, 2010). These techniques have also been applied to classify papers that have aged under different conditions. For example, PCA is effective in identifying markers of oxidation and photodegradation in cellulose, while HCA can group samples according to their aging methods, revealing distinct patterns associated with both natural and artificial aging processes (Kumar and Rajasekaran, 2020; Silva et al., 2024).

In this study, ATR-FTIR spectroscopy along with chemometric techniques was used to distinguish naturally aged from artificially aged paper samples. The samples were new white A4 (80 gsm) paper that was stained with coffee, tea, black vinegar, and dark soy sauce. Three different aging methods were used: air-drying (30 days at room temperature), oven-drying (1 day at 60°C), and UV exposure for approximately 4.3 hours to simulate one month of natural aging. ATR-FTIR spectra were analysed to observe the chemical changes induced by these aging methods, using chemometric techniques like HCA and PCA to discriminate and compare the effects of natural and artificial aging. This approach facilitated the identification of distinct aging markers and patterns associated with each treatment.

1.2 Problem Statement

In forensic science, the determination of a document's authenticity and age plays a crucial role in investigations involving fraud, forgery, and historical validation (Aniq et al., 2024). Documents may be artificially aged to mimic the appearance of naturally aged paper, posing challenges for forensic examiners tasked with distinguishing genuine documents from counterfeits. Naturally aged paper undergoes complex chemical degradation processes influenced by environmental factors such as light, temperature, humidity, and pollutants over extended periods. These changes primarily affect the cellulose and lignin structure, resulting in chemical signatures that reflect the natural aging process (Małachowska et al., 2020).

Artificial aging methods, such as oven drying and UV exposure, are widely used in laboratory settings to simulate the effects of natural aging within shorter timeframes. However, such methods often fail to replicate the full spectrum of chemical changes observed in naturally aged paper, leading to inaccuracies in comparative analysis. Furthermore, the presence of staining agents, often used in artificially aged papers, can obscure these chemical markers, complicating the differentiation process.

Existing analytical techniques, such as ATR-FTIR spectroscopy, provide a non-destructive means of analysing paper's chemical composition. While effective in detecting functional group changes associated with aging, ATR-FTIR analysis often generates complex datasets that are challenging to interpret. Chemometric techniques, such as HCA and PCA, offer powerful tools for extracting meaningful patterns from spectral data. However, these methods are not without limitations. When applied to simulated aging processes, PCA and HCA often fail to produce clear clustering of paper samples by the aging method (Kaldhone and Kumar, 2023). This issue is exacerbated when the chemical variability among artificial aging treatments (e.g., air drying, oven drying, UV exposure) is subtle, leading to overlapping spectral features and ambiguous classifications.

This study addresses these challenges by integrating naturally aged old paper samples into the analysis of simulated aging processes. It aims to evaluate the effectiveness of ATR-FTIR spectroscopy combined with chemometric techniques to discriminate between natural and artificial aging. By incorporating naturally aged old papers as a baseline, this study seeks to identify unique chemical markers that differentiate natural aging from artificial methods and to refine the clustering and classification of samples. These advancements are expected to enhance the reliability of forensic document examination and contribute to the development of standardized protocols for paper aging analysis.

1.3 Significance of the Study

This study is significant as it addresses a key issue in forensic document examination which is the ability to accurately discriminate between naturally and artificially aged paper. Document authenticity is vital in various legal, historical, and fraud investigations, where the age and origin of a document can determine its evidentiary value. By developing a robust methodology using ATR-FTIR spectroscopy and chemometric techniques, this research offers several contributions to the field of forensic science especially those involving document examination.

Firstly, the study deepens our understanding of chemical changes that occur in paper during natural and artificial aging. Natural aging happens due to prolonged exposure to environmental factors, whereas artificial aging methods like oven-drying and UV exposure aim to simulate these changes over shorter periods. By comparing these processes, this study helps to clarify whether artificial aging can reliably mimic natural aging, thus providing insights into the limitations and potential applications of accelerated aging techniques in forensic investigations.

Secondly, the utilization of chemometric tools, such as HCA and PCA, adds a quantitative dimension to the analysis of aging processes. These techniques enable the identification and discrimination of unique spectral markers associated with different aging methods, improving the accuracy and reliability of forensic evaluations.

Moreover, this study also introduces the use of naturally aged paper as a reference, providing a practical benchmark to validate the effectiveness of artificial aging methods.

Furthermore, the inclusion of staining agents, such as coffee, tea, black vinegar, and dark soy sauce, simulates real situations where documents are deliberately manipulated to appear older or aged. By considering these factors, the study ensures that the proposed methodology is relevant to real-world forensic cases, including those involving fraudulent documents.

Lastly, this study contributes to the development of standardized protocols for forensic paper analysis. A well-defined framework for distinguishing between natural and artificial aged strengthens the accuracy and credibility of forensic document examinations. Additionally, it also minimizes dependence on subjective evaluations and expensive or destructive techniques, making it more accessible to forensic laboratories with limited resources.

1.4 Research Objectives

1.4.1 General Objective

To analyse paper samples aged under different conditions (air-drying, oven-drying, and UV exposure) using spectroscopic and chemometrics analysis.

1.4.2 Specific Objectives

1) To simulate natural and artificial aging processes (air-drying, oven-drying, UV exposure) on paper samples.
2) To analyse the naturally aged and simulated aged paper using ATR-FTIR Spectroscopy.
3) To discriminate simulated aged papers with naturally aged papers using HCA.
4) To discriminate simulated aged papers with naturally aged paper using PCA.

CHAPTER 2: LITERATURE REVIEW

2.1 Paper Composition and Aging

Paper is mainly made up of cellulose, hemicellulose, and lignin, with cellulose being the key structural element (Area and Cheradame, 2011). Cellulose, a polysaccharide consisting of β-1,4-glucopyranose units, forms the core of paper's fibrous structure, providing strength and stability while hemicellulose is a branched polymer that contributes to the amorphous regions of the paper, affecting its flexibility and water retention properties (Area and Cheradame, 2011). While lignin is a complex aromatic polymer that serves as a binder within the paper matrix but it is more susceptible to degradation, especially under UV exposure (Kumar and Rajasekaran, 2020).

The aging process of paper involves both chemical and physical changes that are influenced by environmental factors and the paper's composition. Natural aging occurs gradually over time due to exposure to light, air, humidity, and temperature fluctuations (Kaldhone and Kumar, 2023). The main chemical reactions involved are the oxidation and hydrolysis of cellulose and lignin, which lead to discoloration, brittleness, and the weakening of fibres or loss of mechanical integrity (Kaldhone and Kumar, 2023; Silva et al., 2022). In contrast, artificial aging uses controlled conditions like thermal exposure or UV irradiation to mimic the effects of natural aging in a period. Thermal aging accelerates oxidation and hydrolysis through elevated temperatures, while UV exposure causes photodegradation, resulting in the breakdown of cellulose and lignin into smaller, oxidized fragments (Silva et al., 2024).

Environmental factors significantly impact paper degradation. High humidity can enhance hydrolysis, particularly in acidic paper, by breaking down glycosidic bonds leading to the formation of carboxylic acids that further degrade cellulose (Małachowska et al., 2021). While temperature increases the rate of oxidation and weakens cellulose. UV light, a potent degrader of lignin, induces photodegradation that forms chromophores that cause yellowing and embrittlement (Kaldhone and Kumar, 2023). These factors often work together, causing the gradual deterioration of paper over time (Kumar and Rajasekaran, 2020; Silva et al., 2024).

2.2 Staining Agents and Their Effects on Paper

Staining agents like coffee, tea, black vinegar, and dark soy sauce, are commonly used in forensic simulations to mimic the appearance of artificially aged documents. By introducing both organic and inorganic substances, these agents change the composition and appearance of the paper matrix through chemical interactions. For example, coffee contains tannins, polyphenols, and organic acids that can speed up oxidative degradation can cause paper to turn brown (Kumar and Rajasekaran, 2020; LIczbiński and Bukowska, 2022). Similarly, tea contains a lot of catechins and tannins that can form complexes with cellulose and lignin to cause discolouration and structural changes (Kumar and Rajasekaran, 2020; LIczbiński and Bukowska, 2022). While black vinegar and dark soy sauce, on the other hand, introduce acids, sugars, and pigments to the paper matrix. The acidic components may hydrolyse glycosidic bonds in cellulose, weakening the paper's structure, while the pigments will seep into the fibres, giving the paper a mottled or uneven look (Kaldhone and Kumar, 2023; Kumar and Rajasekaran, 2020). During thermal or UV aging, the high sugar content of soy sauce can caramelize and then leave a characteristic brown residue that mimics the natural sign of aging (Kaldhone and Kumar, 2023; Kumar and Rajasekaran, 2020; Silva et al., 2024).

The chemical interactions between these staining agents and paper components during aging vary based on the type of agent and the aging conditions (Kumar and

Rajasekaran, 2020). For example, the acidic compounds induce hydrolysis, whereas the polyphenolic compounds in coffee and tea accelerate oxidative degradation. These interactions can result in spectral markers that can be detected using ATR-FTIR spectroscopy, like alterations in the O-H stretching region (3200–3600 cm^{-1}) and enhanced carbonyl group peaks (1700–1750 cm^{-1}) (Kaldhone and Kumar, 2023; Silva et al., 2018). It is crucial to comprehend these effects for forensic applications because they shed light on the processes via which artificial aging alters paper and helps distinguish between naturally aged and artificially aged documents (Kaldhone and Kumar, 2023; Kumar and Rajasekaran, 2020; Silva et al., 2024).

2.3 Analytical Techniques for Paper Aging Studies

Attenuated Total Reflectance Fourier Transform Infrared (ATR-FTIR) spectroscopy is a non-destructive technique widely used to analyse chemical changes in paper (Farid et al., 2021; Lee et al., 2016). By measuring the absorption of infrared light at specific wavelengths, ATR-FTIR provides detailed information about functional group changes and molecular degradation over time. This makes it particularly useful for investigating the degradation of cellulose, hemicellulose, and lignin, which are the main components in paper (Kapoor et al., 2021; Xia et al., 2019).

For forensic document examination, ATR-FTIR offers several advantages which are non-destructiveness, molecular sensitivity, and also ease of use. It allows analysis to be repeated without compromising the sample's integrity (Alkhuder, 2022). Moreover, it can also detect subtle chemical changes in functional groups, such as hydroxyl (-OH) and carbonyl (C=O), which are the key markers of aging (Kapoor et al., 2021). ATR-FTIR also offers minimal sample preparation as compared to other spectroscopic methods (Alkhuder, 2022). Hence, ATR-FTIR has proven invaluable in identifying spectral changes associated with natural and artificial aging, as well as in detecting the presence of chemical agents used for forgery.

Recent studies highlight the utility of ATR-FTIR combined with chemometric techniques to differentiate between naturally aged and artificially aged paper (Kaldhone and Kumar, 2023; Kumar and Rajasekaran, 2020). This combination enhances spectral interpretation and classification, providing robust evidence in forensic investigations. The degradation of cellulose and lignin during aging processes produces distinct spectral changes detectable by ATR-FTIR. For cellulose, hydroxyl (-OH) groups show broad peaks in the 3200–3600 cm^{-1} region that reduce with aging as hydrogen bonding weakens and cellulose crystallinity increases. Similarly, oxidation products such as aldehydes and ketones that are produced during degradation are indicated by peaks in the 1700–1750 cm^{-1} range for carbonyl (C=O) groups (Zięba-Palus et al., 2017). For lignin, the aromatic rings, which are visible in the 1500–1600 cm^{-1} region, may decrease as a result of photodegradation, which forms the formation of chromophoric groups like quinones that contribute to yellowing, while ether bonds (C–O) in the 1000–1200 cm^{-1} region show the breakdown of lignin's structural network (Małachowska et al., 2020). The fingerprint region (1000–1200 cm^{-1}), which is crucial for glycosidic bonds and other cellulose-lignin interactions, is where these alterations are frequently seen as shifts or intensities (Małachowska et al., 2021).

Although ATR-FTIR spectroscopy is quite good at forensic paper analysis, other spectroscopic methods have also been explored for possible uses. For example, UV-Vis spectroscopy can be used to track the creation of chromophoric groups during the breakdown of lignin, but it is less effective for in-depth study due to its lack of molecular specificity (Kapoor et al., 2021; Silva et al., 2024). Raman spectroscopy, on the other hand, provides excellent resolution and complements ATR-FTIR by providing further structural information, however, its use may be restricted by fluorescence or pigmentation in certain samples (Zięba-Palus et al., 2017). Furthermore, although Mass spectrometry like TOF-SIMS and Py-GC/MS, provides detailed molecular breakdown of paper components, its inherent destructiveness makes it less suitable for forensic document analysis, where non-destructive techniques are preferred. Therefore, ATR-FTIR remains the preferred technique for studying paper degradations as well as forensic applications due to its balance of non-destructive analysis, molecular sensitivity, and simplicity of sample preparation (Kaldhone and Kumar, 2023; Silva et al., 2022; Xia et al., 2020).

2.4 Chemometric Techniques in Forensic Science

Chemometrics, such as multivariate analysis and other statistical techniques, are increasingly being used while gathering forensic data, according to forensic literature (European Network of Forensic Science Institutes Drugs Working Group, 2021). Principal Component Analysis (PCA) and Hierarchical Cluster Analysis (HCA) are widely (and, at times, improperly) applied as "unsupervised classification" methods to evaluate associations between chemical characteristics and the aging processes of forensic documents. The goal of unsupervised learning is to uncover patterns within datasets without relying on predefined labels or outcomes (Sauzier et al., 2021). These algorithms work independently to infer patterns with minimal human guidance. While unsupervised methods are not directly applicable to classification or regression tasks, they are particularly well-suited for exploring the underlying structure of data (Sauzier et al., 2021). Additionally, some unsupervised approaches allow new samples to be projected onto an existing dataset, which can be valuable for comparison (Sauzier et al., 2021). However, this process should not be regarded as true classification, as it does not presume the existence of distinct classes (Sauzier et al., 2021). These techniques are often used to differentiate naturally aged paper from artificially aged or manipulated documents based on their chemical and physical properties (Granato et al., 2018).

HCA encompasses a set of algorithms that group objects (samples) based on their relative similarity (Sauzier et al., 2021). The choice of algorithm and parameters depends on the dataset's characteristics and the analysis's objectives. A widely used method forms clusters by iteratively merging objects based on separation distance. This agglomerative, "bottom-up" approach begins with each object as a single cluster and progressively groups the closest pairs until a single cluster remains (Sauzier et al., 2021). The clustering results depend on the selected distance metric (used to measure similarity) and linkage criteria (determining how clusters are connected). Different metrics and linkage methods can produce varying cluster patterns. HCA results are typically visualized as a dendrogram, which provides a clear representation of the hierarchical relationships and relative similarities between objects, often reflected in their original spectral data (Sauzier et al., 2021).

In forensic applications, HCA is used to classify paper samples by their spectral profiles, revealing clusters that correspond to different aging methods or conditions. For example, naturally aged papers tend to form distinct clusters due to their unique chemical signatures, such as advanced oxidation patterns developed over extended periods. In contrast, artificially aged papers typically cluster separately, reflecting the more uniform but distinct degradation patterns induced by controlled processes such as oven drying or UV exposure. This method complements PCA by providing a clear hierarchical structure for classification (Farid et al., 2021; Xia et al., 2020, 2021) .

PCA analysing complex multivariate data, such as spectra or chromatograms, can be challenging, especially when identifying features responsible for similarities or differences between samples (Sauzier et al., 2021). In HCA, such insights often require visual comparisons of the original data, which may be impractical for large datasets. PCA simplifies this process by reducing data dimensionality (Sauzier et al., 2021). It identifies orthogonal sources of variation, or principal components (PCs), which are linear combinations of the original variables. Each successive PC captures the maximum remaining variance, allowing samples to be represented using PC scores instead of the original variables (Sauzier et al., 2021). PCA visualizations, such as score plots, reveal patterns and relationships not immediately apparent in raw data, while scree plots indicate the PCs containing the most meaningful information. However, as with HCA, a complete understanding of the trends often requires revisiting the original dataset (Sauzier et al., 2021). In forensic paper analysis, PCA can identify patterns and groupings related to aging processes, such as natural versus artificial aging. The first few principal components often explain the majority of the variance, making it easier to interpret large datasets (Kaldhone and Kumar, 2023; Zięba-Palus et al., 2020).

Chemometric techniques, including PCA and HCA, play an integral role in forensic science, particularly in the analysis of aged documents. These methods enhance the ability to discriminate aging processes by identifying unique spectral markers associated with each type of aging. They also enable the classification and validation of artificially aged paper simulations by comparing them to naturally aged paper, ensuring the reliability of accelerated aging methods (Kaldhone and Kumar, 2023; Xia et al., 2020; Zięba-Palus et al., 2020). Moreover, chemometric techniques significantly improve the forensic analysis of questioned documents by objectively identifying inconsistencies in aging patterns, which can support cases of fraud detection (Farid et al., 2021). These methodologies collectively contribute to the development of standardized approaches for forensic document examination, addressing challenges in distinguishing naturally and artificially aged papers with greater accuracy and reliability.

2.5 Accelerated Aging in Forensic Research

Accelerated aging techniques, such as thermal and UV treatments, are commonly employed to replicate the chemical and physical changes observed in naturally aged documents over a shorter time (Silva et al., 2022). Thermal aging involves exposing paper samples to elevated temperatures, typically between 60°C and 105°C, to accelerate oxidation and hydrolysis processes (Małachowska et al., 2021; Silva et al., 2022). This method mimics the effects of long-term exposure to ambient heat, leading to the breakdown of cellulose and lignin. UV treatment, on the other hand, involves irradiating samples with ultraviolet light to simulate photodegradation (Silva et al., 2022). UV exposure primarily affects lignin, producing chromophores that cause yellowing and weakening of the paper structure (Małachowska et al., 2020; Strlič and Kolar, 2004). Both methods aim to reproduce the effects of natural environmental factors, such as sunlight and temperature, on document aging.

Accelerated aging methods are designed to replicate the chemical transformations occurring in naturally aged papers. For instance, both thermal and natural aging lead to oxidation, resulting in the formation of carbonyl (C=O) groups, while UV exposure mimics the effects of prolonged sunlight, degrading lignin and cellulose (Silva et al., 2022). However, accelerated aging processes often produce changes at a faster rate and with greater uniformity compared to the slower, more complex natural aging process. While there is some correlation between the chemical markers of accelerated and natural aging, discrepancies may arise due to the controlled nature of accelerated methods that do not fully replicate environmental variability.

Despite their utility, accelerated aging techniques have limitations. One major challenge is the inability to replicate the full complexity of natural aging, which involves a combination of environmental factors such as humidity, air pollutants, and fluctuating temperatures (Hajji et al., 2016). Additionally, the uniform conditions in accelerated aging studies may produce spectral markers that differ from those of naturally aged documents, complicating direct comparisons. These methods also risk over-simplifying the aging process, leading to results that may not fully align with real-world scenarios, thereby affecting the validity of forensic conclusions.

2.6 Forensic Applications of Aging Analysis

The ability to distinguish between naturally aged and artificially aged documents is crucial in forensic investigations, particularly in cases of fraud or forgery. Forgers often attempt to artificially age documents using methods such as staining or thermal exposure to mimic the appearance of authentic historical records. Identifying whether a document's aging is natural or artificial helps establish its provenance, authenticity, and evidentiary value. Forensic examiners rely on analytical techniques like ATR-FTIR spectroscopy and chemometric tools to detect discrepancies in aging patterns and validate the integrity of questioned documents.

Several forensic cases demonstrate the importance of aging analysis in document examination. For example, studies by (Kaldhone and Kumar (2023) have shown that ATR-FTIR spectroscopy, combined with chemometric analysis, can successfully differentiate between naturally and artificially aged papers by detecting unique spectral markers. In one case, artificial aging through UV exposure was identified by the presence of specific degradation products not typically found in naturally aged samples (Silva et al., 2022). Such findings provide critical evidence in cases involving forged contracts, altered wills, or historical document disputes.

2.7 Gaps in Current Research

While ATR-FTIR spectroscopy and chemometric techniques like PCA and HCA have shown promise in forensic document examination, few studies have comprehensively integrated these methods for analysing paper aging. Existing research often focuses on one aspect, such as spectral analysis or clustering, without addressing their combined potential to enhance forensic accuracy. This gap limits the development of reliable methodologies for distinguishing natural and artificial aging.

There is a pressing need for a standardized framework to simulate and compare natural and artificial aging processes. Current accelerated aging methods lack consistency in experimental conditions, such as temperature, UV intensity, and duration, leading to variability in results. A standardized methodology would ensure reproducibility and enable more accurate comparisons between naturally aged and artificially aged documents. Such a framework would also provide forensic practitioners with robust tools for assessing document authenticity in legal and investigative contexts.

CHAPTER 3: METHODOLOGY

3.1 Research Design

This study aimed to analyse and discriminate paper samples subjected to several simulated aging processes and natural aging using ATR-FTIR spectroscopy and chemometric techniques. Paper samples were stained by dabbing using cotton wools with various staining agents such as coffee, tea, dark soy sauce, and black vinegar to simulate the appearance of naturally aged papers. Then these paper samples were artificially aged using three methods which are air-drying, oven-drying, and UV-exposure. After that, these were compared to unstained and untreated new papers and naturally aged papers. For consistency, all the paper samples except the naturally aged paper samples are sourced from a standard manufacturer and are from the same batch of paper. ATR-FTIR spectroscopy was used to analyse seven random spots on each sample, collecting detailed spectral data indicative of aging-related chemical changes.

Chemometric techniques were employed sequentially to analyse the spectral data. HCA was performed first to group samples based on spectral similarities, followed by PCA to visualize patterns and identify key factors differentiating the aging processes. Figure 3.1 shows the study flowchart.

3.1.1 Study Scope

The study focused on analysing paper samples artificially aged using controlled laboratory conditions to simulate 30 days of natural aging. Fresh new papers and naturally aged papers were included for comparison. A total of 29 samples were analysed: 15 artificially aged (one sample per aging method per staining agent), 7 fresh new papers, and 7 naturally aged papers.

3.2 Research Tool and Methodology

3.2.1 Instruments Used in Staining and Aging the Paper Samples

Table 3.1: List of Instruments Used

Instruments	**Brand/Source**
ATR-FTIR Spectrometer	Bruker Corporation, USA
UV Lamp	Industry-standard
Laboratory Oven	Memmert, Germany
Sample Holders and Glassware	PYREX, England

Table 3.1 shows the instruments used in this study. The methodology for this study involved the preparation and analysis of paper samples using various aging treatments and chemometric analysis to discriminate between naturally and artificially aged samples. New A4 office papers (80 gsm) were stained using coffee, tea, black vinegar, and dark soy sauce to simulate realistic aging conditions. To ensure consistency in the staining process, a standardized method was employed. Coffee and tea stains were prepared by two coffee bags and two tea bags, respectively, in 100 mL of hot distilled water (approximately 90°C) for 5 minutes to extract the staining agents. Black vinegar and dark soy sauce were used undiluted. Each staining agent was applied to the paper samples using a clean, lint-free cotton wool pad, ensuring uniform coverage without over-saturation. The stained papers were air-dried at room temperature (25 °C) for 24 hours to allow complete absorption of the staining agents. Visual inspection was conducted to ensure even staining, and any samples with uneven distribution or excessive residue were excluded. This standardized approach minimized variability and ensured reproducibility in the staining process.

The samples were then subjected to three distinct aging methods. Artificially aged papers were processed as follows where air-dried paper samples were kept in the analytical laboratory for 30 days at room temperature (25 °C). Whereas oven-dried paper samples were heated for 1 day at 60°C to simulate 30 days of natural aging, UV-exposed paper samples were exposed to UV light for ~4.3 hours, equivalent to 30 days of natural aging. Naturally aged paper samples were sourced from old documents (the 1940s), and fresh unstained and untreated paper samples were analysed without additional treatment. These naturally aged old paper samples represented a range of storage environments and timeframes, serving as a comparative baseline. All samples were stored in a controlled environment to minimize contamination or external interference.

3.2.2 ATR-FTIR Spectroscopic Analysis

ATR-FTIR spectra were acquired using a Bruker Lumos IR spectrometer. This Lumos IR spectrometer is a kind of FTIR microscope that measures reflection, transmission, and attenuated total reflection (ATR). It was linked with the OPUS 7.5 software and the computer, which work together to obtain the IR spectra of the paper samples.

The FTIR spectrometer was filled with liquid nitrogen that is approximately -200 °C and allowed to stand for approximately 30 minutes before analysis began. During the usage of the spectrometer, thermal processes heated the mercury cadmium telluride (MCT) detector that was used in it. Consequently, liquid nitrogen needs to be used to regulate heat, cool down the detector, and to lower down the instrument noise while using it. Then, the FTIR spectrometer was calibrated using the Joystick software before scanning the paper samples. The controller and stage of the device were calibrated using this joystick program to ensure the exactness of the sample measurement.

The spectrometer was operated in the mid-infrared range (4000 to 600 cm^{-1}), with a resolution of 4 cm^{-1}, and each sample underwent 64 scans per spot to obtain high-quality and reliable data. Furthermore, the pressure of the ATR scan was set at low-pressure mode as it is the optimal choice for analysing paper samples. It balances between good contact with ATR crystal while preserving the sample's physical and chemical integrity, ensuring reliable and reproducible spectral data across all paper types. Additionally, the samples were positioned on the spectrometer's stage, and a weight was placed around the scanning area to prevent any wind obstacles that would interfere with the spectra's outcome. Spectral data were collected at seven random spots on each sample to enhance representativeness and the background was measured prior to each scan to ensure accuracy. After scanning, the collected data were saved in .dpt format for subsequent chemometric analysis.

3.2.3 Chemometric Analysis

Every spectrum data set stored in dpt format was imported to Microsoft Excel. From the raw spectral data, only 1500 cm^{-1} to 600 cm^{-1} fingerprint regions were selected for this data analysis, a process known as the feature selection technique. The fingerprint region (1500–600 cm^{-1}) of the ATR-FTIR spectrum was selected for analysis due to its richness in unique spectral features, such as C–O, C–C, and aromatic ring vibrations, which can give characterization by identifying molecular changes in cellulose and lignin during aging. This feature selection serves as a preprocessing step for data reduction, minimizing the dataset size while preserving valuable information essential for HCA and PCA. Limiting the analysis to this region also addresses Minitab's data handling limitations, as it cannot efficiently process the entire spectrum (e.g., 4000 cm^{-1} to 600 cm^{-1}, which could include thousands of wavenumbers). By focusing on the fingerprint region, the study ensures computational efficiency and highlights key spectral markers relevant to distinguishing between different paper samples.

HCA was performed first to group samples based on their spectral similarities. This provided an initial classification of samples by aging method and staining agent. PCA was then applied to visualize clustering patterns among the aging methods and identify the key factors contributing to spectral differences among the samples. Graphs were plotted as absorbance versus wavenumber, and dendrograms and PCA scatter plots were generated to interpret the clustering and separation of the samples.

Before performing HCA and PCA using Minitab Version 17, the raw IR spectral data of the samples underwent preprocessing to ensure consistency and improve the reliability of the results. Preprocessing involves auto-scaling also known as standardization to transform the raw data into standardized values. This step is essential for addressing variations in the data that may arise due to multiple factors, such as sample preparation, instrument sensitivity, or baseline shifts, which could otherwise obscure meaningful patterns in the dataset.

Auto-scaling adjusts the data to have a mean of 0 and a standard deviation of 1 for each variable (e.g., spectral intensity at a specific wavenumber). The formula for auto-scaling is:

$$Z = \frac{x_i - \bar{x}}{SD}$$

Where:

- Z is the standardized value.
- x_i is the original intensity value for a given variable (e.g., absorbance at a specific wavenumber).
- $\bar{x}$ is the mean intensity of the variable across all samples.
- SD is the standard deviation of the variable across all samples.

This formula ensures that each variable contributes equally to the analysis by removing differences in scale or magnitude. The process adjusts for "run-to-run" variations, such as differences in instrument performance or slight inconsistencies in sample preparation, thereby reducing bias and enhancing the comparability of the data.

Auto-scaling is necessary to remove bias as variations caused by experimental factors, such as instrument calibration or environmental conditions can be minimized. Moreover, it enables equal weighting by preventing high-intensity peaks from disproportionately influencing the results as each variable contributes equally to the analysis. Additionally, by standardizing the data, patterns in the PCA or clusters in the HCA become more representative of true chemical differences rather than artifacts of data magnitude.

After auto-scaling, the standardized data was used for HCA and PCA instead of the raw spectral data. This ensures that the clustering in HCA and the variance captured by PCA reflect genuine chemical differences between the paper samples rather than variations introduced by data inconsistencies. Using pre-processed data improves the clarity and reliability of the analysis, allowing for a more accurate comparison of spectral features across samples.

3.2.4 Study Flowchart

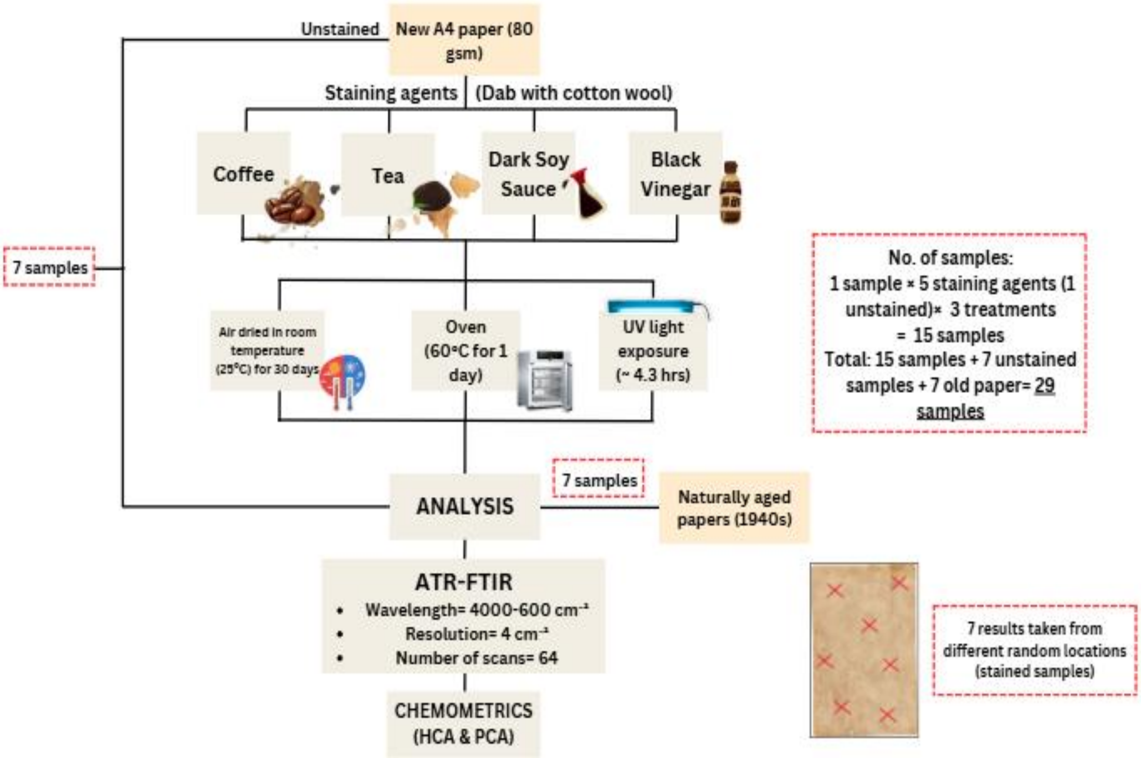

Figure 3.1: Study Flowchart